SPRINGER
REFERENCE

Franz J. Hock
Michael R. Gralinski
Editors

Drug Discovery and Evaluation: Methods in Clinical Pharmacology

Second Edition

Springer

Drug Discovery and Evaluation: Methods in Clinical Pharmacology

Franz J. Hock • Michael R. Gralinski
Editors

Drug Discovery and Evaluation: Methods in Clinical Pharmacology

Second Edition

With 249 Figures and 161 Tables

Springer

Editors
Franz J. Hock
CorDynamics
Dieburg, Germany

Michael R. Gralinski
CorDynamics
Chicago, IL, USA

ISBN 978-3-319-68863-3 ISBN 978-3-319-68864-0 (eBook)
ISBN 978-3-319-68865-7 (print and electronic bundle)
https://doi.org/10.1007/978-3-319-68864-0

General Introduction

Drug discovery and evaluation is a multidisciplinary process. The discovery of a new drug starts with experiments in isolated organs or in biochemical assays, for example, in vitro receptor binding's studies. New chemical compounds have to be compared with known drugs used in specific therapies. Positive results have to be confirmed in various animal tests. The therapeutic advances may be higher potency, fewer side effects, longer activity, or a new mode of action. Many methods are described in the literature and are reviewed in *Drug Discovery and Evaluation: Pharmacological Assays*, the first book of this series.

The strategy of drug development has changed in recent years. Instead of sequential studies in toxicology and pharmacokinetics, the parallel involvement of various disciplines has been preferred. Exposure to the body is investigated by pharmacokinetic studies on absorption, distribution, and metabolism at an early stage of development and contributes to the selection of drugs. The term safety pharmacology, formerly general pharmacology, has been coined to describe a specific issue in addition to traditional toxicology tests.

These studies are reviewed in the second book of the series as *Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays*.

Clinical pharmacology and clinical pharmacokinetics belong together. There is no pharmacodynamics without pharmacokinetics and vice versa. We, therefore, combined both disciplines in the third book of this series as *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*, with the aim of demonstrating the mutual dependency to the reader.

An important objective of clinical pharmacology is the early and ongoing assessment of the **safety and tolerability** of a new drug. This is done by assessing the type, frequency, and severity of side effects; assessing in which patient population these side effects may occur at which dose or exposure; for what duration; and whether these side effects are reversible. The importance of an adequate selection of animal models, assessing the significance of the preclinical data obtained in the first-in-man study, has recently been shown quite dramatically.

The first dose step in the first-in-man study with a humanized monoclonal antibody induced a cytokine release syndrome in all actively treated healthy volunteers, all of whom suffered life-threatening, acute shock and subsequent multiorgan failure. Obviously, these severe and serious adverse results were

not predicted by the animal studies conducted prior to human studies. Because of this incident, the regulators worldwide changed several processes so that this mishap further could be prevented.

Pharmacogenomics has an interesting input to drug development. Genomic information should enable the pharmaceutical industry to target specific patient populations that are more likely to respond to the drug therapy or to avoid individuals who are likely to develop specific adverse events in clinical studies. In this volume, the possibilities of pharmacogenomics-guided drug development are discussed. Another new and promising development is **personalized medicine**. These new areas are discussed in this edition.

This second edition of *Drug Discovery and Evaluation: Methods in Clinical Pharmacology* is completely new organized and extended. It contains besides the former chapters several new ones, such as Clinical Studies in Infants and Geriatric Population, Traditional Chinese Medicine, Space Pharmacology, Nanotechnology in Medicine, etc.

Franz J. Hock
Michael R. Gralinski

Acknowledgment

We would like to express our gratitude and our sincere thanks to all authors who contributed their knowledge to this book. Furthermore, we personally would like to thank the founder of this series H. Gerhard Vogel, who passed away in 2011. He was the Editor-in-Chief of the first edition and all of the *Drug Discovery and Evaluation* titles at Springer consisting of *Pharmacological Assays*, *Safety and Pharmacokinetic Assays*, and *Methods in Clinical Pharmacology.*

Contents

Volume 1

About the Editors

Franz J. Hock Since retiring from Aventis in 2002, Dr. Hock has leveraged his experience as a freelance consultant specializing in safety pharmacology. Dr. Hock was a research scientist at HOECHST, Hoechst Marion-Roussel, and Aventis from 1976 to 2002. He initially worked on methods in general pharmacology and nephrology, before becoming Head of Laboratory and devoting to pharmacological methods for drugs influencing memory and learning. He was ultimately Head of Laboratory for General/Safety Pharmacology at the Frankfurt site of Aventis Pharma Deutschland GmbH.

Dr. Hock received his M.Sc. in Neurobiology from the Technical University Darmstadt and his D.Sc. in Zoology from the University Kassel, Department of Biology, Institute of Neuroethology and Biocybernetics.

He received the degree of Fachpharmakologe DGPT ("Certified Expert Pharmacology") in 1981. In 1983, Dr. Hock spent a sabbatical year at the University of California, Irvine, at the Center for the Neurobiology of Learning and Memory (Director Prof. Dr. James L. McGaugh).

He lectured for several years to student of Biology at the University of Kassel and the Technical University Darmstadt. He has published over 100 original papers on methods in pharmacology and on new compounds. Furthermore, he held 28 patent applications to protect or broaden the application of lead structures.

He is currently a member of the Task Force General/Safety Pharmacology German/Swiss Pharmaceutical Companies and is also member of several national and international scientific societies. Dr. Hock is founding member of "Safety Pharmacology Society," "Neurowissenschaftliche Gesellschaft e.V.," and "European Behavioural Pharmacology Society." For several years, he is serving as a member of the Program Committee of the Safety Pharmacology Society. Dr. Hock is member of several domestic and international scientific societies.

Michael R. Gralinski Dr. Gralinski is an internationally recognized authority and advisor in cardiovascular pharmacology. He has founded multiple contract research and consulting companies and is currently the Chief Executive Officer of CorDynamics, a position he has held since 2002. He has more than 25 years of experience in pharmaceutical research and development with deep expertise in cardiovascular pharmacology and toxicology experimental design, data analysis, and interpretation.

Prior to CorDynamics, Dr. Gralinski garnered extensive management experience working for several pharmaceutical organizations including Warner-Lambert/Parke-Davis, G.D. Searle, Pharmacia, and Pfizer. His past responsibilities included directing product safety and toxicology operations, leading resolution efforts for cardiovascular issues involving discovery and development, and marketed compounds including Rezulin, Celebrex, and other COX-2 inhibitors as well as cultivating relationships with international colleagues throughout small and large biopharma. He was involved in the genesis of the Safety Pharmacology Society and was elected to multiple positions on its Board of Directors including Treasurer, Vice-President, and President.

Dr. Gralinski earned a Bachelor of Science degree in Pharmacology and Toxicology from the University of Wisconsin-Madison and a Doctor of Philosophy in Cardiovascular Pharmacology from the University of Michigan.

Contributors

Khaled Abo-EL-Sooud Pharmacology Department, Faculty of Veterinary Medicine, Cairo University, Giza, Egypt

James Adjaye Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Karel Allegaert Intensive Care and Department of Pediatric Surgery, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands

Department of Development and Regeneration, KU Leuven, Leuven, Belgium

Hesham S. Al-Sallami School of Pharmacy, University of Otago, Dunedin, New Zealand

Jens Atzrodt Integrated Drug Discovery, Medicinal Chemistry, Isotope Chemistry and Metabolite Synthesis Departments (ICMS), Sanofi-Aventis Deutschland GmbH, Frankfurt a. M, Germany

Georgi Banishki University of Aberdeen, Aberdeen, UK

Carolin Bendadani Federal Office of Consumer Protection and Food Safety, Berlin, Germany

Smita Bhuyan Department of Microbial drugs, Helmholtz Center for Infection Research, Braunschweig, Germany

Liora Bik Dermatology, Erasmus University Medical Center, Rotterdam, The Netherlands

Jaine K. Blayney Stratified Medicine Group, Centre for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast, UK

Henning H. Blume Frankfurt am Main, Germany

SocraTec C&S GmbH, Oberursel, Germany

Gerd Bode Institute for Pharmacology and Toxicology, University Medical Center, University of Goettingen, Goettingen, Germany

Stacey Boyer Clinical Research Operations, Cavion, Charlottesville, VA, USA

Evelyn Breitweg-Lehmann Federal Office of Consumer Protection and Food Safety, Berlin, Germany

Michael E. Burczynski Department of Translational Medicine, Global R&D, Teva Pharmaceuticals, Frazer, PA, USA

Ronald Burnette University of Wisconsin School of Pharmacy, Madison, WI, USA

Sandra Cvijic Department of Pharmaceutical Technology and Cosmetology, University of Belgrade-Faculty of Pharmacy, Belgrade, Serbia

Anibh M. Das Department of Paediatrics, Hannover Medical School, Hannover, Germany

Clayton Dehn Highpoint Clinical Trials Center, High Point, NC, USA

Volker Derdau Integrated Drug Discovery, Medicinal Chemistry, Isotope Chemistry and Metabolite Synthesis Departments (ICMS), Sanofi-Aventis Deutschland GmbH, Frankfurt a. M, Germany

Hartmut Derendorf Department of Pharmaceutics College of Pharmacy, University of Florida, Gainesville, FL, USA

Robert Jan Doll Centre for Human Drug Research, Leiden, The Netherlands

Tanja Eisenblaetter München, Germany

Sebastian Felgner Department of Molecular Immunology, Helmholtz Center for Infection Research, Braunschweig, Germany

Anthony G. Fenech Department of Clinical Pharmacology and Therapeutics, University of Malta, Msida, Malta

Ingo Fietze Interdisciplinary Sleep Medicine Center, Charitécentrum für Herz- Kresilauf- und Gefäßmedizin CCM11, Charité – Universitätsmedizin Berlin, Berlin, Germany

Felix Frueh Opus Three LLC Life Sciences, Pescadero, CA, USA

Matthias Germer Corporate R&D, Biotest AG, Dreieich, Germany

Federico Goodsaid Opus Three LLC Life Sciences, Pescadero, CA, USA

Roman Görtelmeyer Scientific Consulting, Frankfurt am Main, Germany

Biological and Clinical Psychology, University of Mannheim, Mannheim, Germany

Nina Graffmann Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Nicolas Grandchamp Biosource / GEG Tech, Paris, France

Geert Jan Groeneveld Centre for Human Drug Research, Leiden, The Netherlands

Bruno Hagenbuch Department of Pharmacology, Toxicology and Therapeutics, The University of Kansas Medical Center, Kansas City, KS, USA

Teodora Handjieva-Darlenska Department of pharmacology and toxicology, Medical Faculty, Medical University of Sofia, Sofia, Bulgaria

Chihiro Hasegawa School of Pharmacy, University of Otago, Dunedin, New Zealand

Translational Medicine Center, Ono Pharmaceutical Co., Ltd., Osaka, Japan

Justin Hay Centre for Human Drug Research, Leiden, The Netherlands

Rainer Heissing Heissing Pharmacovigilance Consulting GmbH, Partenheim, Germany

Lynne Hughes Neurology Center of Excellence, IQVIA, Reading, Berkshire, UK

Gulam Mohammed Husain Pharmacology Research Laboratory, Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India

Paul Hutson University of Wisconsin School of Pharmacy, Madison, WI, USA

Sabine Illsinger Department of Neuropaediatrics, Children's Hospital Oldenburg, Oldenburg, Germany

Rüdiger Kaspera Clinical Pharmacology and Exploratory Development, Astellas Pharma Europe B.V., Leiden, The Netherlands

Munawwar Husain Kazmi Department of Ilmul Advia (Unani Pharmacology), Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India

Mohammad Ahmed Khan Pharmacology Research Laboratory, Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India

Dino Kocijancic Department of Molecular Immunology, Helmholtz Center for Infection Research, Braunschweig, Germany

Andreas Kovar Translational Medicine and Early Development, Sanofi, Frankfurt, Germany

Deborah Lee CNS-TAU, Takeda Pharmaceutical International Co., Cambridge, MA, USA

Björn Lemmer Institute of Experimental and Clinical Pharmacology and Toxicology, Medical Faculty, Mannheim Ruprecht-Karls-University of Heidelberg, Mannheim, Germany

Birgit Liebscher Federal Office of Consumer Protection and Food Safety, Berlin, Germany

Claudia Loewe Integrated Drug Discovery, Medicinal Chemistry, Isotope Chemistry and Metabolite Synthesis Departments (ICMS), Sanofi-Aventis Deutschland GmbH, Frankfurt a. M, Germany

Zheng Lu Clinical Pharmacology and Exploratory Development, Astellas Pharma, Northbrook, IL, USA

Jochen Maas R&D Metabolism and PK Germany, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

Anita Chen Marshall Sequoia Healing Center for Acupuncture and Herbal Medicine, Alameda, CA, USA

American College of Traditional Chinese Medicine (ACTCM), California Institute of Integral Studies (CIIS), San Francisco, CA, USA

Bastyr University, Seattle, WA, USA

John A. Burns School of Medicine, University of Hawaii, Honolulu, HI, US

Soraia Martins Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Andrew McEwen Reverisco Solutions, Peterborough, UK

Mitul A. Mehta Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

Alex Mentink Centre for Human Drug Research, Leiden, The Netherlands

Guy Montay GMPK/PK Sanofi-Aventis, Vitry-Sur-Seine, France

Kenneth T. Moore Cardiovascular and Metabolism Medical Affairs, Janssen Pharmaceuticals Inc., Titusville, NJ, USA

James Gilmour Morrison GW Pharmaceuticals, Cambridge, UK

Michael A. B. Naafs Naafs International Health Consultancy, Oldenzaal, Netherlands

Yoichi Naritomi Analysis and Pharmacokinetics Research Labs., Astellas Pharma Inc., Ibaraki, Japan

Lisa Nguyen Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Petar Nikolov Gastroenterology, Ardea Medical Center, Sofia, Bulgaria

G. Nikolova University Hospital Acad I. Penchev, Medical University, Sofia, Bulgaria

M. Nikolova University Hospital Alexandrovska, Clinic of Nephrology, Medical University, Sofia, Bulgaria

Milena Nikolova-Vlahova Clinic of Nephrology, University Hospital Alexandrovska, Medical University, Sofia, Bulgaria

Pieter Okkerse Centre for Human Drug Research, Leiden, The Netherlands

Vinay Pawar Department of Molecular Immunology, Helmholtz Center for Infection Research, Braunschweig, Germany

Institute of Immunology, Medical School Hannover, Hanover, Germany

Daniela Penn Medisis, Neustadt, Germany

Thomas Penzel Interdisciplinary Sleep Medicine Center, Charitécentrum für Herz- Kresilauf- und Gefäßmedizin CCM11, Charité – Universitätsmedizin Berlin, Berlin, Germany

Vanessa Petroni Magri Department of Clinical Pharmacology and Therapeutics, Faculty of Medicine and Surgery, University of Malta, Msida, MSD, Malta

Italo Poggesi Global Clinical Pharmacology, Quantitative Sciences/Model-Based Drug Development, Cologno Monzese, Italy

Sanjive Qazi AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA

Bioinformatics Program, Gustavus Adolphus College, St. Peter, MN, USA

Md Shaifur Rahman Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Diether Rüppel PKDM/TMED, Sanofi-Aventis Germany, Frankfurt, Germany

Lorraine M. Rusch Highpoint Clinical Trials Center, High Point, NC, USA

Gerard Sanderink Translational Medicine and Early Development, Sanofi-Aventis R&D, Vitry-sur-Seine, France

Juergen Sandow Centre of Pharmacology, Frankfurt-Main University, Glashuetten, Germany

Enya E. Scanlon Stratified Medicine Group, Centre for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast, UK

Jörg Schüttrumpf Corporate R&D, Biotest AG, Dreieich, Germany

Werner Seiz R&D Translational Medicine and Clinical Pharmacology, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

Najam A. Sharif Global Alliances and External Research, Global Ophthalmology Research and Development, Santen Incorporated, Emeryville, CA, USA

Pieter Siebenga Centre for Human Drug Research, Leiden, The Netherlands

R. Slavchov Cambridge University, Cambridge, UK

Ferdi Sombogaard Department of Clinical Pharmacology and Pharmacy, Amsterdam University Medical Centers – VUmc, Amsterdam, The Netherlands

Lucas-Sebastian Spitzhorn Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany

Petra Starck-Lantova The University of Bonn/German Society for Regulatory Affairs, Bonn, Germany

Peter Stopfer Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany

Agnieszka J. Szczepek Department of ORL, Head and Neck Surgery, Charité University Hospital, Berlin, Germany

Lenore Teichert Biostatistics, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

V. Tenev Department of Psychiatry, University Hospital, University of Iowa, Iowa city, IA, USA

Petra A. Thürmann Philipp Klee-Institute of Clinical Pharmacology, HELIOS University Clinic Wuppertal, Wuppertal, Germany

University Witten/Herdecke, Witten, Germany

Hok Bing Thio Dermatology, Erasmus University Medical Center, Rotterdam, The Netherlands

Marie Trad Neurology Center of Excellence, IQVIA, Paris, France

Fatih M. Uckun AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA

Mohd Urooj Pharmacology Research Laboratory, Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India

Guido van Amerongen Centre for Human Drug Research, Leiden, The Netherlands

Anne-Ruth van Troostenburg Gilead Sciences International Ltd, Cambridge, UK

Ivana I. Vranic Clinical Center of Serbia, Faculty of Medicine, University of Belgrade, Belgrade, Serbia

Jan Wahlstrom Amgen, Inc, Pharmacokinetics and Drug Metabolism, Thousand Oaks, CA, USA

Tianli Wang Clinical Pharmacology and Exploratory Development, Astellas Pharma, Northbrook, IL, USA

Willi Weber Frankfurt, Germany

Roland Wesch R&D Metabolism and PK Germany, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

Lea Wettlaufer Medisis, Neustadt, Germany

Beat Widler Widler & Schiemann AG, Zug, Switzerland

Willi Weber has retired.

Larry Wienkers Amgen, Inc, Pharmacokinetics and Drug Metabolism, South San Francisco, CA, USA

Steven G. Woolfrey Independent Pharmaceutical Consultant – Clinical Pharmacology and Pharmacokinetics, Newcastle Upon Tyne, UK

Virginia Wotring Center for Space Medicine and Department of Pharmacology and Chemical Biology, Baylor College of Medicine, Houston, TX, USA

Daniel F. B. Wright School of Pharmacy, University of Otago, Dunedin, New Zealand

Wei Yin Quantitative Clinical Pharmacology, Takeda Pharmaceuticals, Cambridge, MA, USA

Yichao Yu Department of Pharmaceutics College of Pharmacy, University of Florida, Gainesville, FL, USA

Ming Zheng Clinical Pharmacology and Pharmacometrics, Bristol-Myers Squibb Company, Princeton, NJ, USA

Part I

Human Studies in Clinical Pharmacology

Methodologies of Safety Assessment in Clinical Pharmacology

1

Werner Seiz

Contents

W. Seiz (✉)
R&D Translational Medicine and Clinical Pharmacology,
Sanofi-Aventis Deutschland GmbH, Frankfurt am Main,
Germany
e-mail: werner.seiz@sanofi.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_30

Abstract

An important objective of clinical pharmacology is the early and ongoing assessment of the safety and tolerability of a new drug. This is done by assessing the type, frequency, and severity of side effects, assessing in which patient population these side effects may occur at which dose or exposure, for what duration and whether these side effects can be monitored and whether they are reversible. The terminology for the safety assessment of drugs has some specifics that need to be explained right at the beginning of this chapter.

Introduction/General Considerations

An important objective of clinical pharmacology is the early and ongoing assessment of the safety and tolerability of a new drug. This is done by assessing the type, frequency, and severity of side effects, assessing in which patient population these side effects may occur at which dose or exposure, for what duration and whether these side effects can be monitored and whether they are reversible. The terminology for the safety assessment of drugs has some specifics that need to be explained right at the beginning of this chapter.

Definition of Adverse Events as the Parameter to Assess Safety

The term "side effect" used for marketed drugs is replaced by the term "adverse event" (AE) in studies with investigational drugs. An adverse event is defined as any unfavorable and unintended sign, symptom, or disease temporally associated with the use of a drug, whether or not considered related to the drug.

In this chapter the term "adverse event" is explicitly also used for any abnormal laboratory value, as the consequence of abnormal values will be evaluated in the same scheme as for clinical adverse events.

The term "treatment related" is often added as a modifier in order to remove preexisting conditions from consideration. A further term "serious adverse event" is used to describe any untoward medical occurrence that, at any dose, results in death, is life-threatening, requires hospitalization or prolongation of an existing hospitalization, results in persistent or significant disability or incapacity, or is a congenital anomaly or a birth defect. Serious adverse events have to be reported to the health authorities in an expedited manner, typically.

The severity of an adverse or serious adverse event is classified as either mild, moderate, or severe. Standardized definitions for adverse events and classification of severity have been published by the National Cancer Institute (NCI) of the National Institute of Health (NIH) and are also used for clinical trials in nononcology indications (NCI 2017).

It is important to distinguish between the severity and the seriousness of an adverse event. A severe adverse event is not necessarily serious (e.g., severe abdominal cramps not causing hospitalization), and a serious adverse event is not necessarily of severe intensity (e.g., mild to moderate, prolonged dizziness of an outpatient causing hospitalization) (Herson 2000).

How to Manage the Safety Assessment of a Drug

One of the most critical steps in the development of a new drug is the first administration of a drug to humans, the first dose escalation, the first multiple dosing, and the first switch from healthy patients to the targeted patient population. In order to acquire the safety data in a responsible way, it is necessary to consider all of the following areas for each clinical study and to plan these items in advance:

- Expect, plan, and manage the occurrence of adverse events. This administrative part of the safety method includes the selection of the right preclinical animal models for the prediction of the target organ, the definition of the exposure to the drug at the no-observed-adverse-effect-level (NOAEL), the adequate calculation of the safe starting dose in humans, the decision about the dose escalation and when to stop it, the proper organization of the clinical trial, and the definition of the expected adverse event profile.
- Plan and manage the acquisition of adverse events data. This includes – based on the expected adverse event profile – the selection of the clinical, technical, and laboratory observations, by which the expected adverse events are to be monitored.
- Plan and manage the interpretation of the adverse events data and their impact on the subsequent development or study conduct. In order to avoid bias, the statistical analysis

of the safety data obtained has to be predefined, using commonly accepted criteria. For each parameter assessed, it should be clear prior to the analysis, which deviation is considered relevant and therefore an adverse event. This is usually done in the statistical analysis plan, which has to be finalized prior to closing the database of a study and prior to breaking the randomization code. In clinical pharmacology studies, data monitoring committees (DMC) are typically not included; however, in studies with adaptive designs DMCs might be installed very early in clinical development.

Case Study

The importance of an adequate selection of animal models, assessing the significance of the preclinical data obtained for humans and planning adequately the study conduct in the first-in-human study has been shown quite dramatically a decade ago. The first dose step in the first-in-man study with the biotherapeutic TGN1412, a humanized agonistic anti-CD28 IgG4 monoclonal antibody (present on regulatory T-cells), induced a cytokine release syndrome in all six active-treated healthy volunteers, all of whom suffered from life-threatening, acute shock and subsequent multi-organ failure. At least in one of the participants of the TGN1412 first-in-man study, several fingers and toes were to be amputated finally. Obviously this severe and serious adverse events were not predicted by the animal studies conducted prior to human studies, the dose administered was obviously above the minimum active biological effect level (MABEL) for humans, and all volunteers were already dosed before the first dosed person suffered from the symptoms of the upcoming cytokine release syndrome, that is, within less than 90 min. Although a complete explanation of the event was never unanimously accepted (http://www.circare.org/foia5/clinicaltrialsuspension_interimreport.pdf), at least it appears that the drug was given too fast to each subject (3–6 min infusion time) and to too many subjects within too short a time (every 10 min the next subject was dosed) (Horvath and Milton 2009). As a consequence of this event with TGN1412, the regulators worldwide have changed several processes so that this should not happen again. The European Medicines Agency (EMA) has revised their guidance on first-in-human clinical trials to identify and mitigate risks for trial participants (EMA 2017) after a case of death of a human volunteer in a first-in-human clinical trial in 2016. In that case, a 49-year-old healthy subject who experienced neurological symptoms after the 5th out of 10 planned doses in the first multiple ascending dose study with a nonselective fatty acid amid hydrolase inhibitor was submitted to hospital and died about 7 days later (Brentano and Menard 2016). A major new request by the EMA is the treatment of a sentinel at least 24 h before subsequent subjects are to be treated.

It is self-evident that the evaluation and interpretation of the safety data obtained as a whole is of utmost importance to a drug development program; however, here in this chapter the topic will be the technical description of the most often used clinical, technical, and laboratory methods to acquire safety data and how this will influence decisions on dose escalation or termination of a study. No more thoughts are given to analyze the safety data as a whole and in the context of the already accumulated clinical safety data.

Categorization of Adverse Events for Decision Making

Purpose and Rationale

Adverse events should be categorized in the same way across studies so that the decisions based on these categories are consistent within a development program and across programs. The NCI-CTCAE v5.0 terminology is defining the following five grades of severity for each adverse event (AE):

Grade	Definition
1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated

(continued)

Grade	Definition
2	Moderate; minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL)
3	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL
4	Life-threatening consequences; urgent intervention indicated
5	Death related to AE

Procedure

Each adverse event or finding is classified into one of five categories, where grade 1 indicates a deviation from the norm without an obvious relevance for the subject, grade 2 indicates a mild interference with daily activities for the subject but without need for treatment except non-opioid analgesics, for example. Occurrences of events of grade 2 have to be seen as an alert on reaching doses, where tolerability to the test compound decreases. Grade 3 indicates that the event or finding requires medical or other treatment or prevents daily activities of the subject. Grade 4 is reserved for definitely inacceptable adverse events, which typically, if not occurring in the placebo group, leads to a termination of the study at least of the dose, when the grade 4 event has been observed (e.g., rhabdomyolysis, angioedema). Grade 5 of course is an inacceptable consequence of treatment with an investigational drug and may even terminate a drug development project. If there is a rapid change in a parameter, this also might lead to an increase in grading. For laboratory parameters (chemistry, ECG), the grading is done based on the likelihood for further consequences or risks according to the categories above. In order to categorize laboratory values as abnormal, they have to be different from the normal range, which is specific to each laboratory. The numbers given here are suggestions and are based on published normal ranges (Kratz et al. 2004).

Evaluation

For each subject, the maximal adverse events' grading can easily be assessed. For gradings 3 or 4, unblinding is recommended; for grade 5, unblinding and immediate termination of the clinical trial is a must. An individual should not be further dosed if on active drug and grading 3 occurs. If only placebo-treated subjects suffer from an adverse event and if this event is not study-procedure related, it has no impact on further study conduct. If placebo- and active-treated patients suffer from grading 3, but with less than 50% of active-treated subjects, doses should be adapted (=lowered), the number of subjects treated at a time need to be reduced, and the time interval between subjects should be increased, in order to minimize the risk for treated subjects. If the dose step is well tolerated, additional subjects could be treated at the dose with the grade 3 events. Finally, if more than 50% of active-treated patients suffer from grade 3 adverse events at a given dose, the dose below is qualified as the maximal tolerated dose. At all grades, clinical judgment is needed based on the nature, reversibility, and monitorability of adverse events.

Critical Assessment of the Method

The categorization of information leads to a loss of information and therefore has to be used with care. Everyone using this method needs to be aware that the full picture and information needs to be taken into account and not just the results from the categorization. The grading system suggested and described here is modified from (Sibille et al. 2010) and not approved by any authority but should be seen as way to consistently aggregate and interpret information. Grading adverse events is in use in oncology and vaccine studies already (Cancer therapy evaluation program 2009; FDA Guidance 2007). These systems use four or five gradings, where grade 5 is always "death" and grade 4 is mostly of life-threatening adverse events, which is not in contradiction with the grading used here.

Modifications of the Method

An alternative to formal categorization of adverse events for subsequent decision making on dose escalation or study progress is the repetitive assessment of the uncategorized clinical and laboratory data by the investigator, the sponsor, and additional experts to achieve a common understanding on how to proceed. The precategorization of events as described here in this chapter, however, does not prevent such an approach and has the advantage to provide a consistent assessment of the information across dose steps, studies, and compounds.

Decision Making on Dose Increase and to Stop the Study

Purpose and Rationale

The decision to stop dose escalation should be based on the observation of adverse events no longer tolerable (by frequency or severity) and by the observed exposure information.

Procedure

The grading of the adverse events and their frequency need to be assessed. As long as no adverse events are observed and the exposure is not above the exposure in the most sensitive species, dose escalation may go on as planned in the protocol. If the exposure is above the NOAEL exposure, careful further dose escalation may be reasonable to define the maximal tolerable dose. If the severity of adverse events is below grade 3 and the exposure of the NOAEL is not reached, dose escalation can proceed. No further dose escalation should be considered, if more than 50% of active-treated subjects suffer from adverse events of grade 3.

Evaluation

In case grade 3 or 4 adverse events do occur, treatment of ongoing subjects should be stopped immediately.

Clinical Adverse Events Monitoring (Report by Subjects)

Purpose and Rationale

Most drug-related adverse events are based on the spontaneous reporting of clinical signs by the clinical trial participants. Subjects participating in a clinical trial can realize these adverse events at any time.

Procedure

The subjects are asked to report any events, signs, or abnormal observations and feelings to the study personnel immediately. In addition, subjects should be asked direct questions from time to time, such as “Did you make any disagreeable or unexpected observations since you took the drug?” The information obtained need to be documented without interpretation at first. The (preliminary) diagnosis and decision about next steps (physical, and if indicated additional laboratory or technical examinations) will be based on the interpretation by the responsible MD on this report.

Evaluation

Categorization of adverse events reported by the subjects is to be done by experienced medical personnel taking into account possible differential diagnosis and their time course. The reference http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm074775.htm gives an example of recommendations by the FDA.

Clinical Adverse Events Monitoring (Physical Examination by the Clinical Investigator)

Purpose and Rationale

Physical examination based on spontaneous reporting of adverse events will be conducted as needed.

Procedure

Physical examination can include auscultation, investigation of reflexes, or orientation, etc.

Evaluation

The investigator needs to decide on the classification of physical findings based on changes to baseline and their relevance.

Critical Assessment of the Method

Typically there will not be many clinical findings on physical examination. If there are some, this indicates already quite substantial effects (e.g., angioedema, rashes, ankle edema). Exceptions are findings in the vital signs of heart rate or blood pressure (see below), where effects are frequently seen.

Timing of Monitoring

Purpose and Rationale

Timing of clinical and laboratory assessments need to be in line with the timecourse of drug concentration over time.

Procedure

Standard monitoring needs to be done at baseline and repetitively after drug administration. For orally administered drugs, this is typically at baseline before drug administration – 30 min, 1, 2, 4, 8, 12, and 24 h after dosing for once daily drugs. Timing has to be tailored to the specific pharmacokinetic and pharmacodynamic profile of a drug.

Vital Signs

Heart Rate

Purpose and Rationale

The heart rate is influenced by the sympathic and parasympathic system, which can be affected by drugs directly or indirectly. Heart rate as a vital parameter has to be quite stable as heart rate effects in patients with ischemic heart disease could lead to angina pectoris. In phase I studies, heart rate typically is most affected by increased vagal tone and subsequent bradycardia and occasional fainting.

Procedure

Continuous ECG monitoring by telemetry or Holter ECG is the method of choice to observe effects on heart rate.

Evaluation

Normal range: 50–80/min in supine position. Grade 3 definition: $<$45/min for bradycardia. Grade 1 definition: 100–115/min. Grade 2 definition: 116–130/min. Grade 3 definition: $>$130/min.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Modifications of the Method

Holter monitoring allows continuous 24-h assessment of heart rate including analysis of cardiac arrhythmias. Holter monitoring should be used whenever there is evidence of pro-arrhythmic potential of a drug. Central analysis of Holter ECGs is recommended to also be able to compare acquired data to larger groups of subjects.

Vital Signs

Blood Pressure

Purpose and Rationale

Blood pressure is dependent on stroke volume, heart rate (stroke volume × heart rate = cardiac output), and peripheral resistance. Decrease in cardiac output and/or resistance decreases blood pressure and vice versa. A decrease in blood pressure is most often a result of either vasodilatation or decrease in heart rate, both of which can occur during increased vagal stimulation.

Procedure

Blood pressure can be measured manually or by a machine, in supine, sitting, or standing positions. For functional assessments the Schellong test is an easy to conduct procedure to measure the effect of a physical challenge on heart rate and blood pressure. After 10 min in supine position, the subject is asked to take a standing position. Heart rate and blood pressure are measured 2 min after end of supine position. Timepoints of blood pressure measurements need to be adapted based on the observed effects.

Evaluation

Normal range: 100–140 mmHg systolic in supine position, 50–85 mmHg diastolic in supine position. Grade 1 definition: 140–159 mmHg systolic in supine position and 90–99 mmHg diastolic in supine position for pressure increase; a quantity of 80–100 mmHg systolic in supine position for pressure decrease. Decrease in systolic blood pressure after 2 min standing by more than 20 mmHg together with increase in heart rate. Grade 2 definition: 160–179 mmHg systolic in supine position and 100–110 mmHg diastolic in supine position for pressure increase; a quantity of 70–80 mmHg systolic in supine position for pressure decrease. Cannot stay standing after 10 min of supine position for pressure decrease. Grade 3 definition: >180 mmHg systolic in supine position and >110 mmHg diastolic in supine position for pressure increase; below 70 mmHg systolic in supine position for pressure decrease or syncope during Schellong test.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Modifications of the Method

Twenty-four-hour ambulatory blood pressure monitoring (ABPM) is the method of choice for any compound with known or suspected effect of blood pressure as the effect over time can be best followed by continuous monitoring. Some drugs affect the nocturnal decrease in blood pressure (so-called dipping); whenever there is evidence that a drug has such an effect, ABPM should be used early in clinical development. Blood pressure (and heart rate) will be measured every 15–20 min during day time (defined as 6 a.m. to 10 p.m.). During night time, the measurement intervals are 30 min. Full 24 h should be measured, if ABPM is used. ABPM allows to calculate precisely peak and trough effects and the duration of effect on blood pressure.

ECG Parameter

PR Interval

Purpose and Rationale

The PR interval in the ECG is the time during which the electrical excitation is conducted from the atria to the AV-node. Prolongation of the PR interval is a potential side effect of drugs affecting repolarization and bears the risk of AV blockade.

Procedure

Evaluators should be trained in ECG analysis. Automated analysis is frequent but needs to be validated in order to rely upon it.

Evaluation

Normal range: 120–200 ms. Grade 1 definition: <0.8-fold LLN or >1.1-fold ULN. Grade 2 definition: >250 ms. Grade 3 definition: AV-block 2nd degree or syncope.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

ECG Parameter

QT Interval

Purpose and Rationale

The QT interval in the ECG is the time during which the electrical excitation and repolarization of the ventricles takes place. Prolongation of the QT time and especially the QTc time (QT time corrected for effect of heart rate) is a risk factor for torsades des point, a ventricular arrhythmia associated with an increased incidence of drug-induced sudden cardiac death. QT prolongation of drugs is one of the most frequent reasons for termination of a drug development program.

Procedure

Evaluators should be trained in ECG analysis. Automated analysis is frequent but needs to be validated in order to rely upon it.

Evaluation

Normal range for QTc: 360–425 ms for men, 380–445 for women, increase below 40 ms. Grade 1 definition: Increase above 40 ms and QTc below 475 ms. Grade 2 definition: 476–499 ms and increase below 60 ms. Grade 3 definition: Above 500 ms or increase exceeding 60 ms.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Glucose

Purpose and Rationale

A sufficient glucose concentration in blood (>2 mmol/L or >40 mg/dL at minimum) is essential for all life processes. Whenever there are signs of decreased consciousness, this vital parameter has to be assessed immediately.

Procedure

Blood glucose should be measured from capillary or venous blood at predefined timepoints and in addition in cases of suspected hypoglycemia or impaired consciousness.

Evaluation

Normal range 3.8–6.4 mmol/L. Grade 1 definition for hypoglycemia: 3.5–3.8 mmol/L. Grade 2 definition: 2.2–3.4 mmol/L. Grade 3 definition: 1.7–2.1 mmol/L.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Potassium

Purpose and Rationale

Potassium concentration in cells is 25-fold higher than in blood. In all cases were potassium is released into the peripheral blood (e.g., during and after hypoxic events) or a decrease in renal excretion occurs, potassium increases will have the potential for cardiac bradyarrhythmias. Hypokalemia can lead to ventricular tachyarrhythmias. Therefore close monitoring of potassium concentration in serum is very important in early phases of development as long as the effect on its concentration in serum is not yet known.

Procedure

Potassium values are measured from serum taken from peripheral veins at predefined timepoints.

Evaluation

Normal range 3.5–5.0 mmol/L. Grade 1 definition: 3.1–3.4 for hypokalemia and 5.1–6.0 for hyperkalemia. Grade 2 definition: 2.5–3.0 mmol/L for hypokalemia and 6.1–6.5 mmol/L for hyperkalemia. Grade 3 definition: 2.0–2.4 mmol/L for hypokalemia and 6.6–7.0 mmol/L for hyperkalemia.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Alanine Aminotransferase (ALT)

Purpose and Rationale
Hepatic damage is one of the most frequent drug-related adverse events and needs to be monitored in every clinical pharmacology study. Transaminases (SGPT/ALT and SGOT/AST), alkaline phosphatase, and total and conjugated bilirubin are the serum assays to detect liver damage.

Procedure
ALT, AST alkaline phosphatase, and bilirubin are taken from serum of peripheral blood at predetermined timepoints and more frequently, if any increases are seen. If increase of ALT is above threefold upper limit of normal (ULN), ALT needs to be followed until normalization (below ULN) or until no further decrease of ALT after termination of treatment is observed.

Evaluation
Any transaminase elevation above the upper limit of normal should be considered as an indicator for hepatic damage. ALT increase is the enzyme specific for liver damage. Normal range is 0–60 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.2-fold ULN. Grade 2 definition: Increase 2.5- to 5-fold. Grade 3 definition: >5- to 10-fold.

Critical Assessment of the Method
Basic method for safety and tolerability assessment. Increases of ALT are very specific to the liver. Alkaline phosphatase can be increased in other diseases as well, for example, bone disease. Depending on preclinical data of potential liver toxicity and upcoming clinical data early in development, the reporting thresholds for increases in liver enzymes should be adapted specifically.

Laboratory Parameter

Aspartate Aminotransferase (AST)

Evaluation
Normal range is 0–40 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.2-fold ULN. Grade 2 definition: Increase 2.5- to 5-fold. Grade 3 definition: >5- to 10-fold.

Critical Assessment of the Method
Supporting parameter for ALT analysis.

Laboratory Parameter

Phosphatase

Evaluation
Normal range is 30–120 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.1-fold ULN. Grade 2 definition: Increase two- to threefold. Grade 3 definition: three- to tenfold.

Critical Assessment of the Method
Supporting parameter for ALT analysis.

Laboratory Parameter

Bilirubin

Purpose and Rationale
Bilirubin assessment together with ALT measurement is used to identify potential risks of hepatic toxicity.

Evaluation
Normal range is 5–27 μmol/L. Grade 1 definition: Increase >1.3-fold ULN.

Hy's law (FDA 2009) is a prognostic indicator that a drug-induced liver injury leading to jaundice has a case fatality rate of 10–50%. Hy's law cases have the three following components:

- The drug causes hepatocellular injury, generally shown by more frequent threefold or

greater elevations above the ULN of ALT or AST.
- Among subjects showing such ALT/AST elevations, often much greater than 3 × ULN, some subjects also show elevation of serum bilirubin to >2 × ULN, without initial findings of cholestasis (serum alkaline phosphatase [ALP] activity >2 × ULN).
- No other reason can be found to explain the combination of increased transaminases and bilirubin, such as hepatitis A, B, or C, preexisting or acute liver disease, or another drug capable of causing the observed injury.

Critical Assessment of the Method
Together with ALT, a very powerful parameter to identify true drug-related hepatic damage.

Laboratory Parameter

Creatinine

Purpose and Rationale
Creatinine is solely excreted by the kidney, primarily by glomerular filtration, and therefore is a good marker of renal perfusion and filtration. Drugs affecting renal perfusion or filtration lead to an increase in creatinine. Increases in creatinine only occur if there is already a significant decrease in renal glomerular filtration rate.

Procedure
Creatinine concentration needs to be measured in plasma and urine. Together with the urine production per time (either within 24 h, or time overnight sampling; for example, 1,500 ml excreted between 8 pm and 7 am), the glomerular filtration rate can easily be calculated.

Evaluation
Normal range: 80–130 μmol/L. Grade 1 definition: >1.1-fold ULN. Grade 2 definition: >1.5-fold ULN. Grade 3 definition: >1.9- to 3.4-fold ULN.

Critical Assessment of the Method
Serum creatinine levels are not very sensitive to large changes in GFR as long as the GFR is still above 60 ml/min/m^2, but then a rapid increase will be observed. A more sensitive method for renal function is the GFR or the fractional excretion of electrolytes.

Laboratory Parameter

Albumin in Urine

Purpose and Rationale
Presence of albumin in urine is an indicator of glomerular damage.

Procedure
Albumin is measured from morning urine.

Evaluation
Normally no albumin is excreted via urine. Any finding of albumin above 300 mg/24 h in urine is indicative of a renal issue that needs to be further evaluated (if prior to treatment with investigational drug the value was negative).

Critical Assessment of the Method
Albumin in urine is always a pathological sign, which needs further analysis.

Laboratory Parameter

Creatinphosphokinase (CPK)

Purpose and Rationale
CPK is released during damage of skeletal muscle, a frequent side effect of lipid lowering compounds like statins.

Procedure
CPK is taken from serum of peripheral blood at predetermined timepoints and more frequently, if any increases are seen. If increase of CPK is above threefold ULN, CPK needs to be followed until normalization (below ULN) or until no further decrease of CPK after termination of treatment is observed.

Evaluation
Normal range: 50–400 IU/L. Grade 1 definition: 480–1,000 IU/L. Grade 2 definition: 1,000–2,000 IU/L. Grade 3 definition: 2,000–5,000 IU/L.

Laboratory Parameter

Hemoglobin (Male Subjects)

Purpose and Rationale
Hemoglobin can be affected by acute bleeding, by chronic suppression of erythrogenesis, or by dilution/concentration due to changes in the intravasal volume.

Procedure
Hemoglobin is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation
Normal range for males: 13.5–17.5 g/dL. Grade 1 definition: 12.0–12.5 g/dL and decrease >1.5 g/dL. Grade 2 definition: 10.0–11.9 g/dL. Grade 3 definition: <10.0 g/dL. Normal range for females: 12.5–15.5 g/dL. Grade 1 definition: 11.0–12.0 g/dL and decrease >1.5 g/dL. Grade 2 definition: 9.5–10.9 g/dL. Grade 3 definition: <9.5 g/dL.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Polymorphonuclear Leucocytes (PMN)

Purpose and Rationale
Immunotoxic effects of drugs on white bloods cells are not uncommon and need to be detected early on in development.

Procedure
PMN count is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation
Normal range: 1.7–7.5 G/L. Grade 1 definition: <0.7-fold LLN or >1.3-fold ULN. Grade 2 definition:1.0–1.3 G/L. Grade 3 definition: <1.0 G/L.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Platelets

Purpose and Rationale
Immunotoxic effects of drugs on platelets are not uncommon and need to be detected early on in development.

Procedure
Platelet count is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation
Normal range: 150–350 G/L. Grade 1 definition: <0.85 LLN. Grade 2 definition: 50–125 G/L. Grade 3 definition: <50 G/L.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Coagulation Parameter

Activated Partial Thromboplastin Time (aPTT)

Purpose and Rationale
Effects on aPPT are seen in cases of decreased hepatic protein synthesis rate.

Procedure
aPTT is assessed from plasma.

Evaluation
Normal range: 22–43 s. Grade 1 definition: 1.1- to 1.3-fold ULN. Grade 2 definition: 1.3- to 1.5-fold ULN. Grade 3 definition: >1.5-fold ULN until minor bleeds. Grade 4 definition: Major bleeds.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Kidney Injury Molecule-1 (KIM-1)

Purpose and Rationale
KIM-1 is a rather new biomarker indicating renal toxicity at the tubular level. KIM-1 has been preclinically qualified as an excellent marker for drug-related renal toxicity. If there is preclinical evidence for renal toxicity at this region and if KIM-1 has been used in nonclinical toxicity studies, this parameter should be monitored.

Procedure
KIM-1 can be measured using commercially available kits.

Evaluation
Look for significant changes from baseline and if those occur, stop treatment, follow KIM-1 until normalization.

Critical Assessment of the Method
There is limited experience with KIM-1 in healthy subjects and in clinical pharmacology studies. The marker is not well established in its performance in nondisease states so far. Therefore descriptive analysis of the marker and analysis of traditional parameters such as serum creatinine or BUN together with KIM-1 in order to get more experience with the marker is advised.

Modifications of the Method
There are several additional bimarkers for assessment of renal toxicity like alpha-GST or NGAL. They are also well qualified in nonclinical toxicity studies with nephrotoxicants. There is only limited information about the normal ranges and the spontaneous variations available currently.

Visual Analogue Scale for Semiquantitatively Assessing Pain and Other Subjective Factors

Purpose and Rationale

A visual analogue scale (VAS) is a psychometric response scale, which can be used in questionnaires. It is a measurement instrument for subjective characteristics or attitudes that cannot be directly measured, for example, pain or subjective assessment of the effectiveness of a treatment.

For the quantification of these subjective factors, the VAS is an instrument that tries to measure the severity across a continuum from none to an extreme amount of the characteristic. For example, the spectrum of pain to a subjective suffering from it appears to be continuous and does not take discrete jumps, as the typical categorization of none, mild, moderate, and severe suggests. In order to reflect this idea of an underlying continuum the VAS was introduced.

Procedure

Operationally, a VAS is usually a horizontal line, 100 mm in length, anchored by word descriptors at each end, for example, “no pain” and “maximum pain.” When responding to a VAS, subjects are asked to indicate their level of agreement to a statement by indicating a position along a continuous line between the two end points mentioned. This continuous (or “analogue”) aspect of the scale differentiates it from discrete scales (e.g., “none-mild-moderate-severe” or “A to F”).

Evaluation

The VAS score is determined by measuring in millimeters from the left-hand end of the line to the point that the patient marks.

Critical Assessment of the Method

As such an assessment is clearly highly subjective, these scales are of most value when looking at change within individuals, and are of less value for comparing across a group of individuals at one time point. It could be argued that a VAS is trying to produce interval/ratio data out of subjective values that are at best ordinal. Thus, some caution is required in handling such data. Many researchers prefer to use a method of analysis that is based on the rank ordering of scores rather than their exact values, to avoid reading too much into the precise VAS score.

However, a VAS is extremely simple to use, easy to teach and understand. Therefore, bias introduced by complexity can be ignored.

For efficacy studies in patients, where pain is a primary or secondary outcome parameter, the VAS is only of limited value.

In practice, computer-analyzed VAS responses may be measured using discrete values due to the discrete nature of computer displays.

The VAS can be compared to other linear scales such as the Likert scale or Borg scale. The sensitivity and reproducibility of the results are broadly very similar, although the VAS may outperform the other scales in some cases [1].

Modifications of the Method

Due to the limitations mentioned above, several additional tools for pain assessment have been developed and validated, such as the McGill pain questionnaire, where several dimensions of pain are assessed. As for all questionnaires, it is very important to have the questionnaire available in the validated version of the native language. Otherwise the outcome of the questionnaires from different languages cannot be compared. These complicated and often patent protected questionnaires do not have a major place in clinical pharmacology studies.

Summary

It should be kept in mind that during the first clinical studies, there is practically no information about the safety and tolerability of a new drug as compared to the knowledge accumulated later on. Nevertheless, only during these initial studies the administration of the drug occurs under such kinds of secure conditions concerning the ability to handle side effects that dose escalation should not be stopped too early. It has to be kept in mind that during phase II and III studies and even more during marketing of a drug, the exposure of the drug to patients might occasionally – due to overdose, poor metabolization, or other causes of accumulation – be much higher than intended. Especially for these cases the company developing a drug should know, which kind of side effects would be expected.

References and Further Reading

Cancer Therapy Evaluation Program (2009) Common terminology criteria for adverse events (CTCAE) v4.0. Accessed 29 May 2009

Brentano CF, Menard J (2016) The BIAL/Bootrial case of death of a human volunteer in the first-in-human study of BIA 10-2474: are we missing the fundamentals. Presse Med 9:719–722

EMA (2017) Strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products. EMEA/CHMP/SWO/28367/07 Rev. 1

FDA Draft Guidance (2007) Guidance on industry drug induced liver injury: premarketing clinical evaluation. CDER, Rockville, MD, USA

FDA Guidance (2007) Toxicity grading scale for healthy adult and adolescent volunteers enrolled in preventive vaccine clinical trials, Rockville, MD, USA

FDA Guidance (2009) Drug-induced liver injury: premarketing clinical evaluation, Rockville, MD, USA

Herson J (2000) Data and safety monitoring committees in clinical trials. Chapman & Hall, Boca Raton, p 48

Horvath CJ, Milton MN (2009) The TeGenero incident and the Duff report conclusions: a series of unfortunate events or an avoidable event? Toxicol Pathol 37(3):372–383

http://www.circare.org/foia5/clinicaltrialsuspension_interimreport.pdf. Accessed 19 Jul 2010

http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm074775.htm. Accessed 19 Aug 2010

ICH (2005) The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential of non-antiarrhythmic drugs. ICH harmonized tripartite guideline E14, final version

Kratz A, Ferraro M, Sluss P, Lewandroski K (2004) Laboratory reference values. N Engl J Med 351:1548–1563

Lewis JH (2006) "Hy's law", the "Rezulin rule", and other predictors of severe drug-induced hepatotoxicity: putting risk benefit into perspective. Pharmacoepidemiol Drug Saf 15:221–229

Lock EA (2010) Sensitive and early markers of renal injury: where are we and what is the way forward. Toxicol Sci 116:1–4

NCI (2017) https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf. Accessed 23 Nov 2018

Sibille M, Patat A, Caplain H, Donazzolo (2010) A safety grading scale to support dose escalation and define stopping rules for healthy subject first-entry-into-man studies. Br J Clin Pharmacol. https://doi.org/10.1111/1365-2125.2010.03741.x

Pharmacodynamic Evaluation: Cardiovascular Methodologies

2

Ivana I. Vranic

Contents

Abstract

Cardiovascular methods to assess pharmacodynamics nowadays evolve very quickly, due to rapid progress in high technology and IT sector. Noteworthy, mathematical approach grows very fast in new algorithms to analyze

I. I. Vranic (✉)
Clinical Center of Serbia, Faculty of Medicine, University of Belgrade, Belgrade, Serbia
e-mail: ivvra.fesc@gmail.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_31

the heart signal. Many areas of multiple organ damage will relay in very complex software and hardware innovations. Basics for this growth is understanding of previously unknown mechanisms of control of physiological functioning like heart stiffness and compliance. Other reasons go to research in Shannon's entropy and derived calculations. On the other hand, some previous methods have been surpassed like arterial pulse methods when it comes to pharmacodynamics research. It is of importance also to take into account rare diseases and various channelopathies that may interfere with pharmacodynamics evaluation on large-scale clinical trials. In phases III and IV of clinical research, those factors may influence final statistical results. New tests and old proven measures of hemodynamic stabilities are required to evaluate new therapeutics during clinical studies to be able to treat more people on pharmacogenetic basis with pharmacogenomic approach. Safety to treat with new drugs comes into the first place, so many requirements in monitoring of data gathered by contract research organization (CRO) are necessary to get the approval of FDA, and European Medicines Agency (EMA) is the European Union's equivalent to the US Food and Drug Administration (FDA). Those approvals mainly rely on pharmacodynamics data pooled out from clinical drug researches. To be more rapidly accessible, adverse effects are collected via wireless technologies and monitored on wider basis across multicentric studies. Therefore, guidelines on consistent methodology toward new therapeutics approach are adopted constantly.

Introduction

Understanding pharmacodynamics of novel testing drug is essential for its clinical utility especially when it comes to safety for use in human population. Therefore, the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) regularly provides updates and guidelines on consistent methodology toward new therapeutics by adopting quality, safety, efficacy, and multidisciplinary guidelines (www.ich.org). ICH is unique in bringing together the regulatory authorities and pharmaceutical industry to discuss scientific and technical aspects of drug registration, and it is discussed by panel of key opinion leaders and adopted worldwide.

Studying pharmacodynamics of novel therapeutic is of importance from the aspect of exposure/response relationships, weighted as pharmacokinetics (PK) relationship to pharmacodynamics (PD). This PKPD relationship is central to utilizing experimental data to enable right decision-making in drug development put into clinical practice.

Also, studying pharmacodynamics is not the same in pediatric and/or elderly population, therefore yielding adequate study population groups is essential for clinical drug investigation.

Novel approach is pharmacogenetics which refers to heterogeneity in drug response, whereas pharmacogenomics tends to evolve into personal medicine, meaning "the right dose, of the right drug, for the right person." Therefore, there is a need to target individual patient's response of those who will benefit the most and suffer the least. Nevertheless, genetic variation in subgroups of patients will create a different response in pharmacodynamics.

Various newly recognized channelopathies, like the ones previously described involving the *nervous system* (i.e., generalized epilepsy with febrile seizures, familial hemiplegic migraine, episodic ataxia, hyperkalemic and hypokalemic periodic paralysis), the *cardiovascular system* (i.e., long QT syndrome, short QT syndrome, Brugada syndrome, and catecholaminergic polymorphic ventricular tachycardia), the *respiratory system* (i.e., cystic fibrosis), the *endocrine system* (i.e., neonatal diabetes mellitus, familial hyperinsulinemic hypoglycemia, thyrotoxic hypokalemic periodic paralysis, familial hyperaldosteronism), the *urinary system* (i.e., Bartter syndrome, nephrogenic diabetes insipidus, autosomal-dominant polycystic kidney disease, hypomagnesemia with secondary hypocalcemia), and the *immune system* (i.e., myasthenia gravis, neuromyelitis optica, Isaac syndrome,

and anti-NMDA [N-methyl-D-aspartate] receptor encephalitis); need new pharmacodynamics aspect evaluation of new innovative therapeutics.

Validation of Cardiovascular Test Criteria

Pharmacological drug testing in clinical setting has five phases. Phase 0 refers to newly introduced step in testing micro dose of drug in short period of time 1–3–7 days, to be able to select its suitability to enter Phase I.

Phase I refers to the first introduction of a drug into humans – usually healthy voluntaries up to 20 or 80 subjects.

Phase II investigation consists of controlled clinical trials designed to demonstrate effectiveness and relative safety (drug vs. placebo).

Phase III trials are performed after the drug has been shown to be effective and are intended to gather additional evidence of effectiveness for specific disease indications and more precise definition of drug-related adverse effects and safety. This phase includes both controlled and uncontrolled studies.

Phase IV trials are conducted after the national drug registration authority has approved a drug for distribution or marketing. These trials may include research designed to explore a specific pharmacological effect, to establish the incidence of adverse reactions, or to determine the effects of long-term administration of a drug.

All cardiovascular tests planned by the protocol should be monitored, in sense of validity, objectivity, and repeatability. Usually, there is an outsourcing contract research organization (CRO) who is responsible for quality control of the equipment.

The principal investigator is responsible for minimizing the individual risk and optimizing the ethical benefit of the study to the group. Public sanctioning of the study is done by Ethical Committee. While considering usability of test chosen, it should provide a tangible and measurable result of drug action. As for quality criteria of new drug, it should unequivocally prove its safety in human population.

Empirical Quality Criteria

The usability of a method can be quantified by a formal assessment of empirical quality criteria based on test-theoretical principles:

Objectivity: Objective is the extent of investigator independence in conducting the test, analyzing its results, and interpreting its data. It should be done in double-blind fashion with strict standardization and between observer agreement and consistency.

Reliability and sensitivity: Standard error is accepted because it reflects physiological variability between subjects and methods used for quantification. A test or method is reliable if it is hardly subject to such variability and yields highly consistent results when repeated, although this does not imply that the results are objective, which is a matter of sensitivity of the test.

Pharmacosensitivity and pharmacospecificity: The capacity to detect drug-induced systematic effects, pharmacosensitivity reflects reliability of method selected for quantification, intrasubject repeatability of drug-related changes. Partially those variations belong also to circadian rhythm, postprandial effects. The ability to separate these is a proof of specificity.

Economy: If a method can be repeated several times in cost efficient and time efficient way with electronic wireless transfer it is better.

Validity: Methods are valid if they measure what they claim or intend to measure.

Assessment of agreement among observers is important in evaluating test objectivity, also agreement on testing the method on different manufacturer's equipment is essential in this regard.

Issues with Cardiovascular Test Methodology and Measurements Validity

It is important though that measurement equipment be calibrated and standardized as to provide intrasubject repeatability and test

objectivity in drug host reaction. In this way also interobserver variability is less present. Correlation and regression analyses are more accurate in drug effectiveness and safety as well.

Clinical Trial Legal Regulations and Good Clinical Practice (GCP)

Clinical trial by definition is any investigation in human subjects intended to discover or verify the clinical, pharmacological, and/or other pharmacodynamic effects of an investigational product(s), and/or to identify any adverse reactions to an investigational product(s), and/or to study absorption, distribution, metabolism, and excretion of an investigational product(s) with the object of ascertaining its safety and/or efficacy. The terms clinical trial and clinical study are synonymous (ICH GCP 1.12).

The regulative that describes standard for the design, conduct, performance, monitoring, auditing, recording, analyses, and reporting of clinical trials that provides assurance that the data and reported results are credible and accurate, and that the rights, integrity, and confidentiality of trial subjects are protected is also very important (ICH E6: GCP).

The need for more efficient approach in clinical trials goes to design, conduct and oversight, recording and reporting as new tools to trace immediate results. It is wireless technology that enables monitoring instantaneously and on long distances.

A document that describes the objective(s), design, methodology, statistical considerations, and organization of a trial is called clinical trial protocol. The protocol usually also gives the purpose and rationale for the trial, but these could be provided in other protocol referenced documents (ICH GCP 1.44). Throughout the ICH GCP Guidance, the term protocol refers to protocol and protocol amendments.

Cardiovascular Tests in Pharmacodynamics

Circadian Blood Pressure

Purpose and Rationale

Arterial blood pressure (ABP) is the pressure derived from circulating blood upon the intrinsic wall of the arterial vessel, also called systemic blood pressure. It is the pressure derived from ejected blood out of the left ventricle during the systole, and pulsatile wave of sequential modulation thereof, during the progression of the pulse wave to the rest of circulation.

Due to the transformation of the pulse wave, the maximum ("systolic" BP [SBP]) and minimum ("diastolic" BP [DBP]) blood pressure are reflections of the central hemodynamics. Since blood pressure declines almost exponentially over the diastole, bradycardia has a direct DBP reducing effect unrelated to central pump function and peripheral vascular resistance; this is often associated with a relatively higher SBP due to a higher preload while longer filling phase.

In BP estimation there are few international guidelines to consult: the eight report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, American hypertension guideline (JNC 8) and the National Institute for Health and Care Excellence (2011, 2015,2016) Hypertension: clinical management of primary hypertension in adults, British hypertension guideline (NICE) are widely known and accepted international guidelines, although there are some countries with local guidelines (Chiang 2017).

Procedure

Blood pressure is still mostly measured noninvasively according to the principle of Riva-Rocci in 1896, i.e., by inflating a cuff around the upper arm up to arterial compressive occlusion and then slowly deflating the cuff while the pressure in the cuff is measured. Originally, the cuff pressure was measured by means of a mercury sphygmomanometer; most present devices use an aneroid manometer or electronic pressure transducer;

nevertheless, blood pressure is still generally reported in millimeters mercury (mmHg).

Measuring pulse signals originating from systolic and diastolic blood pressure can be achieved by palpation, auscultation, or oscillometry, from a suitable vessel site distant from the cuff. Palpatory (systolic) blood pressure is now only used for emergency evaluations; manually operated auscultatory blood pressure with an appropriately adjusted cuff has long been considered to be the method of choice for clinical practice (using aneroid manometers) and clinical trials (using a random zero mercury sphygmomanometer). Blood pressure is now mostly measured by means of oscillometric devices with automated inflation and deflation; the increasingly frequent use of the automated devices has the implication that many clinicians and nurses are no longer sufficiently well acquainted with the manual auscultatory methods, which rely on cautious highly observer-dependent auscultation of the Korotkoff-I sound (first appearance of a clear tapping sound that gradually increases in intensity) for SBP and the Korotkoff-IV (sound muffling [DBP_{KIV}]) or Korotkoff-V auscultatory criterion (sound disappearance [DBP_{KV}]) for DBP. On the other hand, modern technology has introduced newer robust devices that can be self-operated by patients and trial subjects also with devices that measure BP from the wrist.

Home BP monitoring (HBPM) has been heralded as a useful and reliable measure of BP (Chia et al. 2017). Rather than replacing clinic BP measurements and ambulatory BP monitoring (ABPM), HBPM is a complementary tool that helps to eliminate the white-coat effect and identify masked hypertension. HBPM can also improve patient awareness and treatment adherence by giving all patients with hypertension and their family members a more active role in the management of the disease (Shrout et al. 2017).

Evaluation

The shift from manual (mercury-based) auscultatory to automated oscillometric methods has been the subject of controversy, which has only partly been resolved by imposing standardized (cross-) validation procedures.

Automated systems are highly robust and economic since they do not rely on an experienced analyst. This may result in less well-standardized conditions of measurement by lack of experience and discipline (choice of cuff, position of cuff, position of the microphone or oscillometric sensor, inflation speed, deflation speed, adjusted deflation speed when the pulse rate is low or irregular, posture of the patient, resting time, etc.) (Wood et al. 2017).

On the other hand, this makes measurements made primarily for safety surveillance more reliable and useful also for further (efficacy-based or pharmacodynamic) assessments (Butlin and Quasem 2017).

Critical Assessment of the Method

Only invasive blood pressure monitoring (in intensive care unit (ICU)) (Zhou et al. 2017) is superior to noninvasive Riva-Rocci principle (previously discussed). Discrepancies between the two methods can be equal to or less than 5%, provided that all instructions in technical manual are addressed properly.

Ambulatory Blood Pressure Monitoring (ABPM)

Purpose and Rationale

Single time pressure measurement is only a snapshot of the blood pressure pattern, whereas there are normal fluctuations according to activities and circadian rhythm during the day and night. Single time pressure measurement fails to detect blood pressure fluctuations due to autonomic modulation, physical and emotional stress, and the modification thereof by therapeutic intervention (Mc Kinstry et al. 2015).

ABPM typically provides the following three types of information: an estimate of mean BP level, the diurnal rhythm of BP, and BP variability (Thomas et al. 2006).

This seemingly perfect measurement for night blood pressure (NBP), however, has its own vital

limitations. First, all currently available ambulatory BP monitors produce sonorous stimuli, which have been found to disturb sleep significantly in a substantial proportion of patients. Furthermore, the correlation between NBP derived from ABPM and target organ damage tends to be weaker, with the lowest sleep quality mainly resulting from the repeated cuff inflations during overnight BP monitoring. Second, ABPM is not commonly employed in routine clinical practice for evaluating NBP, mainly because of its high cost and inconvenience in performing multiple NBP measurements.

Similarly, the appropriateness and efficacy of antihypertensive interventions should take ambulatory blood pressure measurement/monitoring (ABPM) data into account also with regard to their chronobiological fluctuations (Lee et al. 2015). Indeed, numerous larger-scale outcome studies have shown that ambulatory blood pressure measurement/monitoring (ABPM) yields better predictors of cardiovascular events when compared to timed manual BP readings in the physician's office or at home, even when the latter are taken carefully and in strict adherence with pertinent guidelines.

Procedure

The first device for noninvasive ambulatory BP (ABP) monitoring (ABPM) was developed in 1962 and subsequently modified by Sokolow et al. in 1966 (Sokolow et al. 1966). It used a microphone taped over the brachial artery, a cuff inflated by the patient, and a magnetic tape recorder for storing cuff pressure. Presently, most devices are automated and rely on the oscillometric analysis of the vascular sound.

A large variety of devices for ambulatory measurement are available. These devices generally provide for both event-triggered and automated oscillometric measurements according to a present protocol of regular intervals (that may be set differently for the day and night measurements). It is important to use a device that has been validated independently, for instance, according to the protocol of the British Hypertension Society (O'Brien et al. 1993) or that of the US Association for the Advancement of Medical Instrumentation (Association for the Advancement of Medical Instrumentation 1993) or both.

Evaluation

The interpretation of ABPM data should be based on standardized criteria (O'Brien et al. 2000; Verdecchia et al. 2004).

An average daytime ABP <135 mmHg systolic and 85 mmHg diastolic is generally considered normal for adults; levels <130/80 mmHg may be considered optimal. Subjects with daytime systolic average ABP values <130 mmHg can be considered to be at only minimal cardiovascular risk even if the reading in the physician's office was higher (exclusion of white-coat hypertension).

In hypertension management, it is important to analyze both day- and nighttime readings, although the latter may have to be set at broader intervals in order not to disturb sleeping rest. The day–night time fluctuations are generally used to calculate the BP-dip (= (1 ($SBP_{sleeping}/SBP_{daytime}$)) 100), with categories such as nondipper (0–10%), dipper (10–20%), extreme dipper (>20%), and reverse dipper (<0%) (O'Brien et al. 1988). In healthy individuals, NBP decreases by 10–20% and increases promptly on waking. However, certain abnormal diurnal variation patterns have been described in which the nocturnal fall of BP may be more than 20% (extreme dippers), <10% (nondippers), or even reversed (reverse dippers). Nevertheless, in recent years, evidence has shown that nocturnal BP levels rather than circadian BP pattern are more accurate in predicting mortality and morbidity related to BP, independent of mean BP and daytime BP levels.

Further important criteria are the overall BP variability and early morning surges, and the pulse pressure (SBP – DBP). Since most systems also report pulse rate, ABPM data can also be used to assess pulse rate variability.

Critical Assessment of the Method

Serial single time BP measurements should be needed to provide complete and accurate diagnostics of hypertension and the need for treatment thereof. For sure, some specific hours should be outlined and sought for in the evaluation of the efficacy of antihypertensive medication. For

example, exact time measurement, same apparatus, same personal, before drug intake, after 1 h of waking and urination and before breakfast, also after period of fasting for 6 h at least. Two measurements with pause of 1 min should be taken and mean values considered into account. While addressing cardiovascular safety of non-cardiovascular therapeutic, we ought to take diurnal fluctuations as well. Finally, ABPM seems to be very important method in the clinical pharmacological evaluation of cardiovascular effects especially during the night time period and because some people are dippers (hypotensive postural orthostatic tachycardia syndrome (Shibao et al. 2013) and/or Bradbury-Eggleston syndrome) and nondippers (nocturnal hypertension).

Electrocardiography

Standard 12 lead ECG is the good old standard test, well known for its unsurpassed diagnostic capacity when it comes to accurateness, availability, and cost-benefit, as well as multiple feasible repeats.

Purpose and Rationale

The electrocardiogram (ECG, also EKG) is the main noninvasive method among other less frequently used in everyday clinical practice. Back in 1901 Einthoven was first to recognize heart's electroactivity using string galvanometer and was the first to construct forerunner of contemporary ECG machine. He also named P wave, QRS complex, and T wave some time previously in 1895. The Nobel Prize for this achievement was awarded to him in 1924. Needless to say how important his discovery was for the sake of the whole Medicine! Those tiny superficial signals from transthoracic spread of the electrical activity of the heart were to be gathered during systole (contraction phase) and diastole (relaxation phase). What came later was of even bigger importance: that ECG analysis provided information on frequency and origin (sinus rhythm or other), presence of premature atrial or ventricular beats (ectopism), nature of intra-atrial conduction (existence of block), sorting-out atrial depolarization and repolarization timeline, assessing atrioventricular conduction properties (in search of concealed pathway and/or block), or to analyze intraventricular and transventricular conduction (search for accessory pathway or block) and ventricular depolarization and repolarization phase (hypoxic or hypothermal injury).

Also, information about cardiac memory that is new and trendy could be very helpful in analyzing drug effects in clinical trials. Noteworthy, many rare diseases, which affect not more than 5 per 10,000 persons in the European Union and encompass between 6000 and 8000 different entities (which affect more than 30 million people just in the EU), have its own ECG varieties and presentations still waiting to be addressed and fully recognized.

Procedure

Still relying on the pioneer work of Willem Einthoven, modern ECG diagnostics now involve digital recording, analysis, and archiving of the ECG tracings and related data (for instance, P-wave duration, PQ-interval, QRS duration, QT-interval, P-wave, QRS-wave, and T-wave vector amplitude and angle). The highest precision is achieved by recording the signals from the bipolar Einthoven leads (I, II, III), amplified unipolar Goldberger leads (aVR, aVL, aVF), and unipolar precordial Wilson leads (V_1–V_6) simultaneously for a sufficiently long time (10 s at least) and at a sufficiently high writing speed (25–50 mm/s).

Standard ECG recording is done with patient resting (12 leads recorded simultaneously, 25 mm/sec paper speed, 10 mm/mV gain, and filter band settings from 0.05 Hz to 150 Hz).

Evaluation

Modern electrocardiography is no longer confined to the "reading" of ECG tracings recorded on paper and the measurement of relevant time sections (intervals, segments, durations) and amplitudes by means of an ECG ruler. It now usually consists of a sequence of finely tuned electronic data processing steps: capturing the ECG lead signals; obtaining a digital representation of each recorded ECG channel by analog–digital

conversion and a special data acquisition software or a digital signal processing chip; processing the resulting digital signal by a series of specialized algorithms, which first condition the signal by removing noise, base-level variation, etc.; mathematical analysis of the clean signals to identify and measure selected time segments and amplitudes (features) for interpretation and diagnosis; secondary processing such as Fourier analysis and wavelet transform decomposition with vector feature extraction to provide input to pattern recognition-based programs; logical processing and pattern recognition, using rule-based expert systems, probabilistic Bayesian analysis or fuzzy logics algorithms, cluster analysis, artificial neural networks, genetic or evolutionary optimization algorithms, and other techniques to derive conclusions, interpretation, and diagnosis; reporting of the tracings, the data, and the conclusions drawn from the analysis with a proper sourcing of the information and the analysis steps.

ECG in diagnosing of specific conditions like, for example, acute myocardial infarction (STEMI or non-STEMI), myocardial ischemia, arrhythmia, pericarditis, etc. Brugada syndrome, J point, Osborn wave belong to ST segment evaluation. Apart from these, ECG can indicate electrolyte disturbances such as hyperkalemia, hypokalemia, hypercalcemia, hypocalcemia, arrhythmogenic right ventricular cardiomyopathy (condition known for sudden cardiac death (SCD) in young individuals); WPW syndrome can be diagnosed noninvasively by ECG only. The available ECG algorithms differ in their complexity and accuracy, and recent analyses have shown lower accuracy rates than those initially reported by the original authors. In clinical practice, a few electrocardiographic features could be helpful to predict the accessory pathway (AP) site and are summarized in the following table:

	Right-sided AP	Left-sided AP
	Precordial transition (R/S > 1) is ≥ V2[a]	R/S in V1 > 0.5 (early transition)[a]
	Shorter P-Delta interval (P on delta sign)	Longer P-Delta interval Decreased
	Increased preexcitation degree[b]	preexcitation degree[b]
Free wall versus	Late R/S transition (> V3)	Negative Delta in I, aVL
Septal location	Early R/S transition (≤ V3)	Positive Delta in I, aVL
Superior versus	Dominant positive Delta in II, III, aVF	
Inferior location	Dominant negative Delta in II, III, aVF	
QS pattern in V1 for a supero-septal location is suggestive of a **Para-Hisian** AP QS pattern in II for an infero-septal location is suggestive of **sub-Epicardiac** AP[c]		

[a]Maybe be inaccurate with very subtle or minimal preexcitation

[b]May vary depending on properties of the intrinsic normal AV conduction system

[c]Coronary sinus branch or diverticulum

These ECG algorithms and criteria may be limited or inaccurate in some subgroups of patients and specific AP types including: (1) pediatric and congenital heart disease; (2) abnormal patient stature or extreme heart rotation; (3) minimal preexcitation degree; (4) multiple APs; (5) slowly conducting APs (Mahaim family).

Furthermore, it is important to emphasize that ECG prediction of the AP site considers only the ventricular insertion site of the AP fiber. Consequently, this site might not always coincide with the successful ablation site since that many of APs may have some degree of oblique course. Finally, though 12-lead ECG is a valuable noninvasive tool to predict the AP site in WPW patients, intracardiac mapping during the electrophysiologic study and successful catheter ablation remain the gold standard for AP localization.

Critical Assessment of the Method

Although the basic principles of electrocardiography are well known, there is an obvious need for standardization. However, such information usually relates to the conventional recording and interpretation of the ECG signals. In contrast, there is little guidance with regard to the complex electronic data processing that is now inherent to state-of-the-art electrocardiography.

Most ECG devices also print a single- or multichannel signal record on paper. Only such (signed) hardcopy record may be accepted as reliable source documentation. However, caution is indicated since many devices use thermopaper, which generally rapidly fades. The date/time stamp of such devices is usually not reliable since it can be easily accessed by the operator and/or is not automatically synchronized with a reliable time server.

Also, most modern ECG devices provide for an automated analysis of relevant ECG intervals (RR, PQ, QRS, QT) usually based on the averaged signals of a 10-s recording. Such analyses are often judged to be less reliable. This prejudice is unjustified in healthy subjects with mostly normal ECGs: there is generally good agreement between automated and manual analyses; possibly gross differences between automated and manual analyses in healthy subjects mostly relate to either artificial or electrophysiological signal distortions (such as U-waves) that can be easily identified if the tracings are appropriately reviewed by an experienced analyst.

Most analyses, whether automated or manual, are subject to the constraint that it may prove difficult to identify the start of the Q-wave; for this reason, the atrioventricular conduction interval is often reported as PR- instead of the PQ-interval; the PR-interval does not extend from the start of the P-wave to the R-peak, but to the intersection of the iso-electricity ("zero") line with the upstroke of the R-wave. The "PR"-interval thus represents a simplification of the "PQ"-interval whenever the start of the Q-wave is not expressed or cannot be measured reliably. This simplification is highly convenient since it is far more easily standardized and/or automated. It is noteworthy that such a simplification is not also generally adopted for the QT-interval: the QT-interval represents the sum of the ventricular depolarization and repolarization, of which the former is relatively constant, less subject to drug effects, and less likely to be of arrhythmogenic relevance. The measurement of the QT-interval relies on two fiduciary points: the start of the Q-wave and the "end" of the T-wave; both are not sharply expressed; the precision of the estimated repolarization duration could be improved by measuring the "RT"-interval, i.e., from the peak of the R-wave to the end of the T-wave; the former fiduciary point is more easily detected, standardized, and/or automated. Automated ECG analysis usually also reports a clinical "diagnosis" of the condition reflected by the ECG based on the rhythmicity and contour of the ECG cycles using either medical or stochastic algorithms. The ECG contour is stereotypic, and deviations from a "normal" morphology may indeed reflect a more or less specific anomaly of cardiac rhythmicity and ectopism; sinus node pacemaker autonomy and function; intra-atrial, atrioventricular, intra- and transventricular signal spread; myocardial mass; myocardial depolarization and repolarization; myocardial energy balance; etc. Nevertheless, no automated diagnosis should be accepted unless reviewed, confirmed, and/or amended by an experienced electrocardiographist.

Relevant electrocardiographic time intervals and signal durations are affected by heart rate (HR) variations: the AV-nodal conduction time and the PQ/PR-interval shorten with increasing heart rate and this fluctuation may be used as an index of autonomic function. The HR-dependency of the QT-interval is well known and has resulted in several approaches to "correct" the QT-interval for HR below or above 60 bpm, according to Bazett, Fridericia, Framingham's regression; however, these corrections apply a population mean correction factor for all subjects while there is convincing evidence for significant interindividual variability in the HR–QT relationship implying that the best HR correction for QT should be estimated for each individual. This is hardly feasible since it requires a number of "normal" QT measurements at varying HR for each subject; normograms have been proposed to solve this problem. However, rather than to "correct" for HR-variations, there might be interest in investigating the disparity of the RR–QT relationship as a more sensitive index of arrhythmogenic risk. Data from the International Long QT Syndrome Registry indicate that the probabilistic risk of

developing malignant arrhythmias in patients with QT prolongation is exponentially related to the length of the QTc interval, but it remains unclear whether a QT-prolongation predominantly related to HR (i.e., with normal QTc) would be without risk.

The time course of experimental ECG criteria reflects time effects both related and unrelated to the investigational medication. Assuming an additive response model, this results in two important steps in the management of such data: to consider the data both untransformed (U) and as arithmetic changes (D) from predose baseline and to match the courses of these U- and D-data for the time course of the respective criteria during a medication-free control day ("time-matching"). Such time-matching using an extra control day within each treatment (placebo, therapeutic dose, supratherapeutic dose, active control) is costly and the need thereof is controversial.

In the setting of the ICH E14-Guideline, an investigational medication is accepted to be without QT/QTc-effect if the upper bound of the one-sided 95% confidence interval for the largest time-matched mean effect (i.e., of the changes from predose baseline relative to placebo) of the drug on the QTc interval excludes (i.e., is smaller than) 10 ms; the study is normally conducted in healthy volunteers investigating both a therapeutic and a (widely) supratherapeutic dose relative to a positive (active) and a negative (placebo) control in an experimental setting stringently powered to exclude an effect on the QTc interval exceeding 5–10 ms. This has been subject to extensive critique also because of well-founded biostatistical concerns and since a possible effect compartmentalization is not accounted for. When the largest time-matched difference exceeds this threshold, the study is termed "positive." A positive study does not imply that the drug is pro-arrhythmic but influences the evaluations that need to be carried out during the further stages of drug development.

Most ECG systems operate as closed "black-boxes" with device-specific file formats and often nonpublic analysis algorithms. There have been several efforts to develop unified, platform- and device-independent solutions.

Vectorcardiography

Purpose and Rationale

The single equivalent dipole theory that has been developed in terms of body surface potentials, the electric field produced by heart muscle, can be represented at any instant by a single equivalent dipole, and this in turn by a mean instantaneous spatial vector; and the voltage registered in any given lead is directly proportional to the projection of the instantaneous vector on the axis of the lead but is inversely proportional to the cube of the distance between the dipole and lead electrode. The vector concept and its implications were known to Einthoven and his contemporaries as well.

Procedure

Exactly as standard ECG procedure with 12 lead system electrodes. The reference frame used to indicate descriptively the orientation of the vector loops in a body is that recommended by American Heart Association Committee as frontal, transverse (horizontal), and sagittal plane vectorcardiograms.

Evaluation

The spatial vectorcardiogram (VCG) (Fig. 1) may be considered to originate at the center of an imaginary cube and as projected to the sides of the cube to represent the frontal, transverse, and sagittal planes.

Critical Assessment of the Method

Analysis of the VCG should be done in a systematic manner if one is to obtain consistent interpretations. Just as the routine interpretation of ECG tracing, vectorcardiographic analysis requires observation of the P, QRS, and T loop voltage and direction of inscription. The conventional planar VCG records the projected image of the cardiac spatial vector on the frontal, transverse, and sagittal plane. Vectorcardiography is superior to ECG in some situations where acute

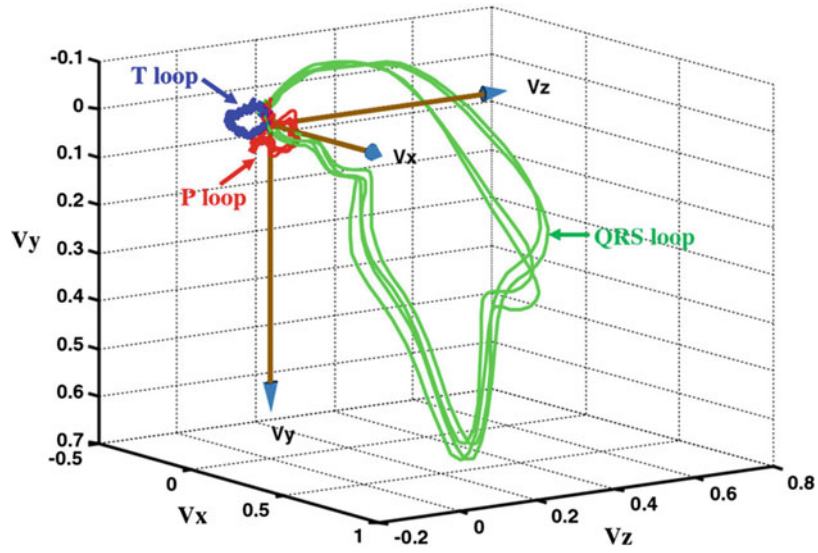


Fig. 1 VCG plot showing vector loops for P, QRS, and T wave (Reprinted from Yang H, et al. Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals. Biomed Eng Online. 2012; 11:16 under license to BioMed Central Ltd. and CC BY License)

myocardial infarction (AMI) could be overseen due to existence of left bundle branch block; ventricular hypertrophy of left or right origin, atrial hypertrophy, some congenital heart disease and valvular heart diseases; also some pulmonary states: cor pulmonale, emphysema, and chronic obstructive pulmonary disease (COPD).

Signal Averaged ECG: Late Potentials

Purpose and Rationale

The high-resolution electrocardiogram (ECG) is defined as a body surface electrocardiographic recording that registers cardiac events not seen in the standard ECG. This is usually done by increasing both the time and the voltage scales of the recording instrumentation. However, as the ECG signal is amplified, there are sources of noise that can obscure very small cardiac signals. There are several possible sources of interfering noise, but the most significant of these noise signals are the electromyographic (EMG) signals from the skeletal muscles.

Computer-based methods can be used to decrease the effects of interfering noise signals. The most common method is known as signal averaging. Hence, the term signal-averaged ECG or SAECG is often used interchangeably with the term high-resolution ECG.

Procedure

There have been several reviews of SAECG in the literature. In addition, there was a combined American Heart Association, American College of Cardiology, and European Society of Cardiology Task Force SAECG report that was published in each group's respective journal. A more recent report by an expert committee of the American College of Cardiology also provides guidelines for clinical use. A technical information report from American Association for the Advancement of Medical Instrumentation was published in 1998 that specifies the technical characteristics of SAECG systems.

Evaluation

SAECG acquisition and analysis are based on the recording of three leads in an anatomically orthogonal configuration and are referred to as XYZ leads, similar to the coordinate axes used in geometry. The most notable of these systems is the Frank lead system, which uses a resistor weighing network and an extra lead position to form its XYZ lead set. Usually 250 accepted sinus beats (QRS complexes) are analyzed, but this can be modified as lesser or greater values. Noise should be between 0 μV and 0.3 μV for 200–300 beats acquisition analysis.

Critical Assessment of the Method

There are three parameters derived from vector magnitude after filter-processed noise is accepted for evaluation. QRS duration as the first one is frequently referred as fQRS or QRSd. It is the distance between QRS onset and QRS offset as measured on timescale in milliseconds (ms). Abnormal QRS duration lies between 110 ms and 120 ms. The other two parameters: Root Mean Square voltage (RMS) and Low-Amplitude Signal duration (LAS) rely primarily on the QRS offset. They both are obtained from the filtered vector magnitude. The focus of those two parameters is waveform of late potentials as a low-level "tail" adjunct to final oscillation of QRS (QRS offset). The threshold of RMS for abnormal values is less or equal to 20 mV, whereas for the LAS values greater than or equal to 20 ms are considered abnormal. If any of the mentioned parameters is abnormal, SAECG is considered positive (abnormal).

The QRS duration is a measure of total ventricular activation time (VAT). VAT is the time from the earliest ventricular activation to the time of latest ventricular activation. Multiple times magnified, in high-resolution mode it delineates the termination of the low-level late potentials.

The Root Mean Square voltage (RMS) is usually seen as shaded region on RMS time function depicting last 40 ms from QRS offset. The RMS voltage of this value is 20 mV. Thus, RMS represents voltage based on time duration after final oscillation of QRS.

The Low-Amplitude Signal Duration (LAS) is a duration based on a voltage measurement at the end of QRS complex. A 40 mV voltage is the most commonly used reference point.

Role in Pharmacodynamics

The innovative therapeutic could impose different effects in human, among which late potentials can be successfully used to identify its capacity to induce proarrhythmic effects in pathophysiological millieu. This is of potential vital interest as for establishing dose-dependent liaison or possible drug interaction contraindications.

Heart Rate Variability (HRV) 5′ Test and 24 and/or 48 Hour

Quantifying the amount of autonomic nervous system (ANS) activity in the human body provides an insight of disease severity in a vast scale of diseases. Heart rate variability (HRV) is calculated from either short-term or 24-h electrocardiograms being an ideal way to predict ANS activity, while giving the open sight into pharmacodynamics of tested drug.

Purpose and Rationale

HRV analysis is based on the RR interval time series, the sequence of intervals between successive fiducial points of R peaks of QRS complexes in the electrocardiogram. Noteworthy, RR intervals are not equally sampled continuous signals, but rather event series on timeline. There are numerous methods and approaches for time-series analysis, some of them being linear and nonlinear.

Procedure

The duration of the recording depends on method used for HRV analysis, but usually 24 h or 48 h time series are adequate for analysis. Also, the aim of the study is detrimental of time period needed as well as stationary issues. Frequency domain methods (Fig. 2) are of short duration of 5–20 min. Typically, nonlinear methods are preferred for short-term measurements (Fig. 3). Electrodes are placed as for standard ECG recording, but special software is available for signal analysis. Patient is supine and then abruptly sitting up. Alternatively, patient is supine, sitting, standing, and squatting for 5 min each, while HRV is being analyzed. For the time domain analysis, it usually takes 48 h as obligatory sample.

Evaluation

In practice high frequency (HF) component domain can be used as a measure of parasympathetic tone and vagal activity, only if respiration is not forced or withheld. Low frequency (LF) component domain is

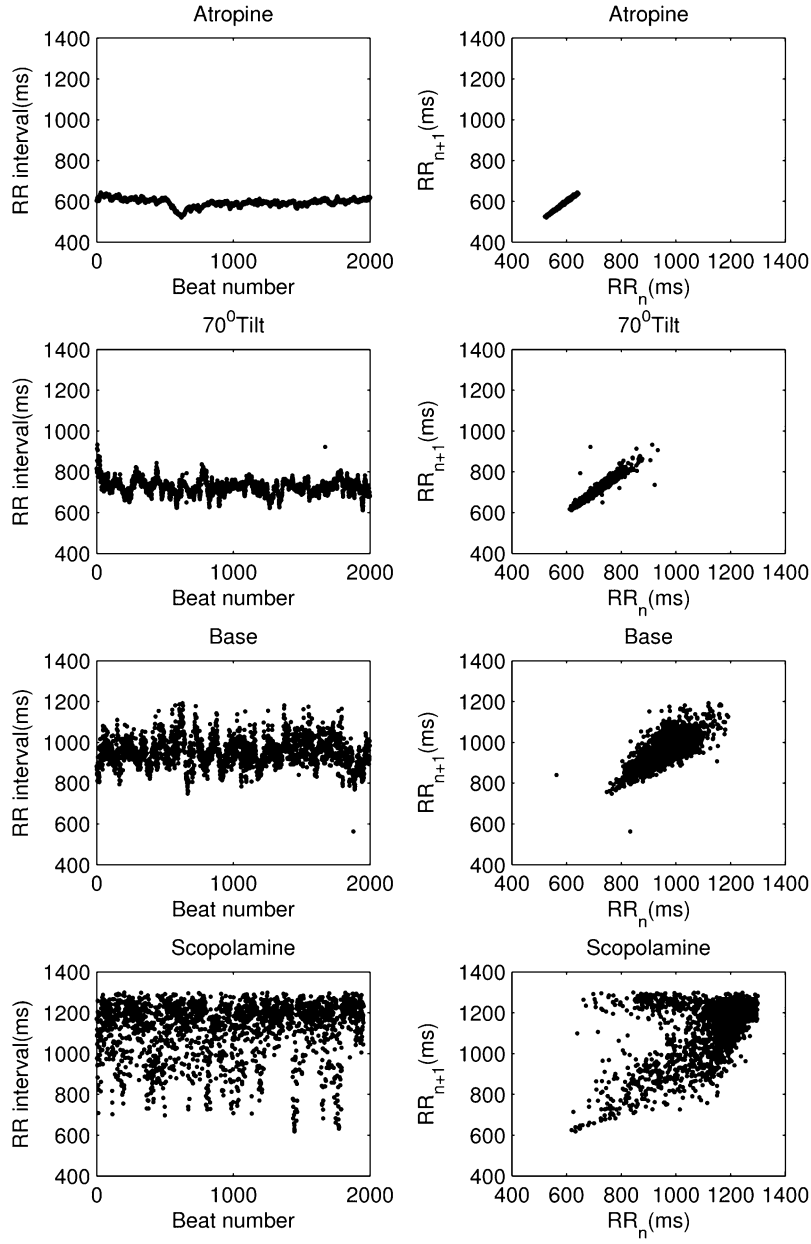


Fig. 2 RR intervals and Poincaré plots during autonomic perturbations. RR interval time series for single subject from all four phases of study with corresponding Poincaré plot (Reprinted from Karmakar CK, et al. Sensitivity of temporal heart rate variability in Poincaré plot to changes in parasympatheticnervous system activity. Biomed Eng Online. 2011; 10:17 under license to BioMed Central Ltd. and CC BY License)

considered to represent sympathetic measure. Power spectral density (PSD) (Fig. 4a, b) decomposes the signal into LF and HF, by two commonly used methods: Fast Fourier transformation (FFT) and autoregressive modeling (AR). Many other mathematical approximates and calculations can be achieved through this, but it is out of the scope of this chapter.

Critical Assessment of the Method

The area under the power spectral density (PSD) curve is divided into three frequency bands: HF, LF and VLF (very low frequency). These parameters depend greatly on the equipment stability and surrounding noise. Since some of the manual filtering of derived signal can modify the outcome results, it may influence true positive and true negative test outputs.

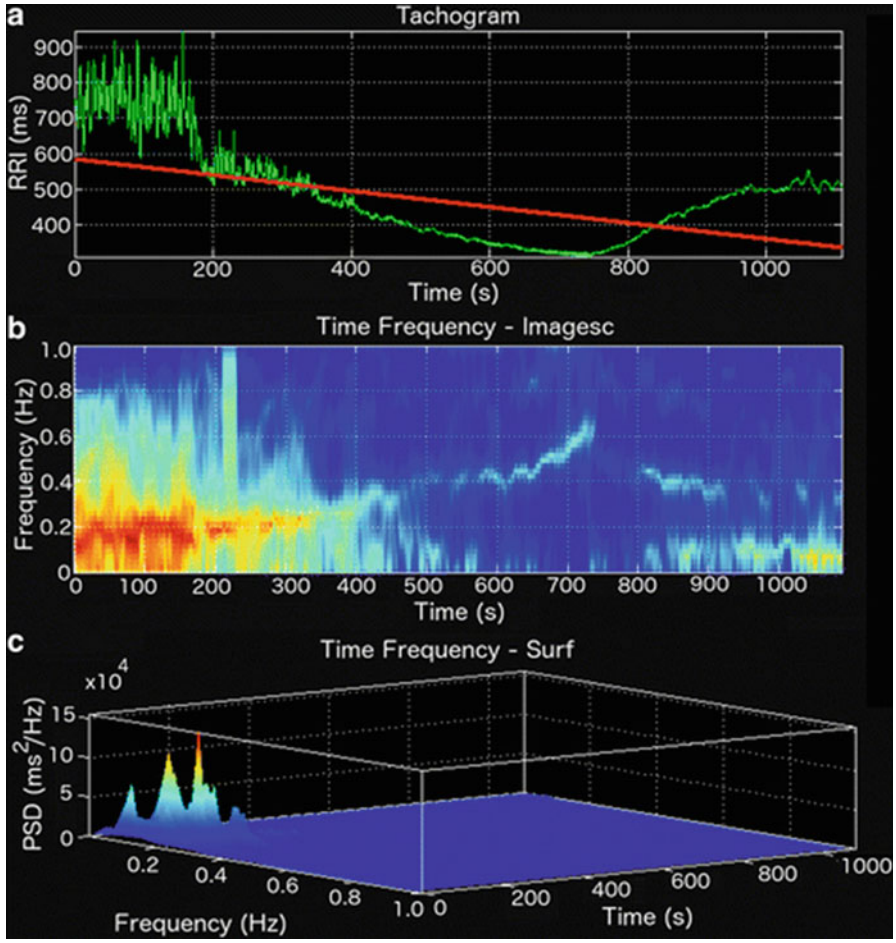


Fig. 3 Non-stationary RR interval (RRi) series during a maximal effort exercise (**a**) and the resulting scaled colors map (**b**) and surface plot (**c**). *PSD* power spectral density (Reprinted from Bartels R, et al. SinusCor: an advanced tool for heart rate variability analysis. Biomed Eng Online. 2017; 16(1):110 under license to BioMed Central Ltd. and CC BY License)

Holter Monitoring 24 Hour ECG

Purpose and Rationale

Since the invention of the continuous ECG monitor in 1961 by Norman J. Holter, the methodologies and applications of continuous recording of the ECG have evolved tremendously. The pioneering work of Bruce Del Mar led to the first commercially available continuous ECG in 1962, and the methodologies have become refined to the degree that the devices now are very light and use solid-state memory to record up to a week's worth of continuous ECGs. The original Holter monitors were primarily used to detect disturbances in the cardiac rhythm, but early studies investigated the presence and significance of ST-segment depression. There are three categories of ambulatory ECG monitors:

1. *Continuous monitors* store the heart's electrical signals for the entire time the patient wears the device. Continuous monitors have two types:
   (a) *Short term*, known as *24-h or 48-h Holter monitors.*
   (b) *Long term*, which can record for more than 48 h. In recent years, new technology has allowed ambulatory ECG monitors to have more memory while still being small and

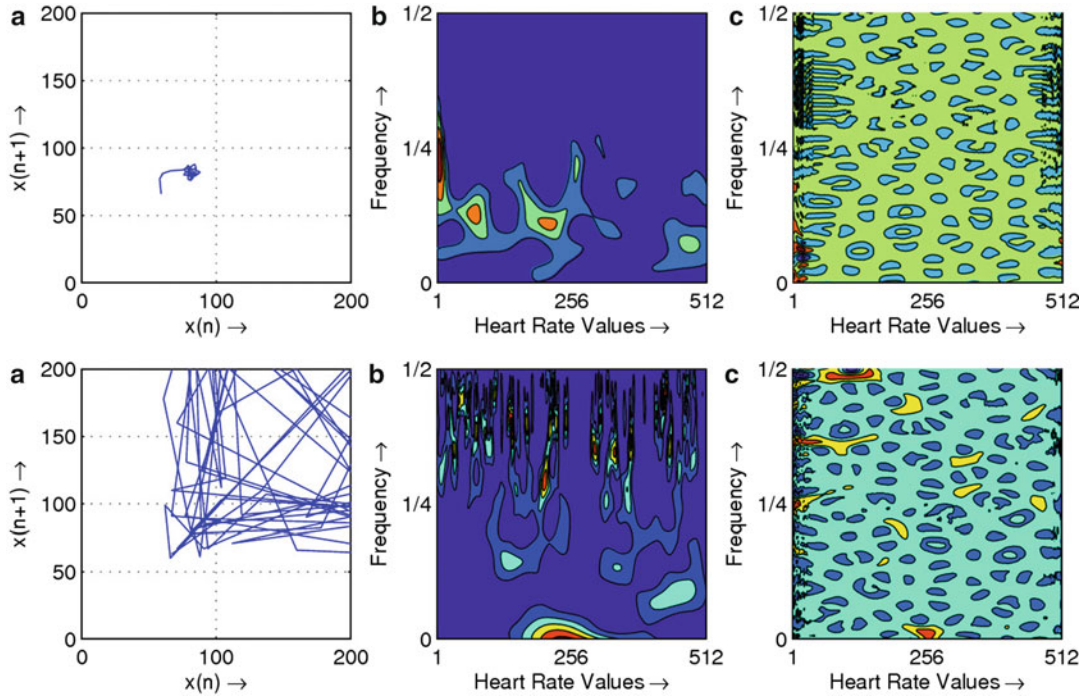


Fig. 4 (**a**) Heart rate in representative subject with Normal; (*a*) Phase space plot (*b*) Scalogram (*c*) Wigner-Ville distribution (Reprinted from Faust O, et al. Analysis of cardiac signals using spatial filling index and time-frequency domain. Biomed Eng Online. 2004; 3(1):30 under license to BioMed Central Ltd. and CC BY License). (**b**) Heart rate in representative subject with AF; (*a*) Phase space plot (*b*) Scalogram (*c*) Wigner-Ville distribution (Reprinted from Faust O, et al. Analysis of cardiac signals using spatial filling index and time-frequency domain. Biomed Eng Online. 2004; 3(1):30 under license to BioMed Central Ltd. and CC BY License)

lightweight; these are known as *efficient-memory Holter monitors* and *patch monitors* (designed without the wires connecting electrodes to the recorder).

2. *Intermittent long-term monitors* store the heart's electrical signals only when the monitor is triggered by a patient or by abnormal heart rhythm. These monitors also have two types:
   (a) *Event monitors*, also known as post-event recorders, which typically store 5–7 min worth of data from the moment triggered.
   (b) *Cardiac loop recorders*, which continuously record new signals, erase old signals and lock in data when triggered. They typically store 1–4 min worth of data. Loop recorders can be either *external*, worn around the waist or wrist, or *insertable* (also known as *implantable*), implanted under the skin in the left parasternal region (near the heart).
3. *Real-time cardiac telemetry systems*, also known as *mobile cardiac outpatient telemetry*, are similar to long-term continuous monitors but can send the data directly to a central monitoring station instead of recording it to be downloaded later.

Procedure

In recent years, innovative engineering and advances in manufacturing have hastened the development of miniaturized medical devices and yielded a variety of cardiac monitors for ambulatory use (Fung et al. 2015). These recently developed wearable, "on-body" ambulatory devices have integrated microelectronics for short- to medium-term (days to weeks) monitoring and are challenging conventional, widely used devices from the last decades that were limited to wearable multi-lead 24−/48-h Holter monitors and event recorders. Further on the

pioneering front, very short-term (seconds to minutes) handheld smartphone-enabled systems are beginning to reshape the field of mobile cardiac monitors as well as the clinician-patient interface. These systems require attachment of an electrode-embedded module to a smartphone that detects electrical impulses from the user's fingertips and transmits signals to the mobile device to generate continuous single-channel ECG for the duration of the contact between the fingers and the sensor. Adhesive Ambulatory ECG (AECG) patch devices typically comprise a sensor system, a microelectronic circuit with recorder and memory storage, and an internal battery embedded in a relatively flexible synthetic matrix, resin, or other material. They are usually intended for medium-term use ranging from days to several weeks, depending on the device. The self-contained adherent unit typically has a low profile and can be affixed to the body surface, usually over the left upper chest area, by means of prefabricated adhesive material.

The main advantages of this kind of AECG system are that they are easy to use, leadless, minimally intrusive to daily activities, water-resistant, hygienic (i.e., single use only), and incur no upfront cost to the clinic for the initial device investment as compared to the wearable, reusable devices (Gulizia et al. 2016). Because of easy application of the adhesive AECG patch to skin and its unobtrusive maintenance-free nature, they have a high study completion rate, implying a high acceptance rate (long wear time) that should translate into improved compliance compared to other short- to medium-term devices such as the Holter monitor.

Evaluation

Guidance has been provided for the continuous ECG monitoring in several clinical settings.

Originally, Holter-ECG analysis was mainly focused on rhythmicity (sinoatrial dysfunction, ectopism, atrial fibrillation, atrial flutter, paroxysmal tachycardia, accelerated rhythms with normal or aberrant configuration), atrioventricular conduction delays and blockade, intermittent changes in QRS-morphology (parasystoles, ectopic rhythms), etc. Improved algorithms also now provide for the analysis of changes of the QT-wave and ST-Twave.

Several electrocardiographic-based methods of risk stratification of sudden cardiac arrest have been studied, including QT prolongation, QRS duration, fragmented QRS complexes, early repolarization, Holter monitoring, heart rate variability, heart rate turbulence, signal-averaged ECG, T wave alternans, and T-peak to T-end (Garcia et al. 2011). These ECG findings have shown variable effectiveness as screening tools (Verrier and Ikeda 2013).

Critical Assessment of the Method

One of the important advantages of continuous ECG recordings is that it collects a vast amount of data under real-life conditions; this permits beat-to-beat analysis for a far more precise and valid interpretation of the QT-related arrhythmogenic risk (Luebbert et al. 2016). Indeed, dynamic beat-to-beat QT interval analysis compares the QT interval to individual cardiac cycles from all normal autonomic states at similar RR intervals, thus eliminating the need for correction functions; in this way, beats with QT intervals exceeding a critical (subject-specific) limit can be flagged as outlier beats for further arrhythmia vulnerability assessment (Coris et al. 2016). Furthermore, such beat-to-beat techniques can also be used to assess the QT–TQ interval relationship known as ECG restitution (Olsen et al. 2015).

Further procedures allow evaluation of highly sensitive prognostic criteria, such as QT-dispersion, heart rate variability, and heart rate turbulence; other methods specifically conceived to quantify arrhythmogenic risk are under development (Abdelghani et al. 2016).

Additionally, continuous ambulatory electrocardiography provides for a better characterization of the diurnal variability and the implications thereof for the timing of drug administration.

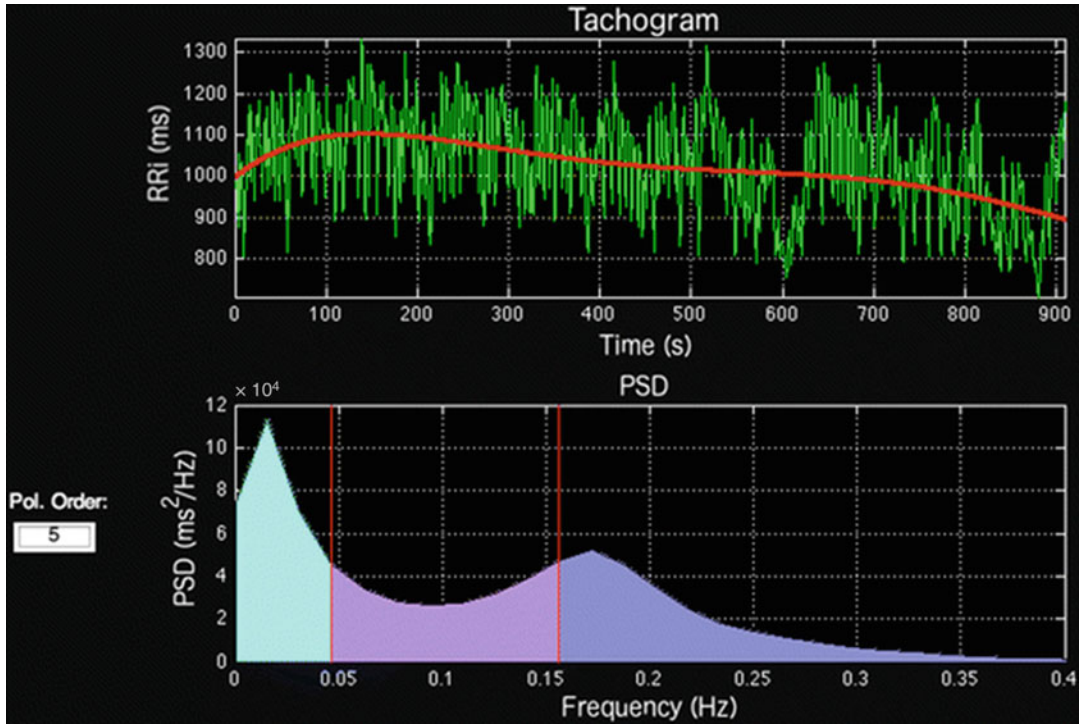


Fig. 5 Example of a RR interval (RRi) series during rest period, detrended with a 5 degree polynomial (*upper panel*) and the corresponding power spectral density (PSD) function estimated with Welch's method (*lower panel*) (Reprinted from Bartels R, et al. SinusCor: an advanced tool for heart rate variability analysis. Biomed Eng Online. 2017; 16(1):110 under license to BioMed Central Ltd. and CC BY License)

Turbulence Onset 24 Hour Holter Monitoring ECG

Purpose and Rationale

Heart rate turbulence (HRT) is a phenomenon that was discovered by Georg Schmidt's group in mid-1990s in Munich. HRT is defined by minute changes in ventricular cycle length following premature ventricular contractions (PVC) (Watanabe 2003). After a premature ventricular contraction, the normal response is a brief initial increase in heart rate, followed by a return to baseline. These changes are the result of premature ventricular contraction–induced hemodynamic disturbances, and the speed at which they happen ultimately provides information regarding cardiovascular autonomic function. Clinical investigations showed that patients in whom this fluctuation of sinus rhythm was absent showed a higher mortality rate. Heart rate turbulence is qualified by two parameters: turbulence onset and turbulence slope (Fig. 5). Turbulence onset is the relative change in the RR interval caused by a premature ventricular contraction, and turbulence slope is the rate of change of the RR interval back to baseline. Heart rate turbulence can also be induced through the use of intracardiac pacing performed in the electrophysiology laboratory or through an implanted pacemaker or an ICD. One contemporary protocol for measuring induced heart rate turbulence involves computing turbulence slope and turbulence onset following 10 ventricular extrastimuli with a coupling interval of 60–70% of the sinus cycle length.

Procedure

In 2008, the International Society for Holter and Noninvasive Electrocardiography (ISHNE) published a consensus statement on the standards of measurement, mechanism, and clinical applications.

Evaluation

Heart rate is influenced by myriad of intrinsic oscillations due to change in posture or activity or mental state or stress (Huikuri et al. 2009). Therefore, a plot of RR interval has a jagged stochastic appearance (tachogram). Computation of HRT uses PVC as anchor point. The PVC tachogram sequence should include two sinus rhythm RR intervals before the PVC, the coupling interval and compensatory pause and 15 subsequent sinus RR intervals. All intervals not having compensatory pause or contaminated by PVC are excluded from the equation:

$$\mathrm{TO} = \frac{(\mathrm{RR}_1 + \mathrm{RR}_2) - (\mathrm{RR}_{-1} + \mathrm{RR}_{-2})}{(\mathrm{RR}_{-1} + \mathrm{RR}_{-2})} \times 100$$

Critical Assessment of the Method

Broadly speaking, HRT predictive capabilities of mortality, cardiac mortality, and also arrhythmic mortality rank with, or even exceed in some occasions, conventional linear HRV (Huikuri and Stein 2013). Limitations concern patients with atrial fibrillation, where this analysis is not feasible.

Symbolic Dynamic Analysis: Theory of Chaos

Purpose and Rationale

The analysis of the symbolic dynamics of the heart rate describes the nonlinear features of HRV. In this technique, the RR intervals are named by different symbols based on the length of the RR intervals. For shorter electrocardiographic recordings, for example, four different symbols can be used, and for longer 24-h recordings, the number of the symbols can be increased, e.g., up to six. After the definition of symbols (alphabets), words, which are from three or four successive alphabets in length and start from each successive beat, are formed. The complexity of the data time series is determined from the distribution of the words using appropriate mathematical methods (Voss et al. 1996).

Procedure

The conversion of a time series into a symbol string may be done using several methods (Fig. 6). The first one divides symbol into two or more value ranges, depending on how many symbols we wish to utilize. Value ranges can be absolute bands or based on signal averages or standard deviation (SD), for example, A, B, C, D, and sequence like: ABCCDABAACBDACDCCBDABBADDCACC. The shape of distribution may itself act as a basis of further analysis, but it is also possible to measure the order related to the distribution in the terms of entropy. The simplest such measure is Shannon's entropy.

Evaluation

This method of symbolic dynamics is a useful approach for classifying the dynamics of HRV. By means of this method, the inner motions of the time series can be investigated (Gimeno-Blanes et al. 2016). Parameters of the time and the frequency domain often leave these dynamics out of consideration. In comparison with all other methods of nonlinear dynamics (NLD) for HRV analysis, symbolic dynamics is the method with the closest connection to physiological phenomena and is relatively easy to interpret.

Critical Assessment of the Method

HR fluctuations can be analyzed using many different methods and approaches. No single method is clearly superior to other techniques. The physiological interpretation of the results is often difficult especially in the case of nonlinear methods, because the unpredictable portion of the HR fluctuation can be due to chaotic behavior and/or stochastic component. The basic idea behind stochastic modeling is that the unpredictable component is not a perturbation but an essential part of the dynamical behavior of the system. However, symbolic dynamics gives a solid basis for Shannon entropy, i.e., with potent modulation analysis in pharmacodynamics.

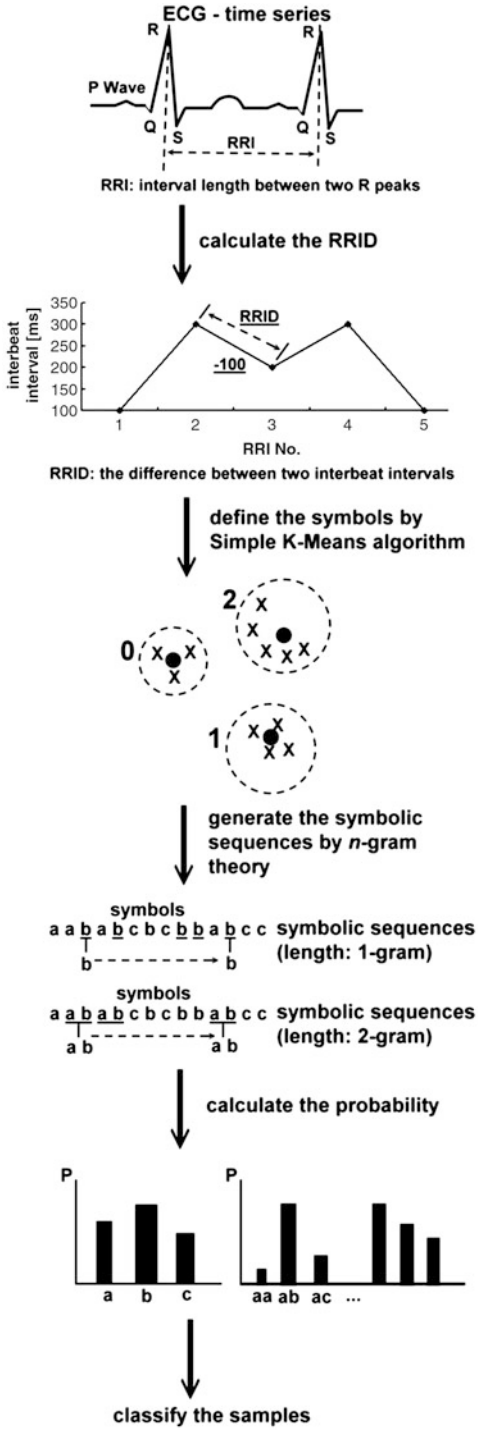


Fig. 6 Concept flow chart of the AIIA method (Reprinted from Huang YC, et al. Using n-gram analysis to cluster heartbeat signals. BMC Med Inform Decis Mak. 2012; 12:64 under license to BioMed Central Ltd. and CC BY License)

Nonlinear Indexes of Cardiovascular Variability

The nonlinear theory has been growing among physiologists and physicians aiming to explain the workings of biological phenomena, highly complex, dynamic, and interdependent, where the system behavior differs from the behavior of its parts or elements (Haugaa et al. 2010).

The exponent of power-law, approximate entropy (ApEn) analysis, and detrended fluctuation (DFA) are nonlinear methods recently introduced to the study of HRV.

Entropy is a measure of randomness or disorder, as included in the second law of thermodynamics, namely the entropy of a system that tends toward the maximum. Different states of a system tend to evolve from ordered configurations to less organized settings. Referring to the time series analysis, the ApEn provides a measure of the degree of irregularity or randomness within a series of data. Entropy was originally used by Pincus (1991) as a measure of system complexity, where smaller values indicate greater regularity, and higher values lead to more disorder, randomness, and complexity of the system. For instance, with a drop in the ApEn, heart rate becomes more regular with age in both men and women.

The DFA is a technique that characterizes the variation pattern through measuring scales. DFA has been specifically developed to distinguish between intrinsic fluctuations generated by the complex system and those caused by external or environmental stimuli acting on the system. The variations that arise due to extrinsic stimulation are presumed to cause a local effect, while the intrinsic variations due to the dynamics of the system are assumed to exhibit a long-term correlation.

The analysis of the *Poincare* plot or Lorenz plot is considered as based on nonlinear dynamics by some authors. The *Poincare* plot is a two-dimensional graphical representation of the correlation between consecutive RR intervals, where each interval is plotted against the next one, and its analysis can be done qualitatively (visually) by evaluating the shape formed by its attractor, which shows the degree of complexity

of the RR intervals, or quantitatively, by fitting an ellipse to the figure formed by the plot from where the indexes are taken: SD1, SD2, and SD1/SD2 ratio. SD1 represents the dispersion of points perpendicular to the line of identity and appears to be an index of instantaneous beat-to-beat variability (i.e., the short-term variability which is mainly caused *by respiratory sinus arrhythmia*), while the SD2 represents the dispersion of points along the line of identity and it characterizes long-term HRV. The SD1/SD2 ratio shows the relationship between short- and long-term RR interval variations. Despite the fact that Poincare plot is primarily considered a nonlinear technique, it has been shown that SD1 and SD2 can be obtained as a combination of linear time domain HRV indexes. Therefore, alternative measures are still needed to characterize nonlinear features in Poincare plot geometry.

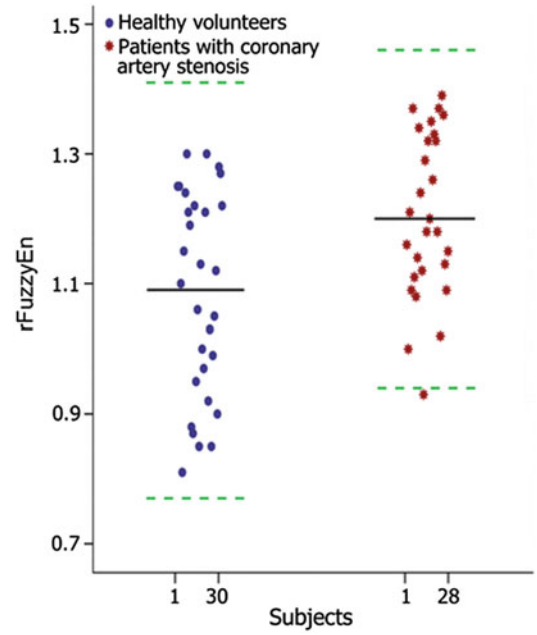


Fig. 7 The Bland–Altman plot of rFuzzyEn of DPV for all subjects in the two groups. The *black solid line* indicates the mean and the *green dotted line* indicates the *upper* and *lower* bounds of the mean ± 2 SD, respectively (Reprinted from Ji L, et al. Analysis of short-term heart rate and diastolic period variability using a refined fuzzy entropy method. Biomed Eng Online. 2015; 14:64 under license to BioMed Central Ltd. and CC BY License)

Fuzzy Logic Concepts

The possibility of using mathematical methods and theories for data analysis has opened up a range of possibilities for the study of pathophysiological behaviors of cardiovascular variability. Large volume of data can be more easily assessed and analyzed with fuzzy logic. In order to better understand the onset and development of important pathologies, the autonomic nervous system activity can be explored through dynamical fuzzy logic models (Fig. 7), such as the discrete-time model and the discrete-event model. Fuzzy logic approaches are able to perform nonlinear mapping or predictions involving more than one cardiovascular parameter and to explore possible relations among these parameters, which normally would not be considered as a possibility. Fuzzy logic represents a flexible system that adequately describes nonlinear and complex systems since the resulting function can be written as a weighted linear combination of the system inputs and, therefore, it can resemble a nonlinear function as needed. For this reason, fuzzy logic methods are a feasible solution to consider in the absence of prior mathematical description between input-output variables.

Considering the Sugeno Fuzzy Logic formulation, the system output z can be modeled from

$$z = \sum i = 1\mathrm{Nwizi} \sum i = 1\mathrm{Nwi}, z = \sum i = 1\mathrm{Nwizi} \sum i = 1\mathrm{Nwi},$$

where N corresponds to the number of fuzzy rules and $zi = \sum j = 1\mathrm{naixj} + \mathrm{cizi} = \sum j = 1\mathrm{naixj} + \mathrm{ci}$ is a linear combination of the system inputs $x_j, j = 1, \ldots n$. The rule weights are obtained as $\mathrm{wi} = \Pi j = 1\mathrm{n}\Gamma Fij(xj) wi = \Pi j = 1\mathrm{n}\Gamma Fji(xj)$ where $\Gamma_{F^i_j}$ is the membership function of rule i and input x_j. Although membership functions may assume different shapes, the Gaussian function is rather a popular choice in the literature due to its symmetry and dependence on mean and variance, which correspond respectively to the center and the width of the membership function.

Fuzzy logic has the singular characteristic to combine empirical knowledge (described as linguistic rules) and knowledge directly extracted from the data, enabling an easier way to interpret the outcomes in a physiological perspective. This mathematical model may be a reliable method to evaluate the influence of the autonomic nervous system over cardiovascular control in healthy and diseased subjects.

The main advantage of the use of fuzzy logic systems comes from their power to deal adequately with the uncertainty. In particular, this approach tolerates imprecise data, and it is focused on the "plausibility" of occurrence rather than the traditional binary response "0" or "1." For example, while a given measurement of a certain biological variable such as stress may convey a person as being "content," the same measurement may reveal a status of "dissatisfaction" for another one. Thus, biological variables that vary from person to person and are closely influenced by external and internal changes direct themselves toward fuzzy logic model of analysis, where the application of methods of investigation based on zero and one, true and false does not apply. Cardiovascular signals are characterized by a great intra- and inter-individual variability, besides imprecise measurements due to limited resolution of acquisition systems. Additionally, it is believed that traditional statistical methods may not capture all the information needed to describe disease in its complexity and dynamics. In this context, fuzzy logic may be a more reliable alternative to traditional methods.

Applications of Fuzzy Logics to the Analysis of Cardiovascular Variability

fuzzy logic approaches have been recently used in the cardiovascular field in different contexts including applications in signal processing and monitoring, classification, prediction, or control. One approach consists of extracting the relevant features from one or more cardiovascular signals, which are then integrated into a fuzzy logic scheme aiming at the identification of the presence or the quantification of a pathological state.

Fuzzy logic methods have been successfully integrated in control systems. For instance during anesthesia, mean arterial pressure was controlled based on the error between desired and measured values, allowing it to control the balance between the unconsciousness and the side effects caused by the hypnotic drug. Also during anesthesia, hemodynamic changes were successfully modeled considering drug dose level alterations as inputs of the fuzzy system. In hemodialysis condition, fuzzy logic has also shown to be capable of effectively control blood pressure trends, using ultrafiltration rate as input. Such a system allowed an overall reduction of 40% of the most severe episodes in hypotension-prone subjects.

Abnormal cardiac rhythms have been identified using artificial neural network and fuzzy interactions based on nonlinear heart period R-R features, such as spectral entropy, *Poincare* SD1/SD2, and Lyapunov exponent. Also based on R-R features, fuzzy logic was used for ECG beat classification to detect arrhythmic and ischemic heartbeats. Fuzzy logic approaches showed efficiency in improving oscillometric cuff pressure measurements by properly detecting outliers and noise artifacts.

With the goal of evaluating autonomic nervous system function, fuzzy logic has been used to choose the optimum subset of time, frequency, and nonlinear variables related to sympathetic and parasympathetic activities on HRV. Fuzzy logic approach has been used in a classification scheme to jointly evaluate results of several autonomic tests, e.g., head-up tilt test and active postural change, using both time and spectral analysis of heart rate and of diastolic blood pressure series. Similar fuzzy logic schemes were used for the information fusion of relevant features extracted from multimodal cardiovascular signals, such as heart period R-R and systolic blood pressure, for the detection of life-threatening states in cardiac care units.

Recently, fuzzy logic methods have been employed to effectively describe blood pressure and heart period R-R coupling and, therefore, have the potential to improve time domain baroreflex sensitivity (BRS) estimation. The

autoregressive linear analysis approach for BRS estimation has limitations when cardiovascular regulation is depressed. Liu et al. proposed a hybrid model consisting of a parallel modular structure with an autoregressive and a fuzzy logic system, to study simultaneously linear and nonlinear heart rate and blood pressure coupling mechanisms (Liu et al. 2008). This approach illustrates the utility of combining more traditional methods with fuzzy logic, which could be of advantage in diseased conditions when cardiovascular system regulation is afflicted.

Time domain BRS methods based on spontaneous data typically assume blood pressure and heart period R-R linearity and provide single slope estimation, regardless of the blood pressure value. In this context, fuzzy logic methods can contribute to establish a BRS dependent of blood pressure level, similarly to time domain blood pressure pharmacological methods. Recently, fuzzy logic has been used to analyze spontaneous R-R series as a function of blood pressure values, comparing performances in real and surrogate data.

Critical Assessment of the Method

The optimized definition and number of symbols have to be validated on larger clinical studies with more patients involved. It is necessary to check which symbol definition has to be adapted by applying symbolic dynamics to patients with atrial fibrillation.

The renormalized entropy (ReEn), as a measure of a relative degree of order, has to find stationary periods in the time series. The influence of instationarities can theoretically lead to misinterpretation due to contradictory results.

Strain Imaging on Echocardiography

Purpose and Rationale

Strain and strain rate are novel imaging techniques that measure changes in length and/or thickness of myocardial fibers. Those methods have been incorporated into routine clinical practice only since recently.

Procedure

Strain is ideally suited to quantify myocardial function regionally, but with the introduction of speckle tracking, a new parameter for global left ventricular (LV) function assessment called "global strain" has been introduced (Iwano et al. 2011). In the longitudinal direction, global longitudinal strain reflects the deformation along the entire LV wall which is visible in an apical image. The measurements from all three apical views are combined to give an average global longitudinal strain (GLS) value (Haugaa et al. 2013).

Evaluation

Strain is defined as the fractional change in length of a myocardial segment relative to its baseline length, and it is expressed as a percentage. Strain rate is the temporal derivative of strain, and it provides information on the speed at which the deformation occurs. Strain is a vector and the complete description of the complex deformation of a piece of myocardium requires three normal and six shear strain components. For practical reasons, the normal strains which are preferred for clinical use are oriented along the coordinate system of the LV; they describe radial thickening and thinning as well as circumferential and longitudinal shortening and lengthening. Lengthening or thickening of the myocardium is represented by positive strain values, whereas negative values represent shortening or thinning. The most commonly used parameter is longitudinal strain, which can be expected to be around 20% in all regions of the LV.

It must be noted that myocardial deformation is load dependent (Klaeboe et al. 2017). Therefore, strain and strain rate measurements must be interpreted considering ventricular wall thickness and shape as well as pre- and after-load.

GLS lower limit of normality has been established in −18% (Kocabay et al. 2014).

Critical Assessment of the Method

Although speckle-tracking echocardiography (STE) (Fig. 8) has significantly contributed to improve the evaluation of LV function and has the potential to improve SCD risk stratification (Leren et al. 2017), certain limitations of the

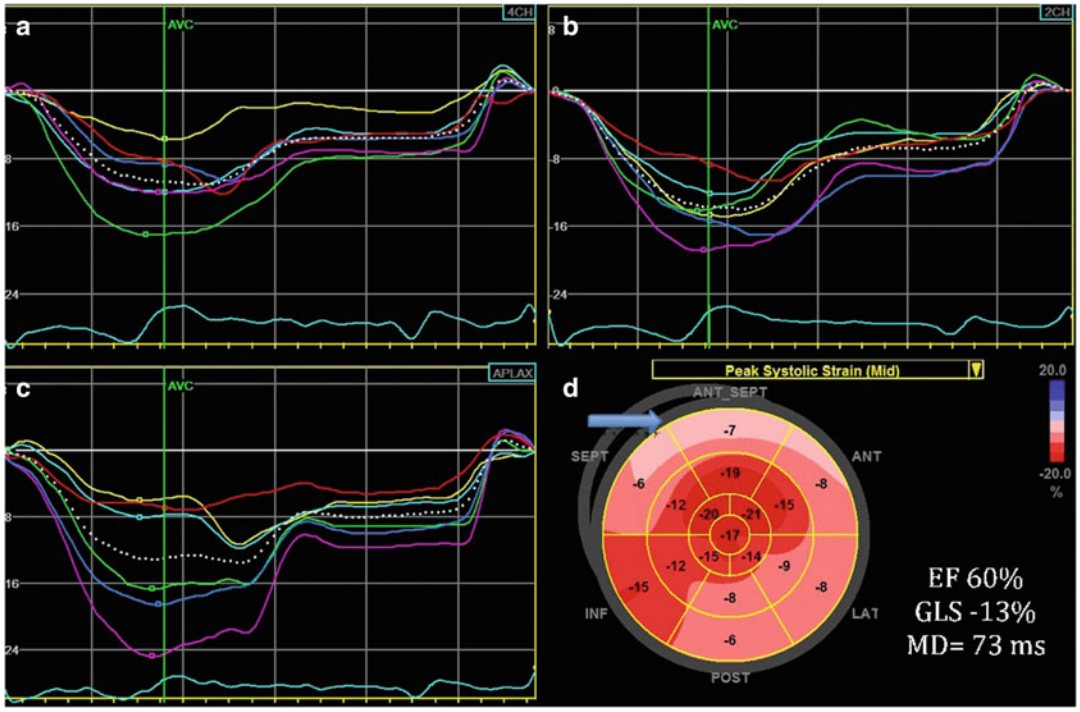


Fig. 8 Global longitudinal strain and mechanical dispersion in a patient with hypertrophic cardiomyopathy (Reprinted from: Book "Sudden cardiac death: Predictors, Prevalence and Clinical Perspectives", Chapter "The role of novel echocardiographic techniques for primary prevention of sudden cardiac death", pp. 267–86, 2017, Editor Ivana I Vranic, with permission from Nova Science Publishers, Inc. New York)

technique should be stated. First, strain has been considered a less load dependent measure; however, variations in loading conditions can lead to different results, this is important in patients with acutely decompensated heart failure in which therapy and improvement in loading conditions might lead to different values. Second, as calculations for strain-derived parameters are derived from 2D images, the presence of artifacts (shadowing, reverberations) can lead to inadequate tracking and inaccurate strain and mechanical dispersion (MD) values (Fig. 9). This is especially true when several segments are not correctly tracked. Third, interchangeability of strain among different vendors, and software vendors is also an issue important to take into account, as values have shown to be different among them. The impact of this issue regarding mechanical dispersion has not been specifically addressed; however, as tracking algorithms differ among vendors, MD might be very likely also affected by this issue. Lastly, the adequate measurement of strain needs training and results from less experienced operators differ from more experienced ones.

Myocardial Mechanical Dispersion

Purpose and Rationale

The diagnosis of mechanical dyssynchrony (Fig. 10) induced by the presence of infarction scar and/or conduction abnormalities in patients with an ejection fraction (EF) of $< 35\%$ may be associated with a greater propensity for inducing serious ventricular arrhythmia (ventricular tachycardia (VT), ventricular fibrillation (VF)) and sudden cardiac death (Claus et al. 2015). The assessment of regional myocardial function using tissue Doppler echocardiography (TDE)

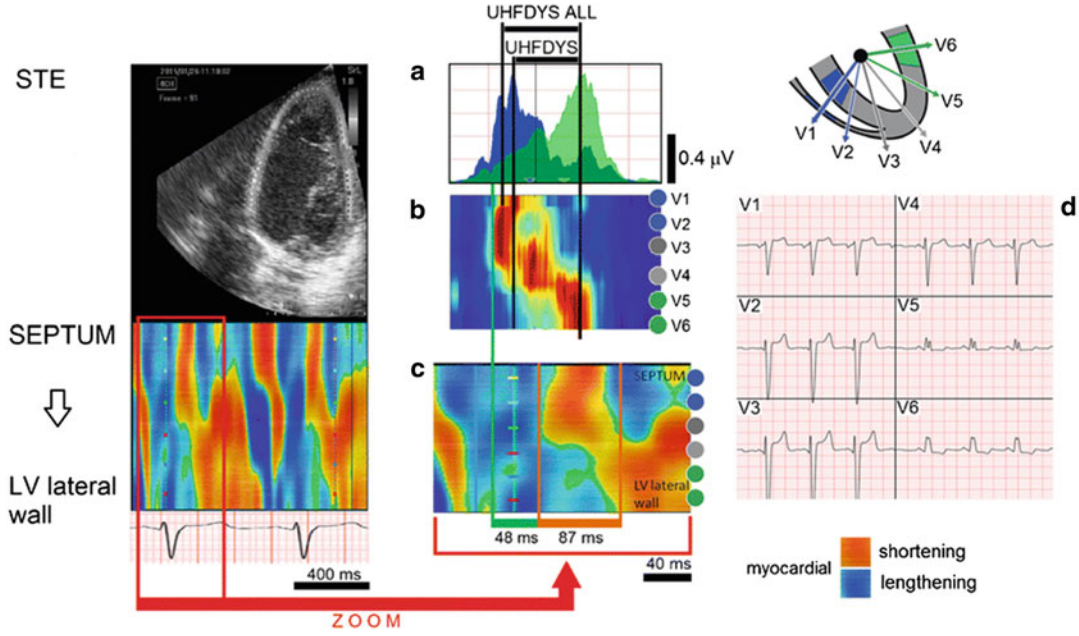


Fig. 9 Electrical and mechanical dyssynchrony coupling demonstrated by UHFQRS and speckle tracking echocardiography (STE) in patient 2 suffering from LBBB. The figure compares the UHF electrical dyssynchrony and the mechanical dyssynchrony of the septum and LV lateral wall. Myocardial shortening is coded by the *orange/red* color and myocardial lengthening by the *blue* color. (**a**) UHFQRS, V1 (*blue*) and V6 (*green*) leads. (**b**) Normalized UHFQRS map. (**c**) Detail from STE map temporally synchronized with A and B. (**d**) V1-V6 ECG. UHFDYS and UHFDYS ALL electrical dyssynchrony are 61 ms and 74 ms, respectively, – *black horizontal bars*. (**a**) The time delay of mechanical motion between the onset of myocardial deformation of the middle septum and the middle lateral wall is 87 ms – *orange horizontal bar*. (**c**) The green horizontal bar defines delay 48 ms between the first electrical UHF activation in V2 lead and onset of mechanical myocardial deformation of the middle septum (Reprinted from Jurak P, et al. Ventricular dyssynchrony assessment using ultra-high frequency ECG technique. J Interv Card Electrophysiol. 2017; 49 (3):245–254 under license to CC BY License)

allows for noninvasive analysis of the regional mechanical dysfunction (LV mechanical dispersion) (Abduch et al. 2014).

Procedure

The time to maximum myocardial shortening, including postsystolic shortening, if present, is measured from the ECG onset Q/onset R-wave in each of 16 segments of left ventricle. The maximum myocardial shortening from a representative strain curve with a shortening duration of a minimum of 50 ms is used in the time analyses. Segments in which no shortening is present are excluded. To quantify LV mechanical dispersion, the SD of the 16 different time intervals to maximum myocardial shortening is used; this parameter is defined as mechanical dispersion. An alternative measure for mechanical dispersion is the difference between the longest and shortest time interval from ECG onset Q/onset R-wave to the maximum myocardial shortening in each individual. This parameter was defined as the delta contraction duration.

Evaluation

Clinical implications of measurements of mechanical dispersion and global strain in post-myocardial infarction (MI) patients add important information about the risk of arrhythmia beyond the EF. Importantly, in patients with a preserved or slightly reduced EF, mechanical dispersion of 70 ms identified post-MI patients with an increased risk of life-threatening arrhythmias. According to current guidelines for primary

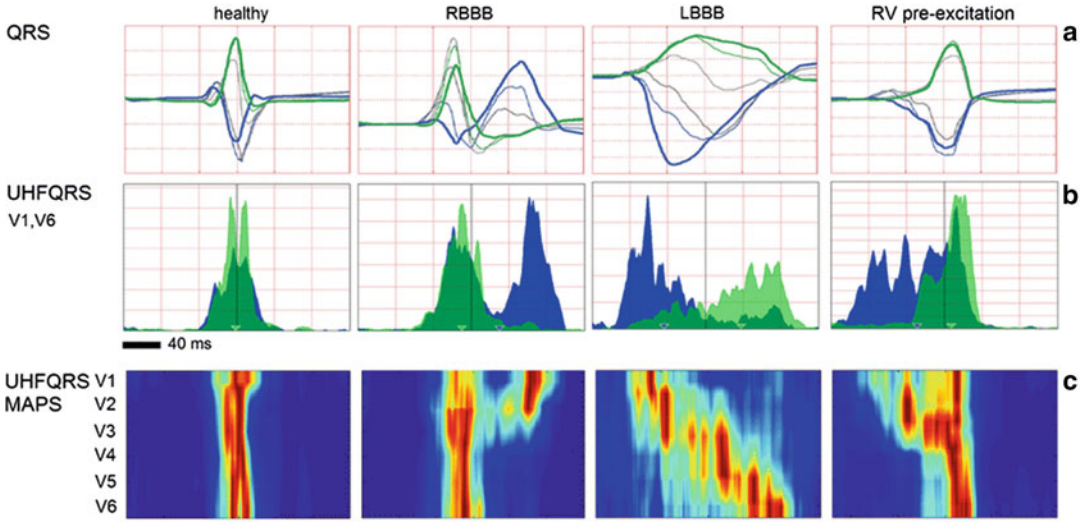


Fig. 10 Examples of different ventricular electrical activation patterns. (**a**) Averaged QRS complexes, V leads. (**b**) Averaged UHFQRS of leads V1 (*blue*) and V6 (*green*). (**c**) UHFQRS maps. From *left*: healthy heart QRSd 81 ms, patient 3 – RBBB, QRSd 139 ms, patient 4 – LBBB, QRSd 190 ms, and patient 5 with WPW syndrome with right lateral accessory pathway QRSd 105 ms (Reprinted from Jurak P, et al. Ventricular dyssynchrony assessment using ultra-high frequency ECG technique. J Interv Card Electrophysiol. 2017; 49(3):245–254 under license to CC BY License)

prevention, post-MI patients with an EF 35% should be considered for ICD therapy. The novel principles might be useful to identify the risk of arrhythmias in post-MI patients with relatively preserved EF who do not fulfill current ICD indications (EF 35%).

Critical Assessment of the Method

Future trials should investigate whether mechanical dispersion and global strain can be used to select additional patients for ICD therapy among the majority of post-MI patients with a relatively preserved EF in whom current ICD indications fail.

Systolic Function

The systole extends from the end of the late diastolic filling (closure of the mitral valve) to the start of the next isovolumetric systolic relaxation phase (closure of the aortic valve); therefore, it includes the isovolumetric contraction phase (until opening of the aortic valve) and the ejection phase(s); the right ventricle contracts first, then shortly followed by the left ventricle.

Performance and energy requirements of the heart muscle and heart pump depend on preload (ventricular filling), heart rate, afterload (the "load" that the heart must eject blood against aortic input impedance as defined by total peripheral resistance, arterial conductivity and distensibility, and wave reflections), and inotropy (load and heart rate independent performance).

Systolic Time Intervals

Purpose and Rationale

Systolic time intervals (STI) are the time equivalents of the electromechanical systolic (forward) pump performance.

Procedure

Relevant segments can be derived from the simultaneous high-speed registration of the electrocardiogram (ECG), phonocardiogram (PCG), carotid pulse mechano-cardiogram, impedance cardiogram (ZCG), or by echocardiography. Although there is some delay between central events and their peripheral reflection, this has relatively little

impact on the accuracy of the estimation of the timing of central events. The preejection period (PEP) corresponds to the duration of the isovolumetric contraction phase from the start of the ECG Q-wave up to the start of the ejection (opening of the aortic valve, between the first and second component of the first PCG heart sound); the left ventricular ejection time (LVET) from the start of the systolic ejection (end of PEP) up to the end of the ejection (closure of the aortic valve, between the first and second component of the second PCG heart sound, nadir of the carotid pulse wave, nadir of the dZ/dt-curve by ZCG, etc.); the total electromechanical systole (QS2) then corresponds to the sum of PEP and LVET.

An increase in HR shortens STIs, LVET, and QS2 in particular, whereas the PEP is less HR-dependent. Accordingly, there are numerous attempts to "correct" STI for HR (STI_c).

Evaluation

The PEP reflects the isovolumetric contraction time (ICT); the PEP is shortened by an increase in HR, an increase in preload (ventricular filling), a decrease in afterload, and by a positive inotropic stimulation. Accordingly, the PEP is particularly sensitive to medications that induce inotropic stimulation and vasodilatation ("inodilators"), provided there is no restriction of venous return. Inotropic stimulation increases the ventricular ejection time (VET) only slightly; accordingly, the shortening of the $QS2_c$ and the reduction of the PEP/VET-ratio ("Weissler-Index"), which are often propagated as "contractility indices," are predominantly defined by the shortening of the PEP. A reduction in afterload shortens the PEP, prolongs the HR-corrected VET with a reduction of the PEP/VET-ratio, whereas the $QS2_c$ is hardly changed. Vasodilatation-induced changes in STI are hardly changed by concomitant beta-adrenoceptor blockade and atropine; therefore, PEP and VETc can be assumed to be (also) highly afterload dependent, whereas the $QS2_c$ is not.

Normally, the electrocardiographic QT-interval is shorter than the QS2. Adrenergic stimulation and other forms of inotropic stimulation prolong the QT-interval relatively to the shortening of the QS2. Accordingly, the shortening of the QS2/QT-ratio has been propagated as one of the many "contractility indices." There have been some early applications in clinical cardiology, but no application in cardiovascular clinical pharmacology.

Critical Assessment of the Method

HR-corrections of STI are based on historic linear regressions in quite small samples. It is doubtful that these equations are stable and universal. Indeed, it is hardly likely nor can it be verified that they can be extrapolated to further subjects and different experimental conditions. Furthermore, these HR-corrected STIs are meaningless mechanically since HR is an intrinsic determinant of pump action, performance, and efficiency. A shortening of the PEP or QS2 should only then be accepted as an index of enhanced "contractility" if a simultaneous change of vascular load can be excluded.

The value of STI in cardiovascular clinical pharmacology relates particularly to their excellent reproducibility and high pharmacosensitivity: STI have been used in clinical cardiology to monitor progressing pump dysfunction including iatrogenic cardiomyopathies; in cardiovascular clinical pharmacology, STI have been used to characterize cardiotonics, negative inotropics, reduction in preload, and stress interventions.

STI have been very important in the late 1980s and throughout the 1990s for the noninvasive characterization of drug effects on systolic performance. Now, such methods appear antiquated also since there are no modern state-of-the art devices to measure and analyze STI.

Myocardial Performance Index (Tei)

Purpose and Rationale

The echocardiographic myocardial performance or "Tei" index (MPI) is the modern analogue of the STI.

Procedure

MPI is based on the estimates of the isovolumetric contraction and isovolumetric relaxation time (ICT and IRT) and ejection time (ET) obtained

by pulsed-wave Doppler (PWD) or tissue Doppler echocardiography of the mitral annulus (TDE).

Evaluation

Doppler echocardiographic ICT, IRT, ET, and MPI are important tools in clinical cardiology for the noninvasive follow-up of patients with myocardial infarction, major cardiac surgery, and after heart transplantation.

Critical Assessment of the Method

These methods have the important add-on advantage to assess both systolic and diastolic function and to be able to distinguish between left and right ventricular function.

The MPI (= (ICT + IRT)/ET) estimates were shown to have high diagnostic accuracy for heart failure, but with distinct and method-specific diagnostic cut-offs. The methods rely on a very high level of analyst expertise: they are observer dependent and not economic; the latter aspects might explain why such methods find little application in the experimental evaluation of cardiac drug effects, in spite of the wealth of information that could be gained.

Noninvasive Estimates of Stroke Volume and Cardiac Output

Purpose and Rationale

The stroke volume (SV) and cardiac output (CO = HR × SV) are the volume equivalents of the systolic cardiac pump function.

Procedure

Several noninvasive methods have been investigated and propagated for the experimental investigation of SV and CO:

- Carbon dioxide rebreathing (indirect Fick method).
- Transthoracic impedance cardiography (ZCG).
- Diastolic pulse contour analysis ("PCA"), i.e., analysis of noninvasive radial artery pulse wave forms by means of a third-order, 4-element modified Windkessel model of the circulation quantifying the Windkessel model criteria: systemic vascular resistance (SVR), large artery "capacitive" compliance (C1), small artery "oscillatory"/"reflective" compliance or "reflectance" (C2), and inductance (L – inertance of blood). This method uses an estimate of SV from the ejection time (ET), heart rate (HR), body surface area (BSA), and age, and all PCA-criteria (SVR, C1, C2, and L) rely on this estimate (and the constraints of its algorithmic simplicity).
- Systolic pulse wave analysis ("PWA"): reconstruction of the pulse wave form of the ascending aorta from distant (carotid/brachial/radial) pulse wave contours by means of a validated general transfer function (GTF) deriving the central augmentation index (AIx), the time to wave reflection (Tr as a measure of central aortic compliance), and algorithmic estimates of central hemodynamics.
- Echocardiographic techniques: M-mode echocardiography, two-dimensional echocardiography, three-dimensional echocardiography.
- Transthoracic pulsed wave Doppler echography of the aorta ascendens, transesophageal Doppler echography, etc.

Evaluation

The older devices required tedious signal analysis and complex nonautomated signal and data processing, which relied on public algorithms; newer methods are mostly highly automated "black-boxes" with proprietary algorithms that often are device-specific nonpublic "adaptations" of the original algorithms.

Critical Assessment of the Method

Invasive measurements of SV and CO are method-specific estimates relying on a "black-box" analysis of the dilution of a controlled injection of dye or a cooled volume of saline ("thermodilution").

In intensive care medicine, newer methods have been introduced that are called "minimally" invasive: they provide for continuous hemodynamic monitoring without repeated central catheter dilution; they monitor systolic function based on wave/contour analysis of (invasive) arterial peripheral pulses with or without calibration with pulmonary artery thermodilution. The surge

of "minimally" invasive methods also illustrates (1) the need for reliable methods for continuous monitoring and (2) the lack of satisfaction with and acceptance of truly noninvasive methods to meet this requirement.

The related constraints are illustrated in the following by the past and present positioning of transthoracic impedance cardiography (ZCG) in the clinical pharmacological characterization of investigational changes in cardiovascular function.

ZCG is based on the observations in the 1930s and 1940s that typical changes occur in transthoracic impedance (Z) to a high-frequency low-voltage alternating current (AC) applied through the thoracic cage during the cardiac cycle; these changes were originally primarily seen as the consequence of volume shifts with an increase in volume and decrease in impedance during systole and a decrease in volume and increase in impedance during diastole; now it is understood that the contour of the time course of the negative velocity of the transthoracic impedance changes (dZ/dt) is analogous with the blood flow velocity in the central large vessels and the differential of the carotid pulse curve, while also including venous and right ventricular components. In clinical cardiology, there was little interest in such rheological plethysmographic concepts because of the various invasive methods that became available. The need for noninvasive monitoring methods in the aerospace industry led to the first impedance cardiographic applications.

The registration of the ZCG signals is not observer-dependent, but the analysis of the signals (delineation of the ejection time and measurement of dZ/dt_{max}) is. Originally, ZCG analyses also included an assessment of STI and therefore required the simultaneous registration of at least three signals (ECG, ZCG, PCG); the ZCG signal has points of repair to delineate the start and the end of the LVET, albeit that these are more easily and accurately identified if the PCG and carotid pulse curve (4-channel method) are recorded as well. In the early 1980s, this approach, which required tedious 3- or 4-channel signal analysis, was frequently used in cardiovascular clinical pharmacology, since it permitted an almost continuous monitoring. In the mid-1980s, an alternative method became popular; it was particularly attractive since it used less inconvenient spot rather than adhesive tape electrodes, was fully automated, and relied only on the ECG and ZCG; furthermore, this method used its own physiologic algorithm and equations to estimate SV, the results of which disagree grossly with those according to the conventional equation by Kubicek applied on the same signals; furthermore, the lack of support information (PCG and/or carotid pulse curve) makes the method less accurate in estimating LVET and, accordingly, SV.

Conventional ZCG is well reliable and highly sensitive for drug effects, inodilatory effects in particular; they may agree with other invasive and noninvasive methods but often appear to overestimate SV and the changes thereof. The alternative methods have a similarly high reproducibility and are sensitive but may be less accurate in estimating LVET and, accordingly, SV. However, all three have limited validity since they yield method- and device-specific estimates of SV that are not unlikely to be affected by substantial method/subject/effect-interaction.

The fate of ZCG is exemplary for most noninvasive cardiovascular methods: they are method- and device-specific estimates that may be very reproducible and sensitive, for drug effects in particular; they have a limited validity since they do not generally agree well with the established golden standards; this per se does not preclude their usefulness, provided this limitation is understood and accounted for, also since the golden standards may prove impractical or impossible to use in similar collectives. However, in order to be useful, these methods need to be accepted as such. In drug development, this means that data generated with such methods need to be useful and acceptable for regulatory purposes. However, with the exception of ICH E14, there is no regulatory need or benefit in pursuing cardiovascular endpoints in early development studies. In the framework of "lean" drug development, this means that there is little demand for such studies. Accordingly, it has become difficult to improve their hardware and software to meet present-day quality standards and to keep the required operational expertise. Due to these latter constraints, it

has become even more difficult to satisfy regulatory requirements. In consequence, several of these methods, although evidenced to be highly informative, are no longer available. Newer methods, especially those related to pulse wave velocity and pulse wave contour analysis or Doppler echocardiography may find a similar fate unless they find high acceptance in clinical cardiology.

Diastolic Performance

Purpose and Rationale

The diastole extends from the end of the systolic ejection (closure of the aortic valve) to the start of the next isovolumetric systolic contraction phase (closure of the mitral valve); therefore, it includes the isovolumetric relaxation phase (until opening of the mitral valve); the rapid filling phase, which begins when LV pressure falls below left atrial pressure and the opening of the mitral valve and involves interaction between LV suction (=active relaxation) and viscoelastic properties of the myocardium (= compliance); diastasis, i.e., when left atrial and left ventricular pressures are almost equal and left ventricular filling is essentially maintained by the flow coming from pulmonary veins using the left atrium as a passive conduit; and atrial systole, which corresponds to left atrial contraction and ends with the closure of the mitral valve. The diastole is far more dependent on the HR than the systole and the diastolic filling lasts longer when the HR is slower.

According to the European Cardiology Society, establishment of the diagnosis of diastolic heart failure requires: (1) the presence of a clinical syndrome of heart failure (dyspnea or fatigue at rest or with exertion, fluid overload, pulmonary vascular congestion on examination, or X-ray); (2) demonstration of an ejection fraction 50%; and (3) demonstration of diastolic dysfunction (The European Study Group on Diastolic Heart Failure 1998). Others prefer the term “heart failure with a normal ejection fraction” (HFNEF), characterized by elevated ventricular filling pressures and abnormal filling patterns to allow for a better distinction between active and passive components, emphasizing that HFNEF may occur with or without impairment of the isovolumetric relaxation (active dysfunction).

Removal of calcium from the myofilaments and uncoupling of actin–myosin cross-bridge bonds govern the rate of myocardial relaxation and thus the rate of ventricular pressure decline. This active component of diastole is typically characterized by the time constant of relaxation (t), determined by fitting a mono-exponential curve to the isovolumetric section of the ventricular pressure curve. Subsequently, the mechanical properties of the ventricle are determined by passive factors, such as the degree of myocellular hypertrophy (myocardial mass), cytoskeletal and extracellular matrix properties, and chamber geometry; this is reflected by the end-diastolic pressure–volume relationship (EDPVR) and the features derived from it: ventricular chamber stiffness (i.e., slope of EDPVR at a given volume [dP/dV]) and compliance (the mathematical reciprocal of stiffness). Both are load dependent and are no measures of load-independent diastolic function (lusitropy). In consequence, diastolic dysfunction may involve either or both active or passive ventricular properties. With an increased t (which is typically observed with all forms of hypertrophy, and with aging), a higher mean left atrial pressure may be required to achieve normal filling volumes, especially at high heart rates. However, an increased t is not ubiquitously associated with elevated mean left atrial pressure and heart failure. Instead, shifts of the EDPVR have been suggested to be a predominant factor of the hemodynamic and symptomatic abnormalities of heart failure in HFNEF: a leftward/upward shifted EDPVR is indicative of decreased chamber capacitance, whereas a rightward/downward-shifted EDPVR (increased ventricular capacitance) occurs in all forms of dilated cardiomyopathy (remodeling). Accordingly, there are various conditions with distinctly different properties of the passive and/or active diastolic components that may result in HFNEF (The European Study Group on Diastolic Heart Failure 1998).

Procedure

An in-depth analysis of diastolic function requires invasive investigations to assess the pressure–volume relation along the overall cardiac cycle, which permits to derive t, end-diastolic stiffness, etc.

Noninvasively, Doppler ultrasound recordings of transmitral and pulmonary venous flow velocities and time intervals are useful alternatives, and Doppler echocardiography has become the primary tool for identifying and grading the severity of diastolic dysfunction in patients demonstrating elevated ventricular filling pressures and abnormal filling patterns.

This involves the determination of the early diastolic velocity (E), atrial velocity (A), deceleration time of E velocity (DT), and the isovolumetric relaxation time (IVRT) from the transmitral Doppler signals. Complementary evaluation of pulmonary venous flow might be of interest; further methods rely on tissue Doppler technology and color M-mode derived flow propagation rate. These investigations are carried out at rest with controlled maneuvers (Valsalva, leg lifting).

Evaluation

In contrast to inotropic changes, lusitropic changes of diastolic function are not regularly investigated and characterized except in patients with post-myocardial infarction dysfunction and other forms of heart failure. Furthermore, there are no drugs that are targeted specifically on improving diastolic function, albeit that ancillary positive lusitropic properties have been demonstrated for some medications.

Investigation of diastolic properties might be of interest in differentiating responsiveness to therapy and lack thereof in the evaluation of treatments of heart failure, but is only rarely used in this context.

In hypertension, diastolic function is also of interest since diastolic dysfunction is inherent to concentric left ventricular remodeling that is commonly seen in hypertensives.

Critical Assessment of the Method

Load-independent diastolic function (lusitropy) suffers from the same conceptual validity constraints as inotropy (load and heart rate independent systolic function). The lack of distinction between true lusitropy and passive components of diastolic performance is obvious. However, this does not preclude that the procedures to characterize ventricular relaxation and filling (even if composite and ambiguous criteria) provide a better understanding of the overall cardiac function.

Echocardiographic Evaluation of Coronary Flow Reserve

Purpose and Rationale

Coronary flow reserve (CFR) can be assessed by echocardiography through direct measurement of coronary blood flow velocity at rest and during adenosine stress test at the window of distal left anterior descending artery, using transthoracic Doppler signal and could be applied and used to calculate the ratio between peak test velocity/baseline velocity which correlates with invasively measured coronary flow reserve (CFR).

Procedure

Routinely performed in larger cardiosurgery centers. CFR is performed with several techniques (PET, MRI, and Doppler echocardiography). Peak stress blood flow measured by ergometry alone is a less powerful predictor of outcomes than CFR, possibly because CFR taken as the ratio of peak stress and rest blood flows may better isolate vasodilator capacity and reduce systematic errors in measurement.

Evaluation

In the evaluation of need for surgery, CFR is measured in culprit arteries. If CFR < 2, then bypass surgery is indicated. CFR > 2.0 is for medicamentous treatment. Noninvasive assessment of coronary vasodilator function provides incremental risk stratification beyond routine measures of clinical risk, including estimates of LV systolic function and the extent and severity of

myocardial ischemia and scar, and results in a meaningful incremental risk reclassification of patients with known or suspected CAD.

Critical Assessment of the Method

A CFR ≤ 2.7 by transthoracic echocardiography has demonstrated good accuracy (87% specific and 82% sensitive) for detecting CAV. In addition, echocardiographic CFR has been reported to have prognostic value for CAV-related major cardiac events (3.3 relative risk of death, myocardial infarction, congestive heart failure, or need for percutaneous intervention at a mean of 19 months). A CFR < 2.9 can detect a maximal intimal thickness of ≥ 0.5 mm by intravascular ultrasound (IVUS) with 80% sensitivity, 100% specificity, and 89% negative predictive value.

References and Further Reading

Abdelghani SA, Rosenthal TM, Morin DP (2016) Surface electrocardiogram predictors of sudden cardiac arrest. Ochsner J 16(3):280–289

Abduch MC, Alencar AM, Mathias W Jr et al (2014) Cardiac mechanics evaluated by speckle tracking echocardiography. Arq Bras Cardiol 102(4):403–412

Bauer A, Malik M, Schmidt G et al (2008a) Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: international society for holter and noninvasive electrophysiology consensus. J Am Coll Cardiol 52(17):1353–1365

Bauer A, Malik M, Schmidt G et al (2008b) Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: international society for holter and noninvasive electrophysiology consensus. J Am Coll Cardiol 52(17):1353–1365

Butlin M, Qasem A (2017) Large artery stiffness assessment using sphygmoCor technology. Pulse 4 (4):180–192

Chia YC, Buranakitjaroen P, Chen CH et al (2017) Current status of home blood pressure monitoring in Asia: statement from the HOPE Asia network. J Clin Hypertens. https://doi.org/10.1111/jch.13058. [Epub ahead of print]

Chiang CE, Wang TD, Lin TH et al (2017) The 2017 focused update of the guidelines of the Taiwan society of cardiology (TSOC) and the Taiwan hypertension society (THS) for the management of hypertension. Acta Cardiol Sin 33(3):213–225

Claus P, Omar AMS, Pedrizzetti G et al (2015) Tissue tracking technology for assessing cardiac mechanics: principles, normal values, and clinical applications. JACC Cardiovasc Imaging 8(12):1444–1460

Coris EE, Moran BK, De Cuba R et al (2016) Left ventricular non-compaction in athletes: to play or not to play. Sports Med 46(9):1249–1259

Fung E, Järvelin MR, Doshi RN (2015) Electrocardiographic patch devices and contemporary wireless cardiac monitoring. Front Physiol 6:149

Garcia EV, Pastore CA, Samesima N et al (2011) T-wave alternans: clinical performance, limitations and analysis methodologies. Arq Bras Cardiol 96(3):e53–e61

Gimeno-Blanes FJ, Blanco-Velasco M, Barquero-Pérez Ó et al (2016) Sudden cardiac risk stratification with electrocardiographic indices – a review on computational processing, technology transfer, and scientific evidence. Front Physiol 7:82

Gulizia MM, Casolo G, Zuin G et al (2016) ANMCO/AIIC/SIT consensus document: definition, precision and appropriateness of the electrocardiographic signal of electrocardiographic recorders, ergometry systems, Holter systems, telemetry and bedside monitors. G Ital Cardiol 17(6):393–415

Haugaa KH, Smedsrud MK, Steen T et al (2010) Mechanical dispersion assessed by myocardial strain in patients after myocardial infarction for risk prediction of ventricular arrhythmia. JACC Cardiovasc Imaging 3 (3):247–256

Haugaa KH, Grenne BL, Eek CH et al (2013) Strain echocardiography improves risk prediction of ventricular arrhythmias after myocardial infarction. JACC Cardiovasc Imaging 6(8):841–850

Huikuri HV, Stein PK (2013) Heart rate variability in risk stratification of cardiac patients. Prog Cardiovasc Dis 56(2):153–159

Huikuri HV, Perkiömäki JS, Maestri R et al (2009) Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics. Philos Trans A Math Phys Eng Sci 367(1892):1223–1238

Iwano H, Yamada S, Watanabe M et al (2011) Novel strain rate index of contractility loss caused by mechanical dyssynchrony. A predictor of response to cardiac resynchronization therapy. Circ J 75(9):2167–2175

James PA, Oparil S, Carter BL et al (2014) 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the eighth joint National Committee (JNC 8). JAMA 311(5):507–520

Klaeboe LG, Haland TF, Leren IS et al (2017) Prognostic value of left ventricular deformation parameters in patients with severe aortic stenosis: a pilot study of the usefulness of strain echocardiography. J Am Soc Echocardiogr 30(8):727–735

Kocabay G, Muraru D, Peluso D et al (2014) Normal left ventricular mechanics by two-dimensional speckle-tracking echocardiography. Reference values in healthy adults. Rev Esp Cardiol (Engl Ed) 67(8):651–658

Lee PY, Liew SM, Abdullah A et al (2015) Healthcare professionals' and policy makers' views on implementing a

clinical practice guideline of hypertension management: a qualitative study. PLoS One 10(5):e0126191. https://doi.org/10.1371/journal.pone.0126191. eCollection 2015

Leren IS, Saberniak J, Haland TF et al (2017) Combination of ECG and echocardiography for identification of arrhythmic events in early ARVC. JACC Cardiovasc Imaging 10(5):503–513

Liu J, McKenna TM, Gribok A et al (2008) A fuzzy logic algorithm to assign confidence levels to heart and respiratory rate time series. Physiol Meas 29(1):81–94

Luebbert J, Auberson D, Marchlinski F (2016) Premature ventricular complexes in apparently normal hearts. Card Electrophysiol Clin 8(3):503–514

Mc Kinstry B, Hanley J, Lewis S (2015) Telemonitoring in the management of high blood pressure. Curr Pharm Des 21(6):823–827

O'Brien E, Sheridan J, O'Malley K (1988) Dippers and non-dippers. Lancet 2(8607):397

O'Brien E, Petrie J, Littler W et al (1993) An outline of the revised British hypertension society protocol for the evaluation of blood pressure measuring devices. J Hypertens 11(6):677–679

O'Brien E, Coats A, Owens P et al (2000) Use and interpretation of ambulatory blood pressure monitoring: recommendations of the British hypertension society. BMJ 320(7242):1128–1134

Olsen FJ, Biering-Sørensen T, Krieger DW (2015) An update on insertable cardiac monitors: examining the latest clinical evidence and technology for arrhythmia management. Futur Cardiol 11(3):333–346

Pincus SM (1991) Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA 88 (6):2297–2301

Shibao C, Lipsitz LA, Biaggioni I (2013) Evaluation and treatment of orthostatic hypotension. J Am Soc Hypertens 7(4):317–324

Shrout T, Rudy DW, Piascik MT (2017) Hypertension update, JNC8 and beyond. Curr Opin Pharmacol 33:41–46

Sokolow M, Werdegar D, Kain HK et al (1966) Relationship between level of blood pressure measured casually and by portable recorders and severity of complications in essential hypertension. Circulation 34(2):279–298

Thomas GP, Daichi S, Haas D (2006) Ambulatory blood-pressure monitoring. N Engl J Med 354:2368–2374

Verdecchia P, Angeli F, Gattobigio R (2004) Clinical usefulness of ambulatory blood pressure monitoring. J Am Soc Nephrol 15(Suppl 1):S30–S33

Verrier RL, Ikeda T (2013) Ambulatory ECG-based T-wave alternans monitoring for risk assessment and guiding medical therapy: mechanisms and clinical applications. Prog Cardiovasc Dis 56(2):172–185

Voss A, Kurths J, Kleiner HJ et al (1996) The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. Cardiovasc Res 31(3):419–433

Watanabe MA (2003) Heart rate turbulence: a review. Indian Pacing Electrophysiol J 3(1):10–22

Wood PW, Boulanger P, Padwal RS (2017) Home blood pressure telemonitoring: rationale for use, required elements, and barriers to implementation in Canada. Can J Cardiol 33(5):619–625

Zhou JC, Zhang N, Zhang ZH et al (2017) Intensive blood pressure control in patients with acute type B aortic dissection (RAID): study protocol for randomized controlled trial. J Thorac Dis 9(5):1369–1374

3 Characterization of Cardiac Electrophysiology Including ECG-Analysis

Ivana I. Vranic

Contents

I. I. Vranic (✉)
Clinical Center of Serbia, Faculty of Medicine, University of Belgrade, Belgrade, Serbia
e-mail: ivvra.fesc@gmail.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_65

Abstract

Rare diseases of cardiac arrhythmias are called channelopathies which represent clinically challenging task to cope with and are considered as unrestricted area for drug innovations. The development of more effective medications that increases quality of life and reduces symptom burden, hospitalizations, and mortality ruminate prudent action. Pathophysiological basis of cellular mechanisms of antiarrhythmic drug actions is still uncovered for interpretation of pharmacovigilance results. Current therapeutic control of arrhythmias is moreover driven by clinical and demographic characteristics of patient groups then to individual, patient-specific and/or phenotype or genotype features. Contemporary limitations of pharmaceutical understanding on cellular mechanisms for clinical pharmacodynamics is the future key to unlock vast possibilities of innovation in electrocardiography industry. Accurate measurement of electrocardiograms (ECG) is critical for effective diagnosis of patient's cardiac functions. Detailed examination of filters' effects on ECG accuracy, reproducibility and robustness covering a wide range of available commercial products can provide valuable information on the relationship between quality and effectiveness of filters, and assessments of patients' cardiac functions. The fact that standard ECG device comprises several limitations when it comes to signal acquisition due to large ambient noise and inherent restrictions of technique applied, it is of utmost importance to modernize electrocardiography or at least to incorporate novel diagnostic techniques such as fragmented QRS (fQRS), heart rate variability (HRV), T peak-T end (TpTe), heart rate turbulence (HRT) and T wave alternans (TWA) that all have predictive value for pharmacovigilance.

Part 1 Pharmacovigilance on Crossroads

Rare diseases of cardiac arrhythmias are called channelopathies which represent clinically challenging task to cope with and are considered as unrestricted area for drug innovations. The development of more effective medications that increases quality of life and reduces symptom burden, hospitalizations, and mortality ruminate prudent action (Skinner et al. 2019).

Physiological grounds of newly understood pathways underlying cardiac arrhythmias have opened up novel possibilities for mechanism-based therapeutic approaches (Schwartz et al. 2017). Existing management of arrhythmias is moreover driven by clinical and demographic characteristics of patient groups then to individual, patient-specific, and/or phenotype or genotype characteristics. The significance of cellular intrinsic parameters that are applicable for clinical pharmacodynamics is the future key to unlock vast possibilities in electrocardiography diagnostics. Therefore, main focus should be on must-to innovations which will be clinically relevant and dealing with this problem. Pathophysiological basis for cellular mechanisms of antiarrhythmic drug actions should be covered as to correctly interpret pharmacovigilance results.

Delayed Afterdepolarization and "Re-entry" Culprit Initiator

Discovery of the cellular mechanisms of delayed afterdepolarization and "re-entry" culprit initiator mechanism is so-called target zone of antiarrhythmic drug action. However, pharmacological progress was largely concentrated on empirical, rather than particular mechanism-based approach (Imbrici et al. 2016). The rapid progress in pacing

device industry and ablation-driven approaches compared to the relatively poor outcome of pharmacological rhythm control treatment may have reduced the interaction between clinical electrophysiologists (whose primary interest is improving the health and quality of life of patients) and basic electrophysiologists (who are constantly driven by novel scientific discoveries) consequently resulting in a gap, which may take years to succumb and reach the clinical setting. Apart from effective and safe rhythm control strategies which are primary, we still need solid diagnostics in everyday clinical practice.

Fundamental Heterogenetic Mechanisms of Cardiac Arrhythmias

Pronounced heterogeneities exist in mechanisms, presentation, treatment, and outcome, between different cardiac arrhythmias and also between patients with the same type of arrhythmia. Nonetheless, arrhythmias require a vulnerable substrate to induce acute initiating triggering event. Both components can be genetic or even acquired due to advancing age or concomitant cardiovascular risk factors that promote abnormal impulse formation (ectopic activity) and/or abnormal impulse conduction (resulting in re-entrant activity) that are considered major fundamental mechanisms of arrhythmia.

DAD Ectopic Activity and "Re-entry"

Ectopic activity is the local generation of action potentials (APs) outside the normal activation sequence, which can serve as an initiator of re-entry–mediated arrhythmias or maintain the fibrillation process when occurring repetitively at high frequency. Polymorphic ventricular tachycardia (VT) or ventricular fibrillation often originate from the Purkinje system, which is characterized by distinctive electrophysiological characteristics that promote the generation of the ectopic excitations and which can be targeted by catheter ablation (Fonseca and Vaz da Silva 2018). In addition, ectopic excitations may trigger premature ventricular contractions which worsen myocardial function and contribute to increased mortality in patients with structural heart disease. Mechanistically, ectopic activity results from abnormal automaticity or triggered activity resulting from early afterdepolarizations (EADs) or delayed afterdepolarizations (DADs). Abnormal automaticity, i.e., spontaneous AP generation from normally quiescent tissue, is often mediated by a reduction in ion currents that stabilize the resting membrane potential (RMP; e.g., the basal inward rectifier potassium current, IK1) or an increased activity of hyperpolarization-activated pacemaker channels responsible for the funny current (Wang et al. 2017). On the other hand, EADs typically result from excessive AP duration (APD), providing time for reactivation of L-type calcium channels. Finally, DADs result from calcium-handling abnormalities, whereby spontaneous calcium-release events from the sarcoplasmic reticulum (SR) through so-called ryanodine receptor type-2 (RyR2) channels activate a transient-inward current, mediated largely by the sodium/calcium-exchanger (NCX), which depolarizes RMP and can trigger a new AP. An increased incidence of spontaneous calcium-release events can be due to RyR2 dysfunction or increased SR calcium load. A classic example of the latter is arrhythmic activity induced by cardiac glycoside intoxication, where inhibition of the sodium–potassium-ATPase increases the intracellular sodium concentration (Garcia-Elias and Benito 2018).

This leads to increased cytosolic calcium by reducing the driving force of calcium efflux via NCX at resting membrane potential and supporting reverse-mode NCX-mediated calcium entry during the AP. The resulting increased SR calcium load not only explains the positive inotropic effect of cardiac glycosides but also favors spontaneous calcium release and DADs. Re-entry is considered the primary arrhythmia maintaining mechanism. Conceptually, re-entry describes a stable circulating excitation. Re-entry can occur around an anatomically defined in excitable core when the product of the effective refractory period (ERP) and conduction velocity (the so-called wavelength) is sufficiently small, allowing the tissue to become re-excitable before the re-entering impulse arrives. As such, structural remodeling, particularly fibrosis, strongly

promotes re-entry (Crotti and Kotta 2017). Re-entry may also occur on purely a functional basis. The classical "leading-circle" concept describes re-entry around a continuously refractory core with a circuit size equal to the wavelength. However, the notion that sodium channel-blocking agents are effective in cardioversion of atrial fibrillation (AF) challenged this concept, since conduction slowing is believed to stabilize re-entry according to the "leading-circle" concept rather than suppressing it. Based on this paradox, the "spiral wave" concept has been developed, wherein re-entry proceeds around an excitable but continuously unexcited central tissue and maintenance of re-entry depends on the balance between the wave front excitatory currents – the "source" and the properties of the unexcited tissue – the "sink." Spiral wave re-entry, often visualized as rotors, represents a major pathophysiological concept in AF and rotor ablation has been suggested as a potential ablation strategy of persistent AF.

Progress to Conduction Disturbances

Cardiac fibrosis, as a hallmark of structural remodeling, is promoting slow, heterogeneous conduction and spatial dispersion of repolarization. This process separates double helicoidal muscular band into parts of inactivity. In parallel, basic science has identified those mechanisms underlying the proliferation of fibroblasts and their differentiation into myofibroblasts, which play a central role to numerous stress signals activated in many cardiovascular diseases, resulting in excessive collagen production and fibrosis (Arbustini et al. 2018).

Re-entry mechanisms that promote slow conduction are also mediated by gap-junction remodeling, resulting in impaired electrical cell-to-cell coupling. The preclinical studies have shown that expression of connexin-43, a predominant constituent of ventricular gap-junction channels, has been reduced under pathological conditions (e.g., post-myocardial infarction), contributing to slowed conduction, and that gene transfer of connexin-43 reduce the occurrence of ventricular tachycardia in pigs, providing an important proof-of-concept for the antiarrhythmic effects of targeting the mechanisms underlying slow conduction (Lau et al. 2015).

Experimental models that successfully identified the major underlying signaling pathways were indicating of targeting these pathways, either using clinically approved drugs (e.g., angiotensin-converting enzyme inhibitors, aldosterone antagonists) or through new targets (e.g., microRNAs such as miR-21 or miR-29) which may limit the development of a proarrhythmic substrate. However, there are increasing evidence that applying such therapy would need early start before extensive fibrosis occurs. Some observations stated that fibrosis and structural alteration can be partly reversible, if an underlying condition is effectively treated.

T-Wave Alternans

Cardiac alternans represents beat-to-beat alternations in the ST segment or the T-wave. At the cellular level, APD alterations underlie T-wave alternans (TWA) on recorded ECG signal. When APD alternans become spatially discordant, a small reduction of cycle length may result in unidirectional block followed by re-entry and initiation of ventricular fibrillation, thus providing a mechanical coupling between TWA and sudden cardiac death (SCD). Basic science has revealed that cardiac alternans is often accused of perturbed cellular calcium signaling. In support of this impression, normalization of SR calcium handling in animal models of heart failure (HF) has been shown to lessen the incidence of cardiac alternans and ventricular arrhythmias. Thus, normalization of RyR2 function and SR calcium behavior may be promising targets to prevent development of alternans and SCD.

Inherited Channelopathies

Inherited channelopathies are key examples of cardiac diseases in which basic science has provided important systematic understanding of

pathophysiology (Batul et al. 2017). This has led to the development of specific diagnostic tests such as the ajmaline experiment in Brugada syndrome (Campuzano et al. 2015) and QTc evaluation during stress test in long and short QT syndromes (LQTS, SQTS), and importantly the continued study and development of mechanism-based therapeutic methodologies such as late sodium channel block in LQT3 and the use of flecainide in catecholaminergic polymorphic ventricular tachycardia (CPVT), which have already partially entered the clinical guidelines.

Long and Short QT Syndrome

Both pathological elongation (LQTS) and hastening (SQTS) of cardiac repolarization can predispose patients to atrial and ventricular tachycardia (AT/VT) and SCD (Lieve and Wilde 2015). Such mutations in KCNQ1 and KCNH2/HERG potassium channel and SCN5A sodium channel genes (recognized in the 1990s) were accused of 15 genes linked to LQTS. Nevertheless, mutations in six genes were found to be liaison to SQTS. An improved understanding of the biophysical significances of these mutations (complete loss of function vs. change in voltage-dependence of the respective currents) and the exact location of the mutations within the channels (pore-region vs. C- and N-terminus) may have clinical importance for risk forecast. Categorization of the biomechanical channel characteristics of the mutated channels and their response to sympathetic activity have helped identify the importance of EADs and sympathetic stimulation as proarrhythmic triggers, providing the scientific foundation for beta-blocker therapy (or left cardiac sympathetic denervation) as a pillar of antiarrhythmic therapy in LQTS. The subsequent generation of genetic (knock-out/knock-in) and transgenic (mouse and rabbit) models of LQTS has provided significant understandings into the major role of spatial and temporal APD heterogeneities as an arrhythmogenic substrate. Mechanisms underlying genotype-specific differences in arrhythmia beginning (Prakash and Sharma 2016) [sustained sympathetic activity (sports), LQT1; sudden sympathetic activation (startle), LQT2] contributing to clinically observed genotype differences in the efficacy of beta-blocker therapy have been recognized. The detrimental role of bradycardia and short-long-short sequences as triggers for VTs has been detected in patients and established in (animal) models. In addition, the potential of genotype-specific therapeutic approaches has been discovered: in genetic murine LQT3 models, with pathologically enhanced late INa, sodium channel blockers mexiletine, flecainide, and the more selective GS967 exert a mechanism-directed, genotype-specific antiarrhythmic effect (Cheung et al. 2016). Similarly, recent clinical data reveal that LQT3 patients benefit from QT-shortening and antiarrhythmic effects of sodium channel blockers mexiletine, flecainide, ranolazine, and eleclazine. The 2015 ESC guidelines for the prevention of SCD thus mention sodium-channel blockers as a potential add-on therapy in patients with LQT3. Moreover, proarrhythmic effects of estradiol due to increased APD heterogeneities and EADs have been shown in transgenic LQT2 rabbit models, contrasting with anti-arrhythmic, protective effects of progesterone due to shortening of cardiac refractoriness and reduced EAD formation. These data suggest progesterone-based therapies might constitute novel antiarrhythmic tactics in female LQTS patients. In SQTS, mechanisms of arrhythmogenesis are less well implicit than in LQTS and therefore effective antiarrhythmic treatment options are scant. The recent development of genetic/transgenic animal models of SQTS (zebrafish, rabbit) has the perspective to close this gap and help to develop mechanism-based antiarrhythmic therapeutic approaches.

Brugada Syndrome

Brugada syndrome (BrS) is regarded as by typical precordial ST-segment elevations on the surface ECG and an increased risk for SCD caused by ventricular arrhythmia (Honarbakhsh et al. 2018). The identification of SCN5A as an important disease gene in BrS suggested that the syndrome was a monogenetic disorder, similar to LQTS. "Loss-of-function" mutations in other genes, implicated in cardiac sodium (GPD1L, SCN1B, SCN3B) or

calcium (CACNA1C, CACNB2) channel expression, were also associated with BrS, as well. "Gain-of-function" mutations in KCND3, by contrast, demonstrated an involvement of increased transient-outward potassium current Ito, in the pathogenesis of BrS. Interestingly, recent data suggested SCN10A and its gene product Nav1.8, primarily expressed in the nervous system and only to a minor extent in the heart, as an important disease gene in BrS, although one study did not confirm this view. However, only ~30% of BrS cases may be explained by the 22 disease-associated genes known so far, pointing to a multifactorial source of the malady. This view was fueled by recent findings that hearts of BrS patients often show subtle structural changes, in particular in the right ventricular outflow tract (RVOT), suggesting a distinct cardiomyopathy. This is pathophysiologically important as such regions – either caused by channelopathy or even more complex mechanisms – may give rise to areas of low voltage and electrical zig-zag conduction that constitute a substrate for micro re-entry as the arrhythmogenic basis of the syndrome. Observations in BrS patients undergoing electrophysiological study and reports of successful radiofrequency ablation of life-threatening parts in the RVOT further substantiated this pathognomonic finding, constructing the support of the depolarization hypothesis. Accordingly, localized slow conduction in the RVOT creates the origin of typical ECG fluctuations in right precordial leads. The repolarization hypothesis, by contrast, postulates an imbalance of ion currents during phase 1 of the right ventricular epicardial AP, mediated by transmural differences in the outward potassium current Ito (Chen et al. 2014) and genetic or drug-induced reduction of INa, underlying the ECG pattern and the induction of arrhythmias in BrS. Regardless which of the two hypotheses more precisely reflects the underlying mechanisms, it is an essential finding of basic electrophysiological exploration that reflects a misalliance of depolarizing current (in particular INa) and Ito, as well as distinct phases of bradycardia, are major drivers of arrhythmogenicity in BrS. Therefore, modulation of the INa/Ito relation has an important role in current clinical diagnostic risk stratification and therapeutic approaches in BrS, including: (1) the application of sodium-channel blockers as a diagnostic test for inducibility of Brugada ECG (ajmaline or flecainide test), (2) the use of quinidine to decrease arrhythmic burden in BrS patients by inhibiting Ito, (3) the use of isoproterenol or other β-adrenergic stimuli in arrhythmogenic storm to increase depolarizing current (in particular L-type calcium current; ICa, L) and avoid bradycardia, (4) avoidance of drugs that decrease Ina or increase Ito (see http://www.brugadadrugs.org), and (5) avoidance of fever to prevent temperature-induced INa/Ito mismatch. However, currently, the only way to safely prevent SCD is the implantation of an implantable cardioverter defibrillator (ICD), representing vital necessity for basic and translational research to provide better mechanism-based diagnostic and therapeutic strategies (Vranic 2017).

Catecholaminergic Polymorphic Ventricular Tachycardia (CPVT)

CPVT is an extremely risky inherited arrhythmogenic disorder branded by episodes of polymorphic VT and SCD in the setting of exercise or emotional stress. In recent years, significant progress has been made to understand the underlying mechanisms, which has positively influenced the therapy of CPVT patients. The molecular basis of CPVT has mostly been ascribed to dysfunction of the RyR2 channel, encouraging calcium-handling abnormalities, DADs, and triggered activity, particularly during increased sympathetic tone when SR calcium load increases and RyR2 channels become hyperphosphorylated, further promoting their dysfunction. In cca. 65% of CPVT patients, this RyR2 dysfunction is due to mutations in the RYR2 gene itself (CPVT1), but mutations in proteins regulating RyR2 function as part of the macromolecular complex that controls SR calcium release (e.g., calsequestrin-2 or calmodulin), which can indirectly promote RyR2 dysfunction, have also been defined.

Current data from animal models of CPVT propose that the primary weaknesses in calcium handling may also cause sinoatrial node (SAN) dysfunction, fibrosis, atrial arrhythmias,

exaggerated left ventricular (LV) hypertrophy, and HF. These results might explain why some affected individuals exhibit bradycardia and atrial arrhythmias, in addition to CPVT. Furthermore, they provide a direct mechanic link between a primary calcium handling disease and structural remodeling in the heart (Spears and Gollob 2015). This could have more general repercussions for other pathologic states.

The identification of the primary arrhythmogenic mechanisms has played a vital role in the treatment of CPVT patients, which comprises inhibition of sympathetic activity (beta-blockers or left cardiac sympathetic denervation) and the use of flecainide. The role of flecainide in CPVT represents a major example of the interaction between basic scientists and clinicians. It was initially established in mice and applied as a proof of concept in two CPVT patients, but has since then become a class IIa guideline recommendation. However, the exact antiarrhythmic mechanism of flecainide remains a topic of debate and may involve a combination of sodium channel block (preventing triggered activity), inhibition of RyR2 (reducing the likelihood of spontaneous SR calcium-release events), or indirect effects on intracellular calcium handling due to changes in intracellular sodium (Posselt et al. 2015). Based on the underlying mechanisms of CPVT, RyR2 dysfunction is considered a promising therapeutic target. Several RyR2-stabilizing compounds, including K201/JTV-519, S107 and carvedilol analogues, have indeed shown promise in preclinical studies. Of particular clinical interest are the effects of dantrolene, an RyR-stabilizing drug approved for the treatment of malignant hyperthermia, which also has antiarrhythmic effects in CPVT patients.

Electrical Conduction Defects

Cardiac conduction defect (CCD) also comprises primary genetic forms, although most cases arise from age-related fibrotic degeneration of the conduction system. The yield of genetic testing in CCD is around 30%. Isolated CCD can be caused by mutations in SCN5A (progressive familial heart block 1) or TRPM4 (progressive familial heart block 2). In association with cardiomyopathy, CCD was linked to LMNA and PRKAG2 mutations. Patients with such mutations are also at risk of ventricular arrhythmias rendering further decision-making regarding ICD implementation important (McNamara et al. 2015). Furthermore, mutations in NKX2–5 and TBX5 cause CCD in association with congenital heart disease. Thus, evidence for genetic forms of CCD can significantly improve stratification of patients, as it constitutes a prognostic indicator for the course of disease with or without a syndrome comorbidity (Ackerman 2015).

Hereditary Sinus Node Disease

Loss or dysfunction of SAN cells results in sinus node disease (SND), embracing sinus bradycardia, SAN block or arrest, and bradycardia-tachycardia syndrome (BTS). In the mainstream of cases, SND is idiopathic and occurs age dependent, either by cellular dysfunction of SAN cells or degeneration of the formerly intact SAN. Loss-of-function mutations in SCN5A are an established pathomechanism (sick sinus syndrome 1). Electrophysiological studies and computer-visualized techniques successfully revealed malformed channels that cause either abnormal slow pacing or even produce SAN send-off block. Besides, mutations in HCN4 underlie a significant proportion of the pacemaker current If in the SAN, which cause hereditary sick sinus syndrome and bradycardia-tachycardia syndrome evenly. It is observed that HCN4 “gain-of-function” mutations are associated with inappropriate sinus tachycardia. Functional investigations showed that HCN4-R524Q mutant channels, heterozygous in nature, carried and affected by family members, have an increased cAMP sensitivity resulting in an augmented funny-current at baseline leading to a hastier resting heart rate. Thus, improper sinus tachycardia has potential genetic background, which increases the possibility of hereditary features. Importantly, contemporary research of If and its underlying ion channels, have built the basis for ivabradine (the first

clinically available If blocker) targeting impulse formation in the SAN. This pharmacological mechanism is successfully depleted to heal patients with chronic stable angina and HF.

Drug-Induced, Reversible "Channelopathies"

Diverse drugs can phenocopy ECG features and arrhythmias resembling genetic channelopathies by interacting with different cardiac ion channels. A variety of them cause, for example, acquired LQTS and BrS-particularly in patients harboring single-nucleotide polymorphism variants or even silent disease-causing mutations. A range of drugs causing acquired LQTS block HERG-encoded IKr. Yet, drugs may also cause acquired LQTS by blocking other currents such as IKs (e.g., isoflurane) or IK1 (e.g., midazolam). Drug-induced BrS is often caused by sodium-channel-blocking class 1 drugs. On the whole, acquired drug-induced "channelopathies" and arrhythmias are much more widespread than rare genetic forms and exploring their mechanisms has important preclinical and clinical repercussions to prevent adverse drug effects.

Pharmacology Safety Methods

Each newly developed pharmacological agent endures broad cardiac safety testing to exclude any proarrhythmic accountability. Until recently, cardiac safety assays predominantly involved high amount of screening HERG blockade early on and a so-called "thorough QT study" as to assess possible repolarization elongation during drug development. However, extensive basic science studies have established that HERG screening alone has limited sensitivity and specificity to identify proarrhythmic composites, which has urged a more integrative tactic to assess the cardiac safety of new compounds. The Comprehensive In vitro Proarrhythmia Assay (CIPA) initiative advocates the use of (1) screening of ion-channel-blocking effects beyond HERG alone, (2) in silico integration of these findings to assess overall effects on ventricular repolarization, and (3) use of integrated biological systems such as induced pluripotent stem cell-derived cardiomyocytes (IPSC-CM).

Ischemic Cardiomyopathy (ICM)

Ischemia-induced release of reactive oxygen species, so-called "free radicals," produce an increase of intracellular sodium. This is due to sodium/hydrogen exchange and inhibition of the sodium/potassium-ATPase initiated activity. This process is driven by hypoxia or secondary depletion of ATP. Within hours after ischemia had arisen, associated cytokine and chemo-attractant mediated endothelial dysfunction, as well as induction of apoptosis, autophagy, platelet aggregation, and micro-embolization have happened. Neutrophil accumulation as well as macrophage and T-cell-mediated cell damage was a matter of devastation process. While acute ischemia/reperfusion injury may primarily lead to ventricular fibrillation, myocardial infarction-related scars (specially the border zone between infarcted and vital myocardium) are an important substrate for VT occurrence in ICM. Infarcted tissue and concomitant changes in excitability of cardiomyocites and conduction instabilities may lead to perpetual reentrant VT.

Pharmacological treatment of ventricular fibrillation is presently incomplete. Class I antiarrhythmic drugs are contraindicated due to augmented risk for arrhythmia-associated sudden cardiac death as shown in the "Cardiac Arrhythmia Suppression Trial" (CAST). After the CAST study had eliminated the use of class I antiarrhythmic drugs in ICM, several antiarrhythmic compounds have undergone evaluation, as the alternative compound amiodarone, besides its antiarrhythmic potential, had been shown to have cumulative toxicity in many body organs. The Class III antiarrhythmic compound, MS551 (nifekalant), was described to have antiarrhythmic properties shortly after the CAST results had been published. This pyrimidine derivative leads to frequency-dependent AP-prolongation. It has voltage- and frequency-dependent inhibitory

properties on HERG-encoded IKr with high affinity for the "open state" of IKr. As described above, this is known to possibly induce an acquired form of QT-interval prolongation with the risk for fatal polymorphic tachycardias including Torsade de Pointes. However, unlike other class III agents, it also has agonistic/facilitating effects on HERG current. Nifekalant has a high potency for destabilization and early termination of spiral wave reentry and to prevent VT/ventricular fibrillation after acute myocardial infarction. It also improves electrical defibrillation efficacy in this setting. However, nifekalant is only approved in Japan for the treatment of life-threatening ventricular tachyarrhythmias. Another class III antiarrhythmic compound, SSR149744C (celivarone), is, like amiodarone, a benzofuran derivative; however, unlike amiodarone, it is not iodinated and thus exerts less organ toxicity. Celivarone has antiadrenergic and angiotensin II antagonistic effects and is a multi-ion channel blocker. Similar to amiodarone, it is a weak sodium channel blocker with additional IKr, IKs, and IK1 blocking properties, but exerts less effective IK, ACh-, IKv1.5-or ICa, L blockade. In the ICARIOS-trial, celivarone showed a 46% (non-statistically significant) reduction in VT/−ventricular fibrillation-triggered ICD shock therapies; however, in the ALPHEE Study, celivarone was not effective for the prevention of ICD interventions or death.

To date, according to the European Society of Cardiology guidelines, only beta-blockers (in the nonacute phase of MI) and amiodarone are recommended in ischemia-related ventricular fibrillation or VT associated with heart failure. Therefore, it seems urgently necessary to search for other treatment options of ischemic cardiomyopathy-associated arrhythmias. A very recent and interesting approach comes from experimental data from a pig animal model which evaluated cardiac remodeling in an ischemia/reperfusion setting with or without administration of the Class III antiarrhythmic compound dronedarone. In a meticulous transcriptome profiling and combined proteome analysis of post-infarction remodeling, the authors found the levels of 879 transcripts in the infarction border zone, seven transcripts in the myocardial infarction area, as well as 51 proteins in the unaffected left ventricle and 15 proteins in the border zone affected by dronedarone treatment. All findings were supported by disease/function charts and an integrated network established by combined "omics." Table 1 shows transcriptomics with the most predominant changes in gene expression. Although dronedarone is not approved for ventricular arrhythmias, this approach highlights the importance of myocardial infarction border zone in ICM. Secondly, it may lead to similar evaluation approaches for future antiarrhythmic compounds that can improve our understanding of arrhythmogenesis on a subcellular transcriptome and proteome level.

Table 1 ECG parameters and their behaviors prior to ventricular arrhythmias

	Behavior		
ECG parameter	Presence	Increase/ prolongation	Decrease
QT_c/QT_vi/ QT_d fQRS		✓	
ER	✓		
VLP	✓		
HRV	✓		✓
iCEB		✓	✓
QT dynamicity		✓	
HRT		✓ (TO)	✓ (TS)
TWA		✓	
TpTe		✓	

Autonomic Nervous System

Autonomic imbalance characterized by increased sympathetic activation and parasympathetic withdrawal, along with changes in density and spatial distribution of the intrinsic efferent innervation of the ventricles, may account for the timing of clinical presentation of arrhythmias after ventricular injury. Sympathetic hyperinnervation in the ventricle occurs post-MI and has been linked to ventricular arrhythmias. Whereas nerve growth factor (NGF) stimulates axon growth, its precursor, pro-NGF, triggers axon degeneration and may be

involved in regional denervation after myocardial injury. Additionally, denervated reperfused infarcts display β-adrenoreceptor supersensitivity. Both sympathetic hyperinnervation and denervation of the ventricles can lead to heterogeneous β-adrenoreceptor activation, either through localized catecholamine release or localized β-adrenoreceptor supersensitivity. This nonuniform sympathetic activation increases the risk of focal triggers and creates gradients of repolarization, increasing the susceptibility for re-entry. Attempts to therapeutically reduce sympathetic activation or sympathetic nerve sprouting (e.g., by cardiac sympathetic denervation or renal denervation) or to increase cardiac parasympathetic tone (e.g., by baroreceptor stimulation) reduce arrhythmias in animal models and selected cohorts of patients.

Nonischemic Cardiomyopathy

Nonischemic cardiomyopathies include genuine dilated or congestive cardiomyopathy (DCM) and special cases such as arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D). While in ICM usually focal areas with post-ischemic/MI fibrosis and scarring can be found that lead to reduced myocardial contractility, as contrary to ARVC/D where it is found as early pathognomonic sign (Vranic 2012), in DCM the impaired myocardial contractility is generally dispersed. Pathophysiology of VT generation in DCM is as complex as its variety of genetic causes. Nowadays, an abundance of genes and alleles that contribute to phenotypic DCM have been identified. Gene alterations of TTN (coding for Titin, 12–25% of DCM), LMNA (coding for Lamina/C, 4–8% of DCM), DES, VLC, and FLNC (coding for cytoskeletal proteins, each around 1% of DCM) significantly contribute to the DCM phenotype. While DSP, coding for desmoplakin, also contributes to DCM, other desmosome changes instead result in ARVC/D. Loss of RNA-binding protein 20 (RMB20), which is a RNA-binding protein of a spliceosome of TTN and calcium/calmodulin-dependent protein kinase II (CaMKII) delta, leads to a clinically aggressive form of DCM. Despite similar left ventricular function, altered calcium handling increased arrhythmic burden (44% vs. 5%) when RBM20 mutation carriers were compared to TTN mutation carriers. Therefore, ICa, L-blockers may possibly reduce arrhythmia burden in this disease entity. Besides playing a role in LQTS3 and BrS, SCN5A mutations can also lead to DCM or ARVC/D. In some cases with a gain-of-function mutation in the sodium channel (i.e., in p.R222Q mutation carriers), class I antiarrhythmic agents have been reported to decrease arrhythmogenic burden and improve left ventricular function. In ARVC/D, several desmosomal and non-desmosomal gene mutations have been identified to induce phenotypic disease. Desmosomal genes include plakophilin 2, desmoglein 2, and desmoplakin gene mutations contributing to half of all cases. Nondesmosomal changes include amongst others gene mutations in the RyR2, phospholamban, Lamin A/C, Desmin, Titin, and transforming growth factor 2. Pharmacological treatment options include beta-blockers in all ARVC/D phenotype patients, as well as sotalol, amiodarone, and mexiletine. In selected patients, epicardial ablation may be considered. Besides antiarrhythmic drugs, experimental data of mouse models have shown reduced structural and electrical remodeling leading to arrhythmia reduction by blockade of the renin-angiotensin-system. The direct renin inhibitor aliskiren was shown to have antiarrhythmic potential by re-establishing normal ventricular conduction velocities due to restoration of connexin 43 expression in a DCM mouse model. Also, the angiotensin II receptor antagonist candesartan was able to partially reverse pro-arrhythmic down-regulation of Kv4.2 (Ito channel protein), KChIP2 (auxiliary subunit of Kv4.2), and Kv1.5 (IKur channel protein) in another DCM mouse model. To date, however, most treatment approaches in DCM are symptomatic or prophylactic and not driven by pathophysiologic understanding of the disease or underlying genetic pathology. Therefore, the main goal for the future is to gain further knowledge about underlying disease pathology at the organ, cellular, and subcellular levels in order to translate this knowledge into mechanism-based therapeutic approaches.

Atrial Arrhythmias

Atrial arrhythmias, particularly AF, are the most common cardiac rhythm disorders and are associated with increased rate of stroke, HF, and death. Great efforts have been undertaken to understand the underlying arrhythmic mechanisms and improve treatment options. Although there are now many sophisticated therapeutic options available, antiarrhythmic therapy remains unsatisfactory. In contrast, pacemakers are the accepted standard therapy for patients with bradycardia. Similarly, patients with accessory-pathway or AV-nodal re-entry syndromes, such as Wolff–Parkinson–White syndrome, are usually successfully treated with ablation therapy.

Atrial Tachycardia (AT)

ATs are subdivided based on underlying mechanisms into focal tachycardias and macroreentrant tachycardias, also designated as atrial flutter. Focal ATs commonly arise in the absence of preexisting structural heart disease and can occur at any age with no gender preference. The autonomic nervous system likely contributes to the initiation of focal AT, which can be triggered by changes in posture, belching, and swallowing. Pathomechanisms of focal AT include abnormal automaticity (19%), triggered activity (25%), and microre-entry (56%). Focal ATs due to triggered activity are likely to be mediated by DADs. Focal ATs may also be caused by microre-entry, which refers to a small reentrant loop which is below the resolution of current mapping systems. Adenosine can be used to distinguish re-entry from other mechanisms underlying AT (Vranic and Matic 2006). Adenosine binding to A1-receptors leads to dissociation of Gi-proteins and to activation of a repolarizing potassium current (IK,ACh), due to direct interaction with the Gi-protein βγ-subunit. The resulting stabilization of the RMP terminates focal AT caused by afterdepolarizations, transiently suppresses AT due to abnormal automaticity, but does not affect AT based on re-entry mechanisms. The current knowledge surrounding underlying AT mechanisms is not yet sufficient to guide antiarrhythmic drug choice, but may help to understand why certain treatment options may be effective in some patients while ineffective in others. There is hope that the future development of selective drugs may provide new therapeutic options for mechanism based therapy of AT.

Remodeling Evolution of Atrial Fibrillation

The "atrial fibrillation creates atrial fibrillation" idea states that the longer AF persists, the greater inducibility and stability of AF, i.e., the better the likelihood that AF will perpetuate. These basic research findings summaries the clinical observation that AF is a highly progressive disease. The developing nature of AF is fundamentally due to atrial remodeling process on the cellular and tissue level so basic research has been instrumental in unravelling these processes. Atrial remodeling comprises processes such as the electrical, contractile, structural, and calcium-handling remodeling. The electrical remodeling is characterized by an intense reduction in atrial remodeling caused by shortening of atrial myocyte APD. The previous is due to altered regulation and expression of ion channels, e.g., diminished ICa, L- and short-lived outward potassium current (Ito), and augmented inward rectifier potassium currents and TASK-1 (K2P 3.1) current. The contractile remodeling is caused by numerous mechanisms including damaged calcium handling and dysregulation of the sarcomeres and myofilaments. Noticeable characteristics of structural remodeling are atrial dilatation, atrial myocyte hypertrophy, and fibrosis, which create a substrate for AF. Calcium-handling remodeling includes altered expression and/or regulation (e.g., phosphorylation) of L-type calcium channels, NCX, RyR2, SERCA2a, or CaMKII. Calcium-handling abnormalities (a) contribute to the electrical, contractile, structural, and intracellular signaling remodeling, (b) and may provide both a trigger (e.g., by spontaneous SR calcium release triggering DADs) and a substrate for AF (by contributing to electrical and structural remodeling), and (c) may be both a reason and result of AF. There are prominent differences in the time

sequence and reversibility of electrical and structural remodeling with important clinical implications. Electrical remodeling is an initial event starting within hours or days after beginning of AF and is reversible upon restoration of sinus rhythm. Fibrosis (structural remodeling), on the other hand, is a late event and is considered poorly reversible. In patients, the degree of atrial fibrosis correlates with the occurrence of postoperative AF as well as with the perseverance and reappearance of AF. Thus, a high degree of atrial fibrosis may be viewed as a marker of critically remodeled atria highly susceptible to AF. These findings about the time course and progression of atrial remodeling may explain the clinical remark that cardioversion and catheter ablation for rhythm control are more successful in younger patients with shorter history of AF and less atrial structural remodeling. They stress the clinical view that, in order to be treated successfully, AF has to be diagnosed early, i.e., before the development of excessive structural remodeling, which may represent a point of no return to sinus rhythm. Modern evidence from cellular electrophysiology designates that calcium-handling remodeling displays important variances between paroxysmal and chronic AF. These findings have important clinical implications: first of all, they indicate that paroxysmal and chronic AF may be diverse entities and that (calcium-handling) remodeling does not automatically represent a gamut from paroxysmal to chronic AF, where a certain adjustments simply develop and worsen over time. Secondly, it has been proposed that optimal handling for paroxysmal and chronic AF could be different and ideally be tailored to the individual remodeling process in a particular patient.

The multifaceted gamut of structural, architectural, contractile, or electrophysiological changes affecting the atria with the potential to produce clinically relevant manifestations has been summarized as atrial cardiomyopathy.

Classification of AF Subtypes

Patient selection for antiarrhythmic therapy relies primarily on symptoms and the duration of AF, i.e., paroxysmal, persistent, long-standing persistent, or permanent. This stratification is incomplete due to lack of data for arrhythmia burden (inherited risk) or the severity of the remodeled atrial substrate. Explicitly, the impact of basic scientific findings on clinical relevance of AF subtypes and therapeutic decision making was poor, despite significant advances in the field. From a cellular-based perspective, the most striking modifications between AF subtypes of patients are verified changes in APD of atrial myocytes. In patients with persistent or permanent AF but without pronounced LV dysfunction, the "classical" mechanism of atrial arrhythmogenesis is detected, categorized by re-entry promoting shortening of atrial APD. In this particular group of AF, inhibition of repolarizing K+ channels that are expressed in human atrium by class III antiarrhythmic drugs is expected to be particularly effective in defeating AF. The coexistence of AF with severe LV dysfunction portrays a clinically significant therapeutic dare attributed to a distinct atrial substrate basis. In affected patients with paroxysmal AF, atrial AP is prolonged. The hypothesis that handlings to activate potassium currents could utilize antiarrhythmic effects in these cases remains to be tested in translational approaches. Noteworthy, HF-associated APD prolongation and APD shortening in chronic forms of AF offset each other, resulting in actually normal APD in patients with long standing persistent AF and severe LV impairment. The upgrading and claim of procedures evaluating the range of atrial remodeling in patients such as electro-anatomic mapping (van der Bijl et al. 2017) or late gadolinium enhancement magnetic resonance imaging (LGE-MRI), as well as mapping techniques applied during ongoing arrhythmias (e.g., mapping of focal or re-entrant activity) may help to mend arrhythmia subtypes and the commencement of customized cure of AF patients. In translational research strategies assessing antiarrhythmic ideas that are useful in patients with paroxysmal AF or following rhythm control by ablation or cardioversion, expenditure of cells obtained from subjects with persistent or even permanent AF exhibit a different cellular electrophysiological mechanism to establish a

limitation which requires a careful consideration. Biopsy-derived cells obtained from specific patient populations targeted by antiarrhythmic interventions (e.g., paroxysmal AF) should be preferred when studying antiarrhythmic concepts.

Vagal AF

AF patients with anatomically normal hearts tend to show a vagal pattern of AF onset (nocturnal AF, AF during rest or after exercise), while patients with primary heart disease tend to illustrate a sympathetic pattern (AF during daytime or during exercise). Other studies in sole AF patients allude that AF onset might be allied with a combined symathovagal activation rather than with an increase in vagal or sympathetic drive alone. Nocturnal arrhythmias might be triggered by autonomic activation due to sleep muddled breathing like obstructive sleep apnea. In the course of obstructive respiratory episodes of intense periapneic vagal activation, followed by combined sympathetic activation at the end of the apnea, following arousal may trigger AF. Furthermore, endurance athletes are more prone to develop AF than nonathletes. Variety of sport characteristics like sort, force, and extent of training influence the risk of developing AF at different mechanisms involving increased vagal tone at rest, prompt cardiomyocyte sensitivity to cholinergic stimulation, structural atrial remodeling, sinus bradycardia, and genetic heritage, which all or in part may contribute to AF in athletes (D'Silva and Sharma 2014). Management of sleep apnea or detraining in athletes may lessen AF burden in these patients.

Novel Antiarrhythmic Drug Examples Contra Classical Ones

Contrary to inherited channelopathies, the ICD remain the cornerstone of antiarrhythmic therapy in patients at high risk for ventricular arrhythmias due to acquired conditions. Without a doubt, with the exception of beta-blockers, antiarrhythmic drugs have not shown a consistent efficacy in the primary management of arrhythmias in this setting, perhaps because currently available antiarrhythmic drugs have predominantly been identified based on chance observations during clinical studies, without a precise understanding of the molecular mechanisms underlying initiation and maintenance of arrhythmias. The discovery of sodium-channel block as the major mechanism underlying antiarrhythmic effects of quinidine led to the development of class I antiarrhythmic agents such as flecainide and propafenone. Better understanding of the central biophysical causes of state-dependent block of sodium channels subsequently directed to the subdivision of agents into classes IA, IB, and IC by Vaughan Williams. Ranolazine is a multichannel blocker, primarily inhibiting IKr and the late component of the cardiac sodium current, which was originally developed as an antianginal drug. Numerous studies have supported the therapeutic potential of ranolazine for ventricular arrhythmias. Conversely, the large, randomized, double-blind, placebo-controlled Ranolazine Implantable Cardioverter-Defibrillator (RAID) trial in which high-risk ICD patients with ischemic or nonischemic cardiomyopathy were randomly assigned to ranolazine or placebo only found a nonsignificant 16% reduction in the primary composite outcome of VT/ventricular fibrillation or death. Still, although more selective blockers of the late sodium current, such as eleclazine, exhibited robust antiarrhythmic effects in animal studies, all subsequent clinical trials have recently been dropped.

Besides beta-blockers, amiodarone is most often used due to contraindications for other antiarrhythmic drugs in the presence of advanced structural heart disease, which is frequent in this population. The antiarrhythmic effects of amiodarone were also an unplanned discovery during amiodarone use in the therapy of angina. Further investigations revealed that potassium-channel inhibition by amiodarone reduces the probability of re-entry by delaying APD and ERP. In addition, amiodarone blocks sodium and calcium channels and constrains the effects of alpha- and beta-adrenoceptors, thus owning properties of all four Vaughan Williams classes of antiarrhythmic agents. So, joint blockade of multiple ion channels has recently been suggested to lie beneath the relatively high anti-

AF efficacy of amiodarone. Still, amiodarone's extracardiac toxicity has motivated the search for alternatives. Dronedarone is the first amiodarone-analog and exhibits reduced toxicity and lipophilicity. Initial clinical studies with dronedarone did not show significant extracardiovascular toxicity and the ATHENA trial demonstrated a reduction in stroke associated with dronedarone use in AF patients.

Nonetheless, dronedarone is less effectual in upholding sinus rhythm than amiodarone and it is contraindicated in patients with HF. A number of alternative antiarrhythmic strategies for patients with ventricular arrhythmias have been anticipated based on preclinical studies. In the middle of these, inhibition of CaMKII, which appears to play a nodal role in both atrial and ventricular arrhythmias, has received significant interest. CaMKII activity is increased in multiple cardiovascular diseases including AF and HF and targets a large number of ion channels and calcium-handling proteins, including L-type calcium channels, RyR2 and phospholamban. New studies have emphasized the antiarrhythmic potential of several new CaMKII inhibitors, particularly for conditions in which CaMKII activity is increased. Alternatively, since RyR2 dysfunction plays a major role in maladaptive cardiac remodeling, e.g., in the setting of HF, RyR2 stabilizing drugs (e.g., K201/JTV-519, S107, carvedilol analogues) may have a more general therapeutic use, outside their role in RyR2-associated channelopathies like CPVT.

Atrial Antiarrhythmic Drugs

Even with the rapid development of radiofrequency ablation strategies for the treatment of AF, their efficiency and safety remain insufficient. Besides, these methods can only be applied in a constrained number of patients and therefore pharmacological approaches remain clinically pertinent, especially due to the large and mounting size of the AF population. The current search for new agents against AF has led to the development of atrial-selective antiarrhythmic drug lines. Vernakalant and ranolazine, for example, are the primary approaches to selectively aim atrial sodium channels by taking advantage of the biophysical modifications between atrial and ventricular channels. In atrial tissue, a higher fraction of sodium channels is in the inactivated state because of the more depolarized resting membrane potential. Since vernakalant and ranolazine bind with higher affinity to activated/inactivated sodium channels than to channels in the closed state, this may contribute to an atrial specific effect on peak sodium current of these drugs. In addition, ranolazine predominantly inhibits late sodium current, which is increased in patients with AF and has been suggested to contribute to AF pathophysiology. Another approach to developing atrial specific compounds aims to target potassium channels which are predominantly expressed in the atrium, such as IKur, IKACh, two-pore K+ -channels, or Ca2+ −dependent K+ -channels. Interestingly, some of the already available antiarrhythmic agents, such as flecainide, amiodarone, quinidine, chloroquine, or verapamil, inhibit some of these atrial-selective K+ -channels, which may contribute to their antiarrhythmic effect in AF. The unique effectiveness and low arrhythmogenic potential of amiodarone has been attributed, among other factors, to its broad spectrum of ion-channel blocking effects. Therefore, identification of specific combinations of ion-channel modulating activities could optimize antiarrhythmic efficacy and atrial selectivity. As a result, the combination of ranolazine or amiodarone with either dronedarone or dofetilide has higher antiarrhythmic efficacy, compared with either drug alone and AF-selectivity of sodium channel blockers can be improved by adding potassium channel blockade. In the HARMONY trial, the combination of ranolazine and dronedarone showed synergistic effects in reducing AF burden in paroxysmal AF patients. However, at present, currently available evidence is insufficient to recommend antiarrhythmic drug combinations, and further research is necessary to define the required channel blocking profile and to validate these approaches in clinical studies. Basic research during recent years on the mechanisms underlying AF pathology also led to the identification of new potential antiarrhythmic approaches, including the normalization of atrial

calcium-handling abnormalities, atrial metabolism, or autonomic-tone manipulation. These are currently not implemented in therapeutic strategies and wait as future perspectives in AF treatment.

Tailored Diagnostic and Therapeutic Methodologies

An enlarged understanding of pathomechanisms underlying specific disease subtypes may open the avenue for more specific, personalized tailor-made therapies. In monogenic channelopathies such as LQTS, for example, this has already led to genotype-specific approaches, e.g., late sodium channel blocker therapy in LQT3. On the other hand, even mutation specific, personalized therapies may develop as (a) different mutations may convey pronounced differences in arrhythmic risk and (b) pronounced mutation-specific differences are witnessed in response to a specified antiarrhythmic drug. The SQTS "gain-of-function" mutations HERG-N588K and KCNQ1-V307 L, for example, alter the extent of IKr/IKs-blocking effects of various beta-blocking agents and HERG-N588K may similarly diminish IKr-blocking (and hence therapeutic) effects of a variety of class I and III agents such as sotalol, quinidine, and amiodarone. Here, the development of patient-specific hiPSC-CM may help to test custom-made anti-arrhythmic tactics.

Apart from monogenic arrhythmia disorders, in more common arrhythmias such as AF, a more detailed classification of the disease into mechanic subtypes may promote patient-oriented rather than widespread therapeutic strategies. miRNAs may help us to characterize the substrate or the electrical phenotype in separate AF patients and forecast the outcome of interventional therapy. Diminished expression of anti-fibrotic miR29, for example, is concomitant with increased atrial fibrosis and vulnerability to AF in a canine ventricular tachy-pacing-induced HF model. Likewise, miR29 expression was abridged in patients with cAF and plasma levels of miR29b could be used to foresee the outcome in pAF-patients. Altogether, sub-stratification of AF phenotypes may allow the grouping of different mechanism-specific drugs, guide antiarrhythmic atrial selective channel modifying drugs, as well as transformers of remodeling processes, inflammation, etc., and interventional therapies. Application of human-induced pluripotent stem cells like cardiomyocytes in experimental and clinical electrophysiology, with use of hIPSC-CM, may theoretically provide the unique opportunity to investigate cardiac electrophysiology.

Preferably, hIPSC-CM might envisage arrhythmogenic risk and efficacy of drugs, so as to contribute to personalized medicine. Question would be, how closely hIPSC-CM look like native human cardiomyocytes, in precise: (a) Which type of cardiomyocytes do hIPSC-CM represent? Ventricular, atrial, nodal, or just a mixture of them? (b) What is the repolarization reserve of hIPSC-CM compared to classic approaches based on animal and undiseased human ventricular tissue? (c) How big is the relatively depolarized RMP as an intrinsic peculiarity of hIPSC-CM? (d) Are there any other reasons for automaticity in hIPSC-CM than low RMP? (e) How can APD be controlled when quantified at altered spontaneous beating rates? Can hIPSC-CM be recycled as biological pacemakers? (f) Do hIPSC-CM summarize basic results on cAMP/PKA-mediated influence to heart pathophysiology? (g) Can hIPSC-CM be used as an unconventional study model for human atrial electrophysiology?

Stem Cells and Gene Therapy

Stem cell-based approaches for the treatment of heart rhythm disorders have almost exclusively focused on the generation of biological pacemaker activity. Recent ability to study human-induced pluripotent stem cells (iPSCs)-derived cardiomyocytes has given promise to a big step forward in this way. It is well known that previously tested stem cell experiments were done in animal studies, regretfully insufficient of testing in human body. Since patient cells have the potential to become testing platforms for disease phenotyping or therapeutic discovery, we are witnessing this baroque époque in cardiology. These methods

technologically advanced to genetically reprogram adult cells into iPSCs, differentiate iPSCs into ex vivo models of adult cardiac tissue, and iPSCs-based progress in discovering the mechanisms underlying pro-arrhythmic disease phenotypes. Despite being discovered less than 15 years ago, several studies have successfully leveraged iPSCs-derived cardiomyocytes to study malignant arrhythmogenic diseases. These models have the capacity to rise our understanding of the pathophysiology principal of these complex illnesses and may recognize custom-made tactics to cure.

Gene therapy, on the other hand, offers greater selectivity than small-molecule-based or interventional treatment. The gene of interest is packaged into viral or nonviral carriers and delivered to the target area via direct injection or using catheter-based techniques, providing the advantage of site-restricted action in contrast to systemic application of drugs. Gene therapy for heart rhythm disorders is currently being evaluated in preclinical stages. To date, no antiarrhythmic gene therapy drug is commercially available or has been investigated in clinical trials. Antiarrhythmic effects against AF and ventricular tachycardia, and restoration or suppression of pacemaker activity were successfully achieved in promising preclinical gene therapeutic approaches. However, in terms of clinical translation, most stem cell and gene therapy approaches are in the early stages of a complex developmental process. That involves extensive research and caution prior to widespread human application, but may herald a shift in contemporary cure.

Part 2 ECG-Analysis and Device Recording Techniques and Limitations

Interestingly, contemporary ECG machine is mainly a digital system (Kligfield et al. 2007). After initial meting out by the front-end module, the analog ECG signal is immediately converted into a digital form, resulting of the appearance of electrocardiogram tracing that is hampered by heart signal recording instrumentation. These apprehensions are primarily established by the bandwidth of filtering frequencies of 150–250 Hz in mainstream of available devices (Ricciardi et al. 2016).

Typical carry out for conducting cardiovascular diagnostics and monitoring procedures are completed by broadly available electrocardiographic devices, even stress ECG systems, trailed by others like bedside – ECG monitors, usually present in the intensive care units (ICU) or the ambulatory recorders for outdoor activities. Recording of the electrical heart activity is becoming more challenging through mobile wearable devices, so we may conclude how important this is in routine medical praxis. The precise analysis of the ECG signals is crucial due to their extensive applicability and also the high-performance expectations of medical professionals. Accurate interpretations of ECGs have always relied heavily on state of the art signal processing.

The preprocessing segments for an electrocardiographic device enclose multiple levels of signal manipulation and detection procedures, which begin by altering analog signals into digital data that is cast-off for analysis, reading, and presentation (display, print out, and storage). It is aimed here to debate a key ECG processing topic: ECG filtering and data accuracy in a modern electrocardiographic device. Four typical filter processes are present in an ECG device: (1) anti-aliasing and upper-frequency cut-off, (2) baseline wander suppression and lower-frequency cut-off, (3) line-frequency rejection, and (4) muscle artifact reduction.

Use of additional filter algorithms features (such as cubic spline technique, time-varying muscle artifact filter, source consistency filter, etc.) to operate the waveform data is out of scope of our intention. Here we first introduce the basis of filtering in the time and frequency domain, because these concepts are fundamental to the discussion of data accuracy (Singh and Pradhan 2018).

Justifications for ECG Filtering

Falsifications of Amplitude and Phase

A filter is generally designed to diminish or eliminate some frequencies from the input data

(Wang et al. 2019). One could expect that a filter gets rid of just the noise without changing the desired signal. Then again in the real world, noise and desired signals often overlap in the frequency domain. As a consequence, when a filter lessens the frequency components of noise, the overlapping desired signal can also be compressed, causing magnitude distortion of the desired signal.

In totaling to the magnitude response, the phase response is another feature of a given filter. Phase distortions occur when a filter's phase response is not a linear function of frequency, so that the phase "shift" is not directly proportional to the frequency. Phase distortion introduced by a filter could produce a significant impact on the data accuracy (examples of phase distortion can be found later in Figs. 3 and 4).

The Nyquist frequency is a key concept in the initial data sampling (with A/D conversion) and resampling processing (such as further down-sampling). To prevent aliasing interference, the signal must be band-limited. The anti-aliasing process is to use a low-pass filter (LPF) to discard the annoying frequencies (identical to and larger than the particular Nyquist frequency) of the input signal before sampling or resampling.

Oversampling and Down-Sampling

Oversampling technique is used in numerous available ECG devices. Oversampling simply refers to an initial A/D conversion sampling rate f_{os}, which is many times higher (e.g., 8000 Hz) than the final data resolution target sampling rate f_s (e.g., 500 Hz) that is used for further processing of the ECG signal. Note that usually $f_{os} \gg f_s$. The oversampled data with higher time resolution could be used for applications like pacemaker detection.

From the oversampling rate f_{os} to the target sampling rate f_s, there is a decimation process, including filtering by an LPF and then resampling (down-sampling) the resulting smoothed signal at a lower sampling rate.

Oversampling has more than a few other advantages such as facilitating pacemaker spike presentation on ECG printouts (with reduced amplitude from an LPF, see Appendix I), quantization error improvement if required concerning the precision of the least significant bit, and lower order analog anti-aliasing filter implementation.

Cutoff Frequency and 3-dB Attenuation

Sometimes it is only cutoff frequency f_c used to define a filter. For example, a 150-Hz LPF means that this filter passes low-frequency signals and has a cutoff frequency at 150 Hz. The term cutoff originates from an ideal filter that can be realized theoretically through a quadrangular function in the frequency domain, where there is not a modification and the cutoff frequency is the function's edge or boundary between a passband (signal is unchanged) and a stopband (signal is completely eliminated).

Not like an ideal filter, a real filter has a transition band from the passband to the stopband. Exhausting the same manifestation, cutoff frequency f_c is defined as the frequency at which a 3-dB attenuation or 30% reduction occurs (but not completely cutoff) from the passband for magnitude response control. Therefore, a 150-Hz filter or a filter with 150-Hz cutoff frequency has a 3-dB "cutoff" or actual attenuation at 150 Hz. In broad terms, the band from the cutoff frequency where the extent of response turns the corner to a frequency (at or before the Nyquist frequency) where the response reaches its bottom maximum attenuation can be viewed as the transition. With viable roll-off specifications applied in a real filter system, the Nyquist frequency could be significantly greater than the upper-frequency cutoff. Since 150 Hz is the recommended cutoff, 250 Hz should be a common choice for the Nyquist frequency and that sets the digital data stream sampling rate at 500 Hz.

In the frequency domain, f_c is significant, but often more provisions should be reflected so that a panoramic picture of a filter process is taken into account by reviewing items such as the following:

1. Flatness of the passband (such as over a range of 1–40 Hz for ECGs)
2. The shape of the transition band
3. The stopband
4. Phase response

Utmost Reliability and the Filter Design Aim

Filters are quantified in the frequency domain, but the time domain retort is what is seen in the signals. Although the result of filtering is arbitrated by the reliability with which it represents the original ECG signal, the concrete intention objective, which is slightly different from a up-front criterion in pursuit of signal reliability, is over and over again to try to find a good trade-off between less magnitude distortion and noise reduction, or a better signal to noise ratio, while still minimizing or avoiding phase distortion.

Low-Pass Filter in Aliasing and Upper-Frequency Cutoff

Frequently two processes for anti-aliasing exists when it comes for filtering: an analog LPF located before A/D conversion and a digital decimation LPF after digitization. The analog LPF is for the initial oversampling process, whereas the digital decimation LPF is for the down-sampling process. If either of them has frail reduction at their individual break bands, aliasing could perform then. One practical scenario is that an ECG unit and alternative medical device are fastened to a patient at the same time. The second device outputs an active signal with frequencies slightly higher than the target Nyquist frequency of the first ECG unit and pairs evident noise to the ECG input. Aliasing artifacts can be observed from the ECG printouts when a weak digital anti-aliasing filter is used.

The analog LPF has three other functions regarding frequency response: provide a flat passband, setup oversampling cutoff f_{oc} greater than 150 Hz (normally it is set much higher than 150 Hz due to oversampling) and minimize nonlinear-phase response. Similarly, a decimation filter provides a flat target passband and controls upper-frequency cutoff $f_c = 150$ Hz with 3-dB (30%) attenuation and, if possible, offers linear-phase response in the passband and even the transition band.

To discard the aliasing noise, an analog LPF should reject frequencies at and greater than $f_s/2$ if the target sampling frequency f_s is used in the A/D conversion. More often, an oversampling technique does apply to an ECG device. For this situation, an analog LPF should exclude frequencies adjacent to the oversampling frequency f_{os} (not at the oversampling Nyquist frequency $f_{os}/2$; see Appendix II) while providing a flat passband larger than the target f_c. Also, it should have a near-linear phase for the passband.

If there is only one stage from f_{os} to f_s, the decimation filter should reject any frequencies at and greater than $f_s/2$, rather than achieving −100% at the fixed 500 Hz without a specific f_s. If $f_s = 500$ Hz, for example, the decimation filter should be designed with linear phase, small ripple in the passband, upper-frequency cutoff f_c at 150 Hz, and strong attenuation at $f_s/2$ (a reaction like filter 1, not 2).

Eventually, the upper-frequency cutoff in IEC 60601–2-51 should be changed from 100 to 150 Hz, which is important especially for pediatric ECGs. In addition, if the upper-frequency cutoff is notably greater than 150 Hz when the sampling rate is 500 Hz, it can cause larger overshoot and rhythm of the step response in the time domain due to the sharp attenuation in the frequency domain when rejecting frequencies toward a 250-Hz target. Slim QRS complexes may also contain overshoot and ringing from this effect.

Primary development of clinical application of ECG focused on measuring waveform changes with highlighting on timing relationship of P, QRS, and T waves. The instrumentation bandwidth was set to 100 Hz and voltage scales were dependent on dynamic range of strip diagram recorders. The 25 mm/s scales and 1 mV/cm are still in use today although leftovers of prior era.

The historic perspective began when signal processing methods used to approximate and analyze high-resolution electrocardiogram. Actually, this required using signal averaged modus operandi. Attainment objective was to improve the signal to noise ratio in an ECG print out. The resolution of signals with the order of 0.5 μV was required to detect late potentials efficiently. The requirement for late potential analysis necessitates computerized data acquisition, QRS detection and alignment, averaging, and noise measurement.

Two fundamental analytical attitudes have been responsible for the analysis of signal

averaged ECG, one of them being time domain and the other frequency domain analysis.

The time domain analysis is dependent of filtering: mainly therefore influencing accurate vector magnitude transformation and measurement techniques (Vranic and Vranic 2019). It is of unique significance if one wants to observe all slurs and notches present in one QRS part of signal.

The frequency domain is concerned with spectral analysis techniques. Still, these events require low noise ECG and still are not possible with the standard ECG measurements. Initial software-based analysis was designed to reproduce the methods used by paramedics. As these programs developed, resolution of the standard ECG remained the same concerning time and voltage, regardless of number of bit sampling (resolution in this context refers to the conversion of an analog voltage to a digital value in a computer and vice versa).

How Much Difference Carries 12-bit, 16-bit, or 24-bit Recorded Signal?

Analog input DAQ devices from various manufacturers is called 12-bit, 16-bit, or 24-bit, which generally means that analog to digital converter returns so many bits. Assuming ADC chip returns 16 bits, it is probably superior than a 12-bit converter, but not at all times. Naive fact that a converter returns 16-bits says tiny about the quality of those bits.

Just to simply declare the resolution of a given machine is not enough. Preferably it is expressive to provide actual measured data that can support the resolution of the device which includes typical characteristic noise.

Looking at a device termed 24-bit one can find that it typically provides 20 bits effective or 18 bits noise-free (like the UE9-Pro). The U6-Pro and T7-Pro provide some of the best performance around from a 24-bit ADC, and they do about 22 bits effective or 20 bits noise-free.

Oversampling Problem

Considering the frequency-domain transfer function of a traditional multi-bit ADC with a sine-wave input signal according to Nyquist theory, F_s is obliged to be at least twice the bandwidth of the input signal.

By detecting the result of an FFT analysis on the digital output, we see a single tone and tons of random noise extending from DC to $F_s/2$ (Fig. 1). Known as quantization noise, this effect results from the following consideration: the ADC input is a continuous signal with an infinite number of possible states, but the digital output is a disconnected function whose number of different states is determined by the converter's resolution. So, the transformation from analog to digital loses some information and introduces some distortion into the signal. The magnitude of this error is random, with values up to ±LSB.

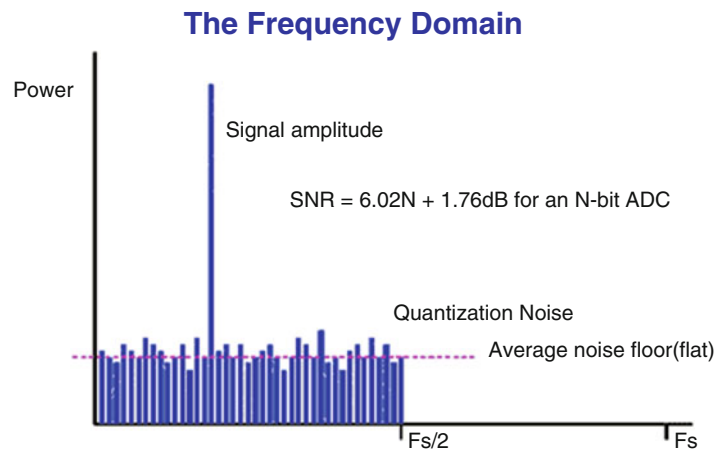


Fig. 1 FFT diagram of a multi-bit ADC with a sampling frequency F_S

By splitting the fundamental amplitude by the RMS sum of all the frequencies representing noise, we get the signal-to-noise ratio (SNR). For an N-bit ADC, SNR = 6.02 N + 1.76 dB. To improve the SNR in a conventional ADC (and consequently the accuracy of signal reproduction), one must rise the number of bits.

Considering the above example with a sampling frequency enlarged by the oversampling ratio k, to kF_s (Fig. 2) delivers FFT analysis that the noise floor has fallen. SNR is the equivalent as previous, but the noise energy has been feast over a broader frequency range. Sigma-delta converters exploit this effect by following the 1-bit ADC with a digital filter (Fig. 3). The RMS noise is less, because most of the noise passes through the digital filter. This action enables sigma-delta converters to achieve wide dynamic range from a low-resolution ADC.

Does the SNR improvement form oversampling and filtering? Noteworthy, the SNR for a 1-bit ADC is 7.78 dB (6.02 + 1.76). Each factor-of-4 oversampling increases the SNR by 6 dB, and each 6 dB increase is equivalent to gaining one bit. A 1-bit ADC with 24x oversampling achieves a resolution of four bits, and to achieve 16-bit resolution you must oversample be a factor of 4, which is not realizable. But, sigma-delta converters overcome this limitation with the technique of noise shaping, which enables a gain of more than 6 dB for each factor of 4x oversampling.

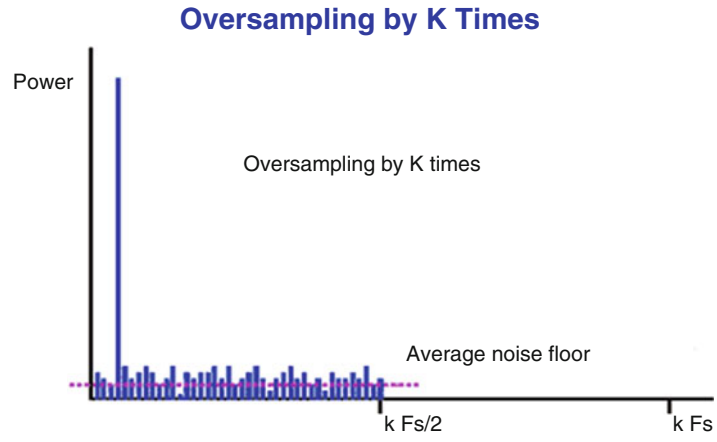


Fig. 2 FFT diagram of a multi-bit ADC with a sampling frequency kF_S

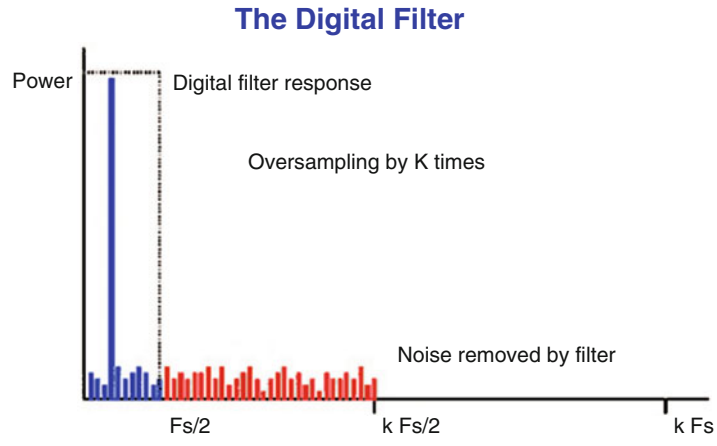


Fig. 3 Effect of the digital filter on the noise bandwidth

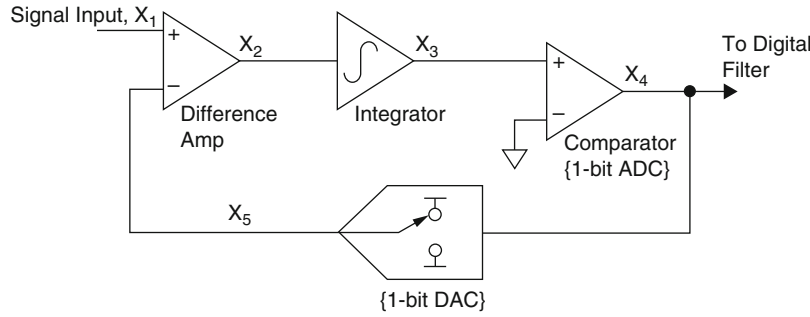


Fig. 4 Block diagram of a sigma-delta modulator

Noise Shaping

To comprehend noise influencing, consider the block diagram of a sigma-delta modulator of the first order (Fig. 4). It includes a variance amplifier, an integrator, and a comparator with feedback loop that contains a 1-bit DAC. (This DAC is simply a switch that connects the negative input of the difference amplifier to a positive or a negative reference voltage.) The purpose of the feedback DAC is to maintain the average output of the integrator near the comparator's reference level.

Compactness of "ones" at the modulator output is proportionate to the input signal. For an increasing input, the comparator generates a greater number of "ones," and vice versa for a decreasing input. By adding the error voltage, the integrator acts as a low-pass filter to the input signal and a high-pass filter to the quantization noise. Thus, most of the quantization noise is pushed into higher frequencies (Fig. 5). Oversampling has changed not the total noise power, but its spreading.

When applying digital filter to the noise-shaped delta-sigma modulator, it removes more noise than ensures simple oversampling (Fig. 6). This type of modulator (first-order) provides a 9 dB improvement in SNR for every doubling of the sampling rate. For higher orders of quantization, we can achieve noise shaping by including more than one stage of integration and summing in the sigma-delta modulator. For example, the second-order sigma-delta modulator of (Fig. 7) provides a 15 dB improvement in SNR for every doubling of the sampling rate. (Fig. 8) shows the relationship between the order of the sigma-delta modulator and the amount of over-sampling necessary to achieve a particular SNR. These interferences distort the original electrical heart signal, which may cause false diagnostics of specific heart condition (Vranic 2017).

Digital and Decimation Filter

The output of the sigma-delta modulator is a 1-bit data stream at the sampling rate, which can be in the megahertz range. The purpose of the digital-and-decimation filter (Fig. 9) is to extract information from this data stream and reduce the data rate to a more useful value. In a sigma-delta ADC, the digital filter averages the 1-bit data stream, improves the ADC resolution, and removes quantization noise that is outside the band of interest. It regulates the signal bandwidth, settling time, and stop-band denial.

In sigma-delta converters, a widely used filter topology that performs the low-pass function is the $Sinc^3$ type (Fig. 10). The main advantage of this filter is its notch response, which (for example) can reject the line frequency when set at that frequency. The notch position is directly related to the output data rate (1/data-word period). The $SINC^3$ filter decides in three data-word periods. With a 60 Hz notch (60 Hz data rate), the settling time is 3/60 Hz = 50 ms. For uses that require lower resolution and a faster settling time, consider an ADC of the MAX1400 family, which provides a choice of filter type ($SINC^1$ or $SINC^3$).

The settling time of a SINC filter is one data word. As in the example above, 1/60 Hz = 16.7 ms. Because bandwidth is reduced by the digital output filter, the output data rate can satisfy the Nyquist criterion even though it is lower than the original sampling rate. This can be realized by preserving specified input samples and abandoning the rest.

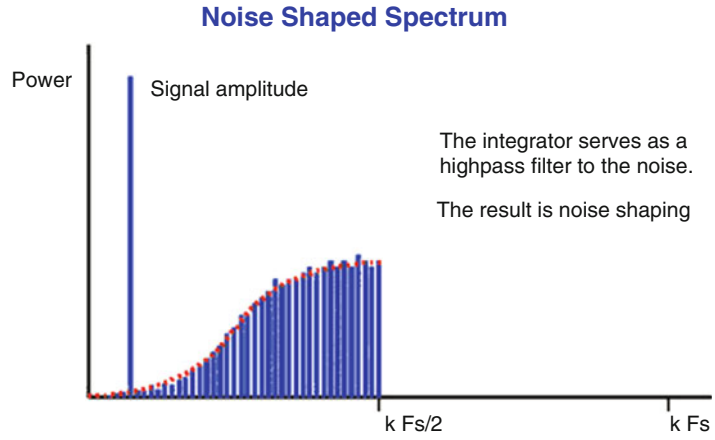


Fig. 5 Affect of the integrator in the sigma-delta modulator

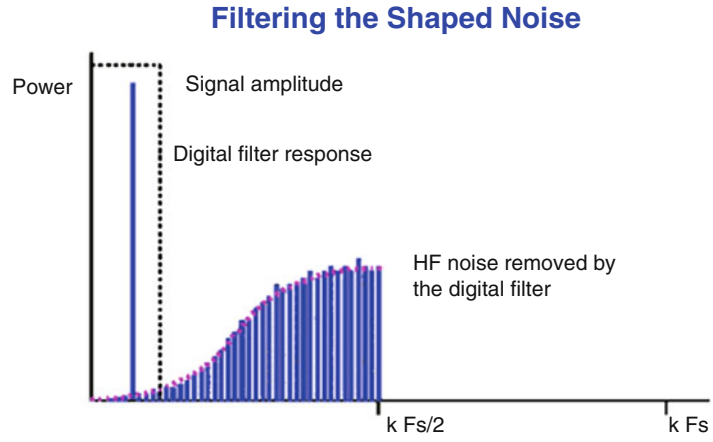


Fig. 6 Effect of the digital filter on the shaped noise

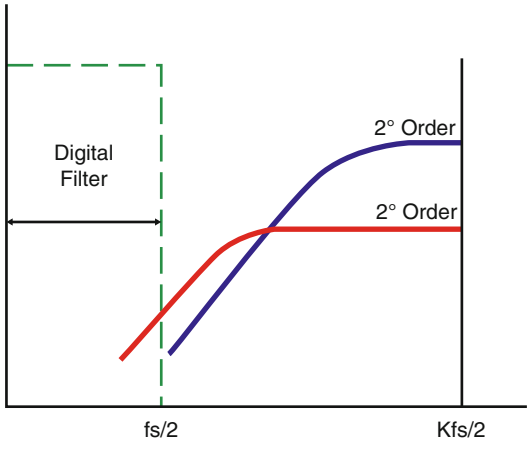


Fig. 7 Using more than one integrator and summing stage to achieve a higher order of quantization noise

This process is known as decimation by a factor of M (the decimation ratio). M can have any integer value, provided that the output data rate is more than twice the signal bandwidth (Fig. 11). If the input has been sampled at f_s, the filtered-output data rate can therefore be cut to f_s/M without loss of information.

What is the exact increase in resolution with signal averaged ECG (SAECG)? There is great variation in display formats, so it cannot be responded by unpretentious scale vectors. Digitalized ECG uses sampling rate of 250 Hz represented by small cube time on ECG paper as 0,04 s (4 ms) increment between sample points.

Quite the reverse, SAECG generally use a 1000–2000 Hz sampling rate that provides 0,01–0,005 increment between sample points

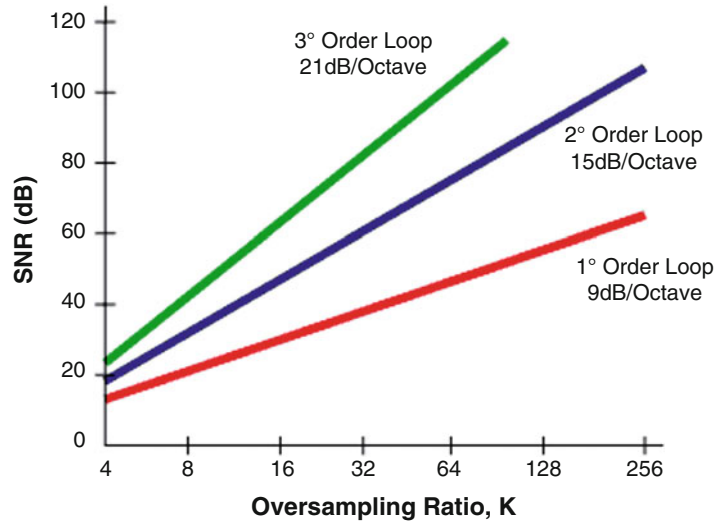


Fig. 8 Relationship between order of sigma-delta modulator and the amount of over-sampling necessary to achieve a particular SNR

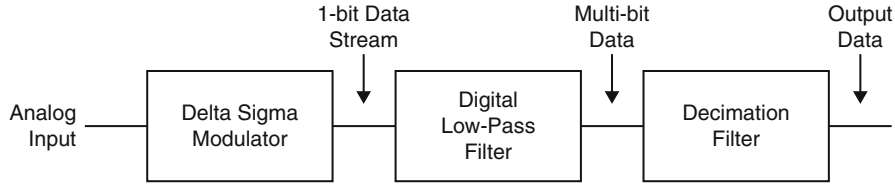
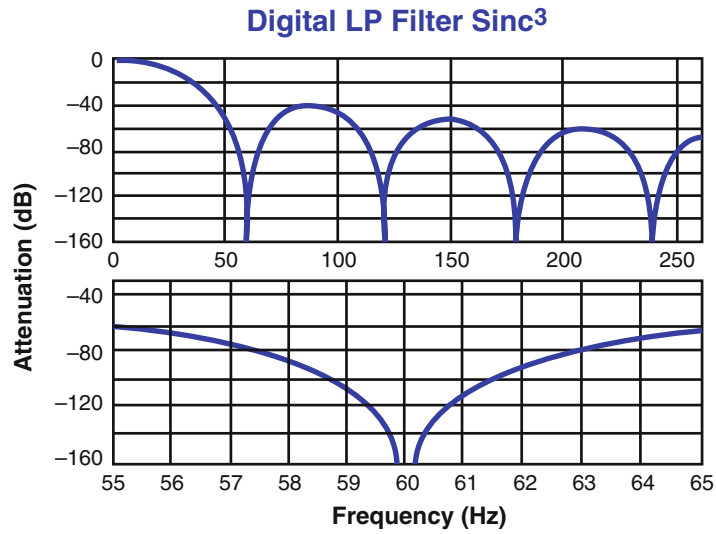


Fig. 9 Digital side of sigma-delta modulator

Fig. 10 Low-pass function performed by Sinc filter



which is 4–8 times that of the standard ECG. Comparing standard ECG by means of vertical scales with 0.5 mm or 50 μV, the scale of SAECG permits the measurements of 0.5 μV which is 100 times improvement over the standard ECG.

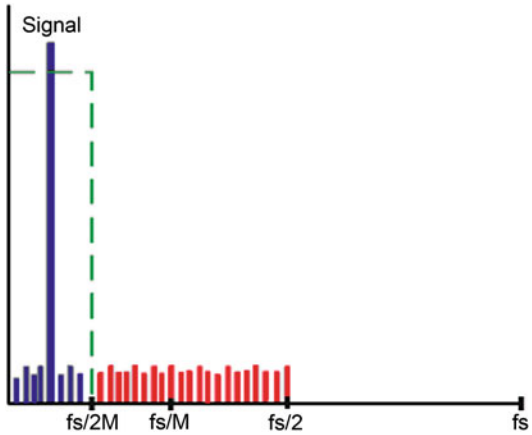


Fig. 11 Decimation does not cause any loss of information

Part 3 Ten Novel ECG Parameters

Investigators have revealed that ECG data that include noninvasive parameters could mirror underlying VT/VF risk (Vranic 2013). The ECG parameters, such as fragmented QRS (fQRS), heart rate variability (HRV), T peak-T end (TpTe), heart rate turbulence (HRT), and T wave alternans (TWA), have predictive value for arrhythmic events. In this case, cardiac monitoring by analyzing the ECG parameters is an effective way to determine future occurrence of the fatal VT/VF (Schneider et al. 2017).

The ten ECG parameters that have potential values to demonstrate underlying ventricular arrhythmias (VA) are heart rate corrected QT interval (QT_c), QT interval variability ($QT_v i$), QT interval dispersion (QT_d), fQRS, early repolarization (ER) (Mazzanti et al. 2017), ventricular late potentials (VLP), HRV, HRT, index of cardioelectrophysiological balance (iCEB), QT dynamicity, TWA, as well as TpTe. The parameters have various behavior changes prior to VA onset (Mandala and Di 2017).

Table 1 shows the changes of the parameters prior to the VA onset that can be grouped into three behaviors, i.e., presence, increase, and decrease. The presence of three ECG parameters, namely VLP, fQRS, and ER, could be associated with an increased incidence of ventricular arrhythmic events. These three parameters appear to precede the arrhythmic event. As listed in Table 1, it is also found that the increased risk of VT/VF occurrence can be represented by either prolongation or reduction of the parameters, based on certain abnormal range (Bartur et al. 2014). The increase or prolongation value of the parameters, including TpTe, heart rate-corrected QT interval (QT_c)/$QT_v i$/QT_d, turbulence onset (TO) of HRT, QT dynamicity, as well as TWA might indicate higher risk for the arrhythmias. Table 1 also reveals that the increased risk of the arrhythmias can also be represented by decrease or reduction in ECG parameters, such as HRV and turbulence slope (TS) of HRT. The iCEB is the most potential predictor for predicting an increased susceptibility to torsades de pointes (TdP) or non-TdP VT/VF based on its increase or decrease value (Mazandarani and Mohebbi 2018).

The increase or decrease of the parameters can be identified according to certain ranges. In reviewed literatures, researchers utilized two methods to define abnormal ranges. Firstly, the researchers match the parameters with predefined threshold values that indicate the abnormal ranges. The parameters that have abnormal prolongation prior to arrhythmias are iCEB, QT dynamicity, QT_c/$QT_v i$/QT_d, TO of HRT, TWA, and TpTe. Whereas, the parameters that have lower values than normal heart condition are HRV and TS of HRT. The parameters, such as fQRS, VLP and ER, in which their presence indicates the increased risk of VA, can also be identified using predefined thresholds (Zink et al. 2017).

Ion channels are membrane proteins that selectively regulate ion fluxes across the membranes of cells and cellular organelles, their gating mechanism depending on changes in membrane voltage, ligand binding, or physical and chemical stimuli (Szilágyi et al. 2018). The presence of distinct ion channel isoforms and their age-dependent and tissue-specific expression allows the fine regulation of many cellular functions, such as cell excitability, contraction, neurotransmitter and hormone release, gene expression, ion and water homeostasis. Given the pivotal roles played by ion channels and their extensive distribution, it is not surprising that mutations in ion channels' genes, or their interacting proteins, cause specific

inherited ion channelopathies, ranging from relatively common ones, such as idiopathic epilepsies, to very rare diseases. Despite the differences in their genetic origin and clinical setting, ion channelopathies share some common features regarding pathophysiology and therapeutic approach. In general, loss- or gain-of-function mutations translate into a principal organ-specific clinical phenotype.

Part 4 Late Potentials that Hide a Lot Information for Pharmacovigilance

Heart signal is captured by standard ECG machine in an old-fashioned way. That means it collects all kinds of noises from the contact of electrodes with the skin (muscle tremor, hair, and other skin coarseness) and also cable interference.

Signal-averaged electrocardiography, vastly known as late potentials, is recorded by three leads referred as Frank leads (XYZ leads), similar to the coordinate axis used in geometry. The fact that this method is based on the recording of three leads in anatomically orthogonal configuration was based on previously limited number of low noise amplifiers and slow computers at the time.

Cardiac late potentials signal and sources therefore must be acknowledged first, before even trying to analyze high-resolution ECG.

High-resolution electrocardiography consumes computer processing to record low-levels signals not normally observed on standard electrocardiographs. Cardiac late potentials occur at the end of or after the QRS complex and require these methods to be quantified. A brief overview of the methods used to record late potentials is presented. These include lead placement, computer-implemented signal averaging, high-pass filtering, and feature extraction for characterizing the late potentials.

The signal-averaged electrocardiogram (SAECG) was introduced over a decade ago and has been used extensively in clinical investigation and clinical practice to study the duration of ventricular activation. Major advantage of the SAECG over simple ECG recording is the ability to record low amplitude cardiac signals, in the microvolt range, which allows a true accurate measurement of ventricular activation delay, a prerequisite for ventricular reentry. We focus here on the currently accepted methodology and clinical applications of the time-domain SAECG and highlight its development from research technique to office test.

The methods for recording the SAECG have been reviewed in depth. Following is a brief overview of the methodology. Obtaining of SAECG is performed in two steps: acquisition and analysis. The acquisition begins with a three-lead, uncorrected XYZ lead set. The XYZ signals are directed to low-noise biophysical amplifiers. The key features of these amplifiers are that they meet the standards for leakage current and patient isolation and must be defibrillation protected. As these amplifiers are differential, they must be able to limit 60-Hz electrical line interference, which is often electrostatically coupled into the system. The signals should be directed through a bandpass filter that attenuates both low and high frequencies within specified limits. For the SAECG, the bandwidth is 0.05–300 Hz. For practical purposes, the 25–40 Hz low-pass frequency is more than adequate for the SAECG. In Holter systems where there is a poor bandwidth at the higher frequencies, caution should be used when these systems are used for late potential analysis. A frequency response <100 Hz may unduly distort, in a nonsystematic fashion, the SAECG. The next step in the SAECG process passes the signals through an analog-to-digital (A/D) converter, which transforms the tiny varying voltages in the digital domain of the computer. The sampling frequency is usually 1000 or 2000 times per second for each lead. The voltage resolution can be as little as 1 part in 400 (12-bit A/D) or as high as 1 part in 65,000 (16-bit A/D). While the 16-bit resolution is superior, there is little difference in these systems from a clinical diagnostic perspective. The first step in implementing the SAECG software is the detection and alignment of the QRS complexes. In some systems, this is a completely transparent process for the user, while in others it is possible to tune the process to enable greater accuracy and flexibility in signal averaging. By considering the shape of the QRS, one can accomplish two things: (1) a more accurate detection of the QRS by eliminating ventricular premature depolarizations, excessively noisy beats, and motion artifacts; and

(2) allow the system to finely align each QRS complex for the purposes of averaging. The most common method to incorporate the shape of the QRS is correlation. This process compares each incoming beat with a preselected template beat. In some commercial systems, the user will select the template, while in others the computer will choose the template. In some cases, the template is an average beat from the first 10 s of signal acquisition, or it can be changed dynamically and be based on the signal average itself after the first 10 or 20 beats are averaged. This initial part of the signal-averaging process could be considered a learning phase. If the template is automatically chosen, the learning process should be underscored in order to avoid cases where premature ventricular beats may be included in the template learning process. Once the computer has detected and aligned a beat, it will be added point by point for each XYZ lead. Dividing the sum by the number of beats will then result in a set of averaged recordings. As each beat is added, the noise is reduced in the signal averaged recordings due to the random nature of the noise. The signals that repeat on a beat-to-beat basis will emerge as the average of the noise goes to zero. This is the primary reason for using the signal-averaging method because very low level signals are usually masked in noise. Thus, standard ECG techniques are not adequate for recording these very low level signals. Theoretically, the noise will decrease by the square root of the number of beats averaged. If 100 beats are averaged, then the noise will be reduced by a factor of 10. In practice, this is only approximate because the characteristics of the noise may vary over time. The most significant source of noise in the SAECG is the signal generated by the chest muscles during normal physiologic processes such as breathing. The initial SAECG studies used a predetermined fixed number of beats for the average, e.g., 2000 or 300 beats per average. This proved to be unsatisfactory because each patient had his own noise and signal-level characteristics. The most common approach used to terminate an average with a measurable and consistent performance is to measure the noise during the averaging process until some predetermined level is reached. There are several approaches used for measuring residual noise as the averaging process is performed, and it is not possible to directly compare noise voltage values from the different matrixes. After one is assured that a high-quality, low-noise SAECG has been obtained from the patient, the next step is the analysis of the SAECG. The primary waveform analyzed for cardiac late potentials is the filtered vector magnitude, which is derived from the averaged XYZ leads. Noteworthy is that at the end of each QRS, one can appreciate small deflections. These are the late potentials, but they are difficult to discern and quantify. Each of these leads would go through processing with a high-pass filter which is implemented by the computer. The gain is further increased and the QRS complexes are further magnified and discrepancies found.

The low-amplitude signal (LAS) is the duration of the signal from >40 μV voltage at the end of the QRS to the QRS offset point. The QRS duration is a measure of the total ventricular activation time. That is, it portrays the time from the earliest ventricular activation to the time of the latest ventricular activation. The RMS 40 and LAS are waveform measures which do not directly relate to the electrophysiology of the heart as the QRS duration does. Essentially, a late potential appeals as a low-level "tail" alter the main body of the QRS complex. The RMS and LAS are designed as descriptors of this late potential tail.

When three orthogonal signal-averaged X,Y,Z leads are bandpass filtered digitally, they are combined into their vector magnitude waveform. The image surface is a geometric representation of the relationship between a fixed-position current dipole inside a volume conductor and the electric potential it produces on the boundary of the conductor. It is a powerful mean of conveying a wealth of information concerning some basic problems in electrocardiography, in terms of familiar heart-vector projection ideas. It gives deep insight into the nature of the relationship between torso surface voltages and the internal heart generator. Applications to the frontal-plane, Wilson central-terminal voltage and spatial vectorcardiography systems reveal shortcomings of current practices.

Several parameters are typically measured from this: QRSD, the filtered QRS duration,

RMS40, the amplitude of the terminal 40 ms of the QRS, and LAS (duration of the low amplitude signal <40 μV). Abnormal values of QRSD generally range from greater than 110 or 120 ms; abnormal values of RMS 40 range from less than 25 to 20 μV; and abnormal values of LAS typically are around 40 ms. These values have been derived empirically from a number of different clinical studies involving recent myocardial infarction, non-sustained ventricular tachycardia and syncope of unexplained origin. The gold standard for abnormality typically is not clinical events, but the induction of VT during EP studies.

High-Pass Filter for Baseline Wander Suppression and Lower-Frequency Cutoff

There are two important milestones for the technical requirements of lower-frequency cutoff in ECG instruments. The first is the 1975 American Heart Association (AHA) recommendations. Based on an analog single-pole filter with insignificant distortion of the ST segment and QT interval, the committee recommended a lower-frequency cutoff (3-dB down) of 0.05 Hz. Although this cutoff is too low to suppress most baseline wandering, it is still a classical reference basis to compare the performance of a digital filter at the current time.

The second is the 1990 AHA recommendations by an Ad Hoc Writing Group Committee of the AHA. Based on the availability of digital filters and electrocardiography systems, the lowest frequency components of the ECG, this document recommended a lower-frequency cutoff (3-dB down) at 0.67 Hz that corresponds to a heart rate of 40 beats per minute (bpm), and a 1 mV·s testing impulse for displacement and slope evaluation. Also, it requires less than 0.5 dB ripple over the range of 1 to 30 Hz.

Here we briefly review available high-pass filters (HPFs) that can be used to suppress baseline wander and setup lower-frequency cutoff. For an analog HPF, distortions of the ST segment will increase as the cutoff frequency increases above 0.05 Hz. There is a similar issue for a digital filter with nonlinear phase. Like its analog counterparts, an IIR digital filter with 0.5 Hz cutoff, for example, can make marked distortions to ST segments, although it has straightforward implementation especially when using design tools. On the other hand, an FIR digital filter can be designed with a linear phase, and the ST-segment distortion due to the phase issue is then completely gone. This filter is better able to preserve the fidelity of the ST-segment levels even when the cutoff is 0.5 Hz or higher. A linear FIR filter has longer delay and needs special design considerations. These three filter designs can work in real-time situations. A fourth filter design uses two IIR filters with identical designs, filtering once in a forward direction and once in the reverse direction. Using signal processing terms, it is a zero-phase filter, a special case of the linear-phase filter. Therefore, it does not have a phase distortion problem. Because this application requires the entire data segment to be available for reverse filtering, it cannot be accomplished in real time.

Compliance Review

With some modification and harmonization, the basic concepts of the 1990 recommendations were adopted into the current American Standard AAMI EC11 and international IEC 60601–2-51, both for diagnostic ECG units, as well as AAMI EC13 and IEC 60601–2-27, both for cardiac monitoring devices. One key examination from these harmonized standards is to check the displacement level (i.e., the ST-segment deviation) surrounding a single 0.3 mV·s impulse (3 mV for 100 ms): this impulse shall not produce a displacement greater than 0.1 mV. Then the question becomes: What should be the reference for the displacement limit measurement – the 0 level or the level of the impulse onset (to mimic the PR segment)?

Therefore, the question is whether the reference point from this analog filter test needs to be updated for linear-phase digital filter testing.

Furthermore, the standards require −3 dB at 0.67 Hz for monitoring ECG devices and − 0.9 dB (10%) at 0.67 Hz for interpretive ECG units, in

addition to the displacement requirement of 0.1 mV or less for a single 0.3 mV impulse (as above). If using the 0 level as the reference, the maximal cutoff frequency for a single-pole filter to meet this requirement is 0.05 Hz; for a linear-phase filter, it is around 0.25 Hz (far from 0.67 Hz). However, if the impulse onset as the reference is used, the results will be quite different, especially for a filter with linear phase.

Line-Frequency Filter for 50−/60-Hz Rejection

Power line frequency interference is a commonly encountered noise contamination in the clinical environment. Data acquisition analog hardware at a very early stage is developed to reduce the line-frequency interference by using common mode rejection circuitry design. However, a certain amount of line-frequency interference voltages still occurs in real clinical ECGs. To produce printouts with clean traces, typically, a line-frequency filter (LFF) is activated almost constantly during ECG acquisitions.

Nota Bene

International standard IEC 60601–2-51 has a section (§51.105.3.2) for notch ringing artifact. The test record ECG ANE20000 created by the European CTS-ECG project is artificial data with (*a*) nonzero J-point values in V2, V3, and V4 and (*b*) elevated and sloped ST segments. These three leads have larger QRS amplitudes, and the largest ringing artifacts would be mostly visible from them. §51.105.3.2 offers a testing procedure, including preparing two printouts with and without notch filter application and comparing the two ECG records. The requirement is "Notch FILTERS for line frequency interference suppression shall not introduce on the ECG RECORD more than 25 μV peak ringing NOISE in any LEAD when tested with the test ECG ANE20000." A 25 μV ringing response is only 1/4 mm tall. To measure peak value is technically correct, but practically the comparison is quite difficult to perform even if using a light table. It is partially due to ST-segment morphology of the testing waveform from V2 to V4, and partially because of the trace's line thickness and the small 1/4-mm limit.

International Standard IEC 60601–2-51. Medical electrical equipment – Part 2–51: particular requirements for safety, including essential performance, of recording and analyzing single channel and multichannel electrocardiographs. Geneva, Switzerland: International Electro-technical Commission (IEC); 2003.

Part 5 Conclusion

Human genome consists of more than 400 genes to encode ion channels, which are transmembrane proteins mediating ion fluxes across membranes. Being put across in total cell types, they are tangled in almost all physiological processes, including sense perception, neurotransmission, muscle contraction, secretion, immune response, cell proliferation, and differentiation. Due to the widespread tissue dissemination of ion channels and their physiological functions, mutations in genes encoding ion channel subunits, or their interacting proteins, are liable for inherited ion channelopathies. These diseases can range from common to very rare disorders and their severity can be mild, disabling, or life-threatening. In spite of this, ion channels are the primary target of only about 5% of the marketed drugs suggesting their potential in drug discovery.

The fact that standard ECG device comprises several limitations when it comes to signal acquisition due to large ambient noise and inherent limitations of technique applied, it is of utmost importance to modernize electrocardiography or at least to incorporate novel diagnostic techniques such as fragmented QRS (fQRS), heart rate variability (HRV), T peak-T end (TpTe), heart rate turbulence (HRT), and T wave alternans (TWA) that all have predictive value for pharmacovigilance. If not, we may declare facing ECG signal distortion (distinct from original heart signal data) not very precise for analysis.

Currently, there are ten ECG parameters that have potential values to detect underlying ventricular proarrhythmias in this respect. Much hidden information are present within heart rate-corrected QT interval (QT_c), QT interval variability (QT_vi), QT interval dispersion (QT_d), fQRS, early repolarization (ER), ventricular late potentials (VLP), HRV, HRT, index of cardio electrophysiological balance (iCEB), QT dynamicity, TWA, as well as TpTe. Each of these could provide with helpful information during clinical drug testing.

References and Further Reading

Ackerman MJ, Zipes DP, Kovacs RJ, Maron BJ, American Heart Association Electrocardiography and Arrhythmias Committee of Council on Clinical Cardiology, Council on Cardiovascular Disease in Young, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, and American College of Cardiology (2015) Eligibility and disqualification recommendations for competitive athletes with cardiovascular abnormalities: task force 10: the cardiac channelopathies: a scientific statement from the American Heart Association and American College of Cardiology. Circulation 132(22):e326–e329. https://doi.org/10.1161/CIR.0000000000000246. Epub 2015 Nov 2. Review

Arbustini E, Di Toro A, Giuliani L, Favalli V, Narula N, Grasso M (2018) Cardiac phenotypes in hereditary muscle disorders: JACC state-of-the-art review. J Am Coll Cardiol 72(20):2485–2506. https://doi.org/10.1016/j.jacc.2018.08.2182

Bartur G, Vatine JJ, Raphaely-Beer N, Peleg S, Katz-Leurer M (2014) Heart rate autonomic regulation system at rest and during paced breathing among patients with CRPS as compared to age-matched healthy controls. Pain Med 15(9):1569–1574. https://doi.org/10.1111/pme.12449. Epub 2014 Jul 24

Batul SA, Olshansky B, Fisher JD, Gopinathannair R (2017) Recent advances in the management of ventricular tachyarrhythmias. F1000Res 6:1027. https://doi.org/10.12688/f1000research.11202.1. eCollection 2017. Review

Campuzano O, Sarquella-Brugada G, Brugada R, Brugada J (2015) Genetics of channelopathies associated with sudden cardiac death. Glob Cardiol Sci Pract 2015(3):39. https://doi.org/10.5339/gcsp.2015.39. eCollection 2015. Review

Chen H, Chatelain FC, Lesage F (2014) Altered and dynamic ion selectivity of K+ channels in cell development and excitability. Trends Pharmacol Sci 35(9):461–469. https://doi.org/10.1016/j.tips.2014.06.002. Epub 2014 Jul 9. Review

Cheung CC, Laksman ZW, Mellor G, Sanatani S, Krahn AD (2016) Exercise and inherited arrhythmias. Can J Cardiol 32(4):452–458. https://doi.org/10.1016/j.cjca.2016.01.007. Epub 2016 Jan 14. Review

Crotti L, Kotta MC (2017) The role of genetics in primary ventricular fibrillation, inherited channelopathies and cardiomyopathies. Int J Cardiol 237:45–48. https://doi.org/10.1016/j.ijcard.2017.03.119. Review

D'Silva A, Sharma S (2014) Exercise, the athlete's heart, and sudden cardiac death. Phys Sportsmed 42(2):100–113. https://doi.org/10.3810/psm.2014.05.2062. Review

Fonseca DJ, Vaz da Silva MJ (2018) Cardiac channelopathies: the role of sodium channel mutations. Rev Port Cardiol 37(2):179–199. https://doi.org/10.1016/j.repc.2017.11.007. Epub 2018 Mar 7. Review. English, Portuguese

Garcia-Elias A, Benito B (2018) Ion channel disorders and sudden cardiac death. Int J Mol Sci 19(3):E692. https://doi.org/10.3390/ijms19030692. Review

Honarbakhsh S, Providencia R, Lambiase PD (2018) Risk stratification in Brugada syndrome: current status and emerging approaches. Arrhythm Electrophysiol Rev 7(2):79–83. https://doi.org/10.15420/aer.2018.2.2. Review

Imbrici P, Liantonio A, Camerino GM, De Bellis M, Camerino C, Mele A, Giustino A, Pierno S, De Luca A, Tricarico D, Desaphy JF, Conte D (2016) Therapeutic approaches to genetic ion channelopathies and perspectives in drug discovery. Front Pharmacol 10(7):121. https://doi.org/10.3389/fphar.2016.00121. eCollection 2016. Review

Kligfield P, Gettes LS, Bailey JJ et al (2007) Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. Circulation 115:1306

Lau JK, Sy RW, Corbett A, Kritharides L (2015) Myotonic dystrophy and the heart: a systematic review of evaluation and management. Int J Cardiol 184:600–608. https://doi.org/10.1016/j.ijcard.2015.03.069. Epub 2015 Mar 5. Review

Lieve KV, Wilde AA (2015) Inherited ion channel diseases: a brief review. Europace 17(Suppl 2):ii1–ii6. https://doi.org/10.1093/europace/euv105. Review

Mandala S, Di TC (2017) ECG parameters for malignant ventricular arrhythmias: a comprehensive review. J Med Biol Eng 37(4):441–453. https://doi.org/10.1007/s40846-017-0281-x. Epub 2017 Jun 28. Review

Mazandarani FN, Mohebbi M (2018) Wide complex tachycardia discrimination using dynamic time warping of ECG beats. Comput Methods Prog Biomed 164:238–249. https://doi.org/10.1016/j.cmpb.2018.04.009. Epub 2018 Apr 20.

Mazzanti A, Underwood K, Nevelev D, Kofman S, Priori SG (2017) The new kids on the block of

arrhythmogenic disorders: short QT syndrome and early repolarization. J Cardiovasc Electrophysiol 28(10):1226–1236. https://doi.org/10.1111/jce.13265. Epub 2017 Jun 27. Review

McNamara DA, Goldberger JJ, Berendsen MA, Huffman MD (2015) Implantable defibrillators versus medical therapy for cardiac channelopathies. Cochrane Database Syst Rev 10:CD011168. https://doi.org/10.1002/14651858.CD011168.pub2. Review

Posselt BN, Cox AT, D'Arcy J, Rooms M, Saba M (2015) Atrial and ventricular tachyarrhythmias in military personnel. J R Army Med Corps 161(3):244–252. https://doi.org/10.1136/jramc-2015-000494. Epub 2015 Aug 5. Review

Prakash K, Sharma S (2016) Interpretation of the electrocardiogram in athletes. Can J Cardiol 32(4):438–451. https://doi.org/10.1016/j.cjca.2015.10.026. Epub 2015 Nov 3. Review

Ricciardi D, Cavallari I, Creta A, Di Giovanni G, Calabrese V, Di Belardino N, Mega S, Colaiori I, Ragni L, Proscia C, Nenna A, Di Sciascio G (2016) Impact of the high-frequency cut-off of bandpass filtering on ECG quality and clinical interpretation: a comparison between 40Hz and 150Hz cut-off in a surgical preoperative adult outpatient population. J Electrocardiol 49(5):691–695. https://doi.org/10.1016/j.jelectrocard.2016.07.002. Epub 2016 Jul 18

Schneider F, Martin J, Hapfelmeier A, Jordan D, Schneider G, Schulz CM (2017) The validity of linear and non-linear heart rate metrics as workload indicators of emergency physicians. PLoS One 12(11):e0188635. https://doi.org/10.1371/journal.pone.0188635. eCollection 2017

Schwartz PJ, Ackerman MJ, Wilde AAM (2017) Channelopathies as causes of sudden cardiac death. Card Electrophysiol Clin 9(4):537–549. https://doi.org/10.1016/j.ccep.2017.07.005. Review

Singh P, Pradhan G (2018) Variational mode decomposition based ECG denoising using non-local means and wavelet domain filtering. Australas Phys Eng Sci Med 41(4):891–904. https://doi.org/10.1007/s13246-018-0685-0. Epub 2018 Sep 6

Skinner JR, Winbo A, Abrams D, Vohra J, Wilde AA (2019) Channelopathies that lead to sudden cardiac death: clinical and genetic aspects. Heart Lung Circ 28(1):22–30. https://doi.org/10.1016/j.hlc.2018.09.007. Epub 2018 Oct 4. Review

Spears DA, Gollob MH (2015) Genetics of inherited primary arrhythmia disorders. Appl Clin Genet 8:215–233. https://doi.org/10.2147/TACG.S55762. eCollection 2015. Review

Szilágyi J, Walters TE, Marcus GM, Vedantham V, Moss JD, Badhwar N, Lee B, Lee R, Tseng ZH, Gerstenfeld EP (2018) Surface ECG and intracardiac spectral measures predict atrial fibrillation recurrence after catheter ablation. J Cardiovasc Electrophysiol 29(10):1371–1378. https://doi.org/10.1111/jce.13699. Epub 2018 Aug 23

van der Bijl P, Delgado V, Bax JJ (2017) Sudden cardiac death: the role of imaging. Int J Cardiol 237:15–18. https://doi.org/10.1016/j.ijcard.2017.03.010. Epub 2017 Mar 6. Review

Vranic I. Ivana (2013) Signaling prodromes of sudden cardiac death. Bosnian J Basic Med Sci 13(1):44

Vranic I. Ivana (2017) Electrocardiographic appearance of aortic stenosis before and after aortic valve replacement. Annals of Noninvasive Electrocardiology 22 (5): e12457

Vranic I. Ivana (2017) Book "Sudden cardiac death, predictors, prevalence and clinical perspective". Nova Science Publishers, New York, p 1–293

Vranic I. Ivana, Simic T (2012) Sudden cardiac death syndrome - arrhythmogenic right ventricular dysplasia/cardiomyopathy as most frequent cause of fatal arrhythmias in the Book Dysplasia: Causes, Types and Treatment Options, Eds. Laurel M. Sexton and Hershel J. Leach. Nova Science Publishers, New York, p 129–164

Vranic I. Ivana, Mihailo Matic, Jovan Perunicic, Tijana Simic, Ljiljana Soskic, Natasa Milic (2006) Adenosine cardioprotection study in clinical setting of paroxysmal supraventricular tachycardia. Prostaglandins, Leukotrienes and Essential Fatty Acids 74(6):365–371

Vranić I, Vranić I, Antić B, Stojanović G, Al-Salami H (2019) Influence of the Main Filter on QRS-amplitude and Duration in Human Electrocardiogram, Meas Sci Rev 19(1):29–34

Wang S, Li L, Tao R, Gao Y (2017) Ion channelopathies associated genetic variants as the culprit for sudden unexplained death. Forensic Sci Int 275:128–137. https://doi.org/10.1016/j.forsciint.2017.03.006. Epub 2017 Mar 21. Review

Wang Z, Zhu J, Yan T, Yang L (2019) A new modified wavelet-based ECG denoising. Comput Assist Surg (Abingdon) 28:1–10. https://doi.org/10.1080/24699322.2018.1560088. [Epub ahead of print]

Zink MD, Brüser C, Stüben BO, Napp A, Stöhr R, Leonhardt S, Marx N, Mischke K, Schulz JB, Schiefer J (2017) Unobtrusive nocturnal heartbeat monitoring by a ballistocardiographic sensor in patients with sleep disordered breathing. Sci Rep 7(1):13175. https://doi.org/10.1038/s41598-017-13138-0.

Pharmacodynamic Evaluation: CNS Methodologies

4

Lynne Hughes, Marie Trad, Stacey Boyer, Deborah Lee, and Wei Yin

Contents

L. Hughes (✉)
Neurology Center of Excellence, IQVIA, Reading, Berkshire, UK
e-mail: lynne.hughes@iqvia.com

M. Trad
Neurology Center of Excellence, IQVIA, Paris, France
e-mail: marie.trad@iqvia.com

S. Boyer
Clinical Research Operations, Cavion, Charlottesville, VA, USA
e-mail: boyer@cavionpharma.com

D. Lee
CNS-TAU, Takeda Pharmaceutical International Co., Cambridge, MA, USA
e-mail: leezwiz@yahoo.com

W. Yin
Quantitative Clinical Pharmacology, Takeda Pharmaceuticals, Cambridge, MA, USA
e-mail: wei.yin@takeda.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_32

Abstract

Drug discovery in neurology is confronted with a high attrition of compounds in both early and mid-stages of the development cycle. The causes of these failures are multiple and mainly based on the uncertainty around the precise pathological processes, lack of reliable biomarkers, and variability of drug penetration through the blood-brain barrier (BBB) into the brain. The first section of this chapter focuses on pharmacokinetic/pharmacodynamic (PK/PD) aspects when developing a medication targeting brain disorders. The biomarker section discusses those markers employed in the most common neurological disorders. It addresses their advantages and limitations, and how they serve in confirming efficacy and, in some instances, the safety of therapeutic interventions.

Pharmacokinetics and Pharmacodynamics in Neurology

Stacey Boyer, Deborah Lee, and Wei Yin

Introduction

Drug development efforts to treat neurological disorders have continuously faced significant challenges, leading to a lower success rate in comparison with other therapeutic areas (Reichel 2009). Of these difficulties, measurement and accurate prediction of the free drug concentrations at the site of action within the brain, and identification of pharmacodynamic (PD) biomarkers for the mechanism of action and drug effect are of particular focus. Since pharmacological activity depends on the free drug concentrations at the site of action, it has become critical to optimize unbound drug concentrations in the central nervous system (CNS), most notably at the drug target. The unbound drug concentrations can be measured in both in vitro and in vivo preclinical models (Reichel 2015). In patients with neurological diseases, obtaining information with regards to substance penetration into the CNS and target exposure requires application of biological sampling and translational techniques (de Lange 2013). Despite these limitations, there is opportunity to apply novel concepts and optimize existing methodologies and tools to assess drug penetration and pharmacokinetics (PK) in the human brain (Reichel 2015; Rizk et al. 2017).

Factors that Control Target Drug Exposure in the CNS

Critical factors that control the free drug exposure at the CNS target include systemic PK, plasma protein binding, rate and extent of CNS penetration, CNS distribution, binding to brain tissues, and elimination (de Lange 2013; Reichel 2015). Systemic PK and plasma protein binding determine the unbound free plasma concentration. Since only the unbound drug is able to pass through membrane barriers (e.g., the BBB), it is important to measure the unbound plasma concentration to understand drug transport to the brain (de Lange 2013). The rate of CNS penetration is controlled by cerebral blood flow and permeability across the BBB. The extent of CNS penetration is not determined by the permeability and the penetration rate: rather, it is measured by the unbound brain to unbound plasma concentration ratio (Reichel 2015). CNS distribution is also the key element in determining the effective concentration at the CNS target. Drug elimination from the CNS via extracellular fluid (ECF) bulk flow into the cerebrospinal fluid (CSF) or across the BBB

reduces drug concentration at the target site (de Lange 2013; Reichel 2015). Drug metabolism at the BBB and blood-CSF barrier may also influence CNS distribution accordingly (de Lange 2013).

Methods to Evaluate Brain Penetration and PK in the CNS

Evaluation of brain penetration and PK in the CNS is conducted via in vitro and in vivo studies, in situ brain perfusion, and integrated quantitative modeling approaches (Alavijeh et al. 2005; Danhof et al. 2007; de Lange 2013; Reichel 2009, 2015; Rizk et al. 2017; Yamamoto et al. 2017). In vitro studies include permeability and transporter assays using Caco-2 and MDCK-MDR1 cells, and binding assays in plasma, brain homogenate, or brain slices (Reichel 2009). In situ brain perfusion in whole animals can also provide a kinetic measure of drug entry rate into the brain (Alavijeh et al. 2005). A notable in vivo study is the determination of the unbound brain/plasma ratio in rodents, which provides information on the extent of brain penetration. In addition, CSF samples are of high value since they can be collected in humans to provide information on unbound drug concentration in the CNS.

To better understand the concentrations over time in ECF, a key compartment for the effect of many drugs, brain microdialysis in rodents can be applied to calculate PK parameters such as Cmax, $t_{1/2}$, and AUC (Reichel 2009; Alavijeh et al. 2005). Physiologically-based PK (PBPK) models have also become increasingly important to distinguish between system- and drug-specific parameters to allow for translational prediction of human CNS PK and associated drug effect from preclinical models accordingly (Yamamoto et al. 2017; de Lange 2013).

Translating PK/PD to Address Neurological Disease

As discussed in further detail in due course, translational approaches to predict exposure and distribution in the human brain present both inter- and intraindividual variation when investigating therapeutic modalities targeting the CNS (de Lange 2013). Addressing these variations is required to understand PK and effects in the CNS, as they differ between diseases and species (de Lange 2013). Consideration must also be paid to the variation in PK properties of different compounds as they relate to defining dose regimens and their pharmacological and/or toxicological effects (Dingemanse et al. 1998). Both PK and PD properties of the drug under study should be well defined in terms of concentration, dose-response relationships, and target site kinetics as they relate to the translation of preclinical models to clinical study (Dingemanse et al. 1998).

Target Exposure and Target Engagement in the CNS

Both target exposure and target engagement are required to foster target-mediated pharmacology and therapeutic effect(s) on the etiology of the neurological disease (Reichel 2015). Target exposure lends itself to PK optimization efforts focused on both unbound and total drug concentrations (Reichel 2015). In the absence of minimally required exposure, the drug will be unable to mediate the necessary required pharmacology to yield the desired effect(s) on the neurological disease under study (Reichel 2015). Target engagement requires binding of the drug to the target protein at concentrations in excess of the pharmacological potency of the drug over time (Reichel 2015). Within the CNS, target site (s) can be incredibly challenging for drug access, as the BBB limits exposure as a result of tight cell junctions and multiple efflux transporters (Reichel 2015). As a result, neurological drug discovery has focused on assessing target engagement via brain tissue sampling in preclinical models of CNS diseases in humans: in addition, indirect methodologies such as imaging modalities can provide demonstration of target engagement (Durham and Blanco 2015).

Translational Approaches

Translation from preclinical models to the clinic must account for drug potency to normalize across species for differences in target receptor inhibition,

PK to inform differences in dose-exposure, and level of the PK/PD relationship to facilitate prediction of CNS penetration and efficacy (Di and Kerns 2015). Limitations include contributing factors in the drug concentration-time profile which influence PK in the CNS, and physiological volumes of CSF: intracellular fluid and interstitial fluid should also be considered when translating exposure to efficacy (Di and Kerns 2015). Integrated PK/PD modeling essentially serves as a valuable tool, merging quantitatively driven in vitro pharmacologic properties of a drug with its respective in vivo PK to investigate the exposure-response (E-R) relationship (Tuntland et al. 2014).

Determining Target Engagement in the CNS via Positron Emission Tomography

While scaling of PK properties is executed via application of standard allometric principles, PD properties are often species-independent as related to receptor occupancy (Melham 2013). A noninvasive approach to determining target engagement in the CNS is the application of positron emission tomography (PET), an imaging technique which enables translation of ex vivo receptor occupancy from preclinical models to humans (Melham 2013). Application of PET provides a quantification of brain metabolism, receptor binding of various neurotransmitter systems, and of alterations in regional blood flow (Politis and Piccini 2012). Translation of research into clinical populations may assist with differential diagnosis and narrowing of complexities in patients with neurological disorders such as mild cognitive impairment and dementia (Johnson et al. 2013).

PK/PD Relationships in Humans

The fundamental understanding of PK/PD or E-R relationships is not always straightforward since it relates to a pharmacological or toxicological response relative to the drug concentration at the site of action. The drug concentration needs to reach a minimal level to obtain a response. E-R relationships can be complex since they are not always direct and rapidly reversible (Dingemanse et al. 1998). In addition, various factors can impact both PK and E-R relationships, including age, sex, race, disease status, chronic drug use, and drug-drug interactions. As more detailed information on the PK and E-R relationships becomes available for CNS drugs, well-defined E-R relationships have been used to optimize dose regimen in neurology clinical trials. This is notably so in the late phase to increase probability of success and to extrapolate the dose regimen in special populations (e.g., elderly, pediatric) (Dingemanse et al. 1998).

PK/PD Effects of Drugs in Neurological Disorders

Approved cholinesterase inhibitors (i.e., tacrine, donepezil, rivastigmine, and galantamine) for first-line treatment for the symptoms of Alzheimer's disease (AD) patients demonstrate a dose-related effect on the desired therapeutic outcome of improved cognition and functional activities, as well as the mechanism-based gastrointestinal adverse events.

The application of antiepileptics to treat neonatal neurological diseases has been very challenging owing to the lack of reliable data to inform safe and effective dose regimens in this patient population. The majority of the drugs utilized are done so in an off-label manner, leading to scaling from adult dose based on body weight or body surface area (Donovan 2015). As a result of these challenges, quantitative modeling approaches become essential to optimize the dose selection for the neonatal population. For example, a population PBPK model simulation suggested that only 10% of the adult dose of lorazepam needs to be given to newborns to demonstrate antiepileptic activity (Donovan et al. 2015; Maharaj et al. 2013).

PK/PD data following administration of antiretrovirals indicate highly variable drug transfer to the CNS, as CNS penetration is dependent on both drug and patient characteristics (Calcagno et al. 2014). The use of drugs with high penetration and compartmental activity has been associated with optimized CSF viral control, and in some cases, enhanced neurocognitive activities (Calcagno et al. 2014).

PK/PD Modeling of Neurology Drug Effects: Utility and Implications

As discussed, developing drugs to treat neurological disorders involving the CNS is particularly challenging, often because of many factors including the complexity of brain function, relative isolation of the CNS from peripheral drug compartments, lack of definitive pathology, inability to clearly monitor the effects of drugs, lack of biomarkers, both inter- and intra-variability of drug delivery to the CNS, difficulty of prediction of appropriate dosing in alternative populations such as children, and inability to predict both alternative on- and off-target effects. In particular, the first hurdle is often ensuring that the drug reaches the target "at the right place, at the right time, and at the right concentration" (de Lange 2013).

In this section, the ability of PK/PD modeling to help predict brain distribution, kinetics, and therapeutic effects, how PK/PD models can be used to predict various effects of differing formulations over a broad dosing range, and how PK/PD extrapolation can lead to regulatory approvals will be discussed.

Accessing the CNS

The first hurdle facing drugs to treat CNS disorders is successful access to the appropriate CNS compartments, i.e., the brain extracellular fluid compartment, brain intracellular compartment, and ventricular and lumbar CSF compartments. Mechanisms of crossing the BBB and the blood-CSF barrier include passive diffusion of unbound molecules across a concentration gradient, facilitated diffusion across a concentration gradient using a helper molecule, vesicular transport, and active transport, which may occur against a concentration gradient but requires energy. Even if the drug gains access to the CNS, drug effects may be independently modified by CSF turnover and ECF bulk flow, extra-intracellular exchange and brain tissue binding, drug metabolism, target interaction, and signal transduction and homeostatic processes, all of which can vary between individuals and disease states. Other individual factors that can play a role in determining drug effect include genetic background (e.g., fast vs. slow metabolizers), sex, age, diet, and long-term drug treatment with the drug of interest or concomitant medications (i.e., up- or down-regulation of receptors). Finally, while most of the elegant and informative PK/PD studies are carried out in animals, interspecies differences are well known.

Because of these many variables, in vivo experiments need to be integrated. Data should be collected examining multiple different conditions in the same animals, employing study designs in which as many conditions that can lead to variability as possible are taken into account. These may include, but are not limited to, protein binding, inhibition of transporters and receptors, and/or existence of the disease state which closely mimics the human disease. In addition, in vivo dialysis methods in animal models may establish the effects that bound and unbound drugs may have in CNS drug delivery.

The Mastermind Approach

To help understand the complexity of the integrated studies and the data expected to result from them, de Lange (2013) proposed using the Mastermind approach. The game Mastermind is a unique code breaking/code making board game in which colored pegs are used by one player to establish a code and by the other player to systematically break the code. In a similar manner, de Lange (2013) hypothesized that the "code" represents a complex PK/PD relationship. By understanding the strict distinction between the properties of the drugs and the properties of the biological systems by careful design of integrated studies and by utilizing mathematical modeling, the "code" can be broken. De Lange (2013) uses this Mastermind approach and detailed PBPK modeling to understand drug effects of acetaminophen, quinidine, and remoxipride.

Developing PK/PD Models from Human Data to Inform Future Clinical Trials

While de Lange (2013) proposed detailed integrated animal studies and mathematical modeling to predict CNS effects in humans, studies in

humans can also be used to build PK/PD models. These models can then be employed to predict time of onset and expected effects of alternative dosing regimens not tested. In one example, Wiltshire et al. (2012) reported such a model based on a Phase 1 single-ascending-dose study of remimazolam, a rapidly metabolized benzodiazepine when compared with placebo and midazolam. In this randomized, single dose-escalation study, 54 healthy subjects in 9 groups received a single 1-min IV infusion of remimazolam (0.01–0.3 mg/kg), while 18 subjects received midazolam or placebo. All data were used for Monte-Carlo simulations of alternative dosing regimens. A 4-compartment pharmacokinetic model of midazolam and a physiologically based recirculation model of remimazolam best fit the data. Simulations based on these models demonstrated that remimazolam delivered extremely rapid sedation and that dosing by body weight offered no advantage over the weight range studies (65–90 kg) suggesting a fixed dose regimen.

Sophisticated PK/PD modeling has been accepted by regulatory agencies for extrapolation of efficacy from one population to another. In 2017, the International Conference on Harmonisation (ICH) revised a guideline that states as follows: "When a medicinal product is to be used in the pediatric population for the same indication(s) as those studied and approved in adults, the disease process is similar in adults and pediatric patients, and the outcome of therapy is likely to be comparable, extrapolation from adult efficacy data may be appropriate." (ICH 2017). In 2012, Pellock et al. demonstrated that antiepileptic drugs from published trials in focal seizures showed clinical responses that were similar in both children and adults based on comparable PD effect sizes. Based on this, the Pediatric Epilepsy Academic Consortium on Extrapolation (PEACE) was formed. After clear demonstration using both animal models and information in humans that the physiological processes which lead to focal seizures in adults are similar to those in children as young as 2 years of age, there was regulatory agreement that sophisticated PK/PD models could be used to extrapolate efficacy from the adult population to the pediatric population (Pellock et al. 2017). However, it was recognized that enough PK data needed to be obtained from pediatric populations to determine actual exposure which then can be used to predict efficacy based on PK/PD models, and that extrapolation could not be used to determine safety.

A novel use of PK/PD modeling led to approval of vigabatrin for pediatric patients with partial seizures (Nielson et al. 2014). Previously, vigabatrin had been approved for pediatric patients between 1 month to 2 years of age with infantile spasms, and as adjunctive therapy in adult patients with complex partial seizures who had responded inadequately to several alternative treatments. Upon approval, the FDA issued a Pediatric Research Equity (PREA) requirement for children with refractory partial seizures. Three previous Phase 3 trials had been initiated but were halted early for administration reasons. These data were pooled with adult pivotal clinical trial data to develop a population dose-response model linking vigabatrin dosage and seizure counts. This model allowed prediction of appropriate dosing based on body weight. Submission of these data along with the model simulations allowed for the successful fulfillment of the PREA requirement and approval by FDA of vigabatrin in pediatric patients with complex partial seizures.

Summary

Discussions to this point in the chapter demonstrate that while more complicated than other systems, PK/PD modeling can be successfully utilized to better understand the significant complications involved in achieving CNS drug delivery to the correct location at the correct time and with the correct concentration. Data obtained in limited clinical studies can be expanded using these models to predict the full extent of dosing effects and even lead to regulatory approval. Attention now turns to biomarkers.

Biomarkers in Neurology

Lynne Hughes and Marie Trad

Introduction

With the increased global life expectancy, the prevalence of neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD) is expected to grow significantly in the coming years. The estimated number of patients with AD is expected to triple in the US by 2050 to attain 13.7 million people afflicted by the disease, which will bring an increased economic burden to the health care systems (Hebert et al. 2013).

Regardless of the efforts invested in drug development in neurosciences, there remains a high unmet therapeutic need in multiple neurological indications. Drug discovery has been characterized with a high attrition rate, in general, and this is exemplified in CNS, as success rates drop from an average 11–8% (Kola and Landis 2004). Multiple factors contribute to this high attrition and failure rate: lack of adequate animal models, poor crossing of drug through the BBB and insufficient drug penetration into the brain, poor knowledge of pathophysiological mechanisms, increased placebo responses, and, most prominently, lack of specific biomarkers to measure therapeutic effect (Pangalos et al. 2007). However, significant recent progress has been achieved in the biomarker field with the development of the "omics" technologies: genomics, proteomics, metabolomics, lipidomics, and immunological and biological epigenetics (Sethi and Brietzke 2016). The advancement of the neuroimaging techniques has further dramatically contributed to successful drug development in indications such as multiple sclerosis (MS), with 16 disease modifying therapies (DMTs) approved to date. This is considered as quite an accomplishment in view of low success rates of investigated DMTs in more prevalent diseases such as AD and PD. This has led to the increased interest in developing highly targeted biomarkers which will both contribute to reducing attrition and costs as well as, potentially, speed up the drug development processes (Frank and Hargreaves 2003).

This chapter reviews the more widely assessed biomarkers in the major neurological diseases including AD, PD, MS, and neuro-orphan indications.

Biomarker Definition

As per the Biomarkers Definitions Working Group, a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathologic processes, or biological responses to a therapeutic intervention. Change in biomarkers after treatment will help identifying a treatment benefit and safety issues potentially related to that same treatment. An exemplary biomarker should be accessible, reproducible, quantifiable, and easy to apply in the studied patient populations. A marker can be a surrogate to replace a clinical endpoint and give indication on disease modification versus symptomatic effect. A PD marker is a biomarker of pharmacologic response. Both surrogate and PD markers are a subset of biomarkers (Biomarkers Definitions Working Group 2001; FDA 2014).

Biomarkers in Neurological Disorders

The need for the development of DMTs in multiple neurological diseases warrants the use of targeted biomarkers that are both diagnostic of early disease and predictive of drug effect. The inclusion in AD clinical trials of prodromal or preclinical/asymptomatic patients necessitates the incorporation of multiple biomarkers both to permit identification of such subjects and to enhance the chances of proving a subsequent drug effect. A combination of biochemical, neuroimaging, and clinical assessments has become a standard approach in clinical trials, especially in early stages of drug development. Choosing the most optimal combination of biomarkers is critical in early proof-of-concept studies to increase clinical trial success. The selection of the most appropriate biomarkers, whether in isolation or combination, becomes even more relevant when

testing DMTs to differentiate symptomatic effects when compared with longer-term disease modification effects.

General Biomarkers in Neurology

This section summarizes the frequently used biomarkers in neurology while focusing more specifically on the major diseases such as AD, PD, and MS (Giacomelli et al. 2017; Henley et al. 2005; Silva and Furie 2009).

Cerebro-spinal Fluid

Although lumbar punctures are invasive procedures and a number of findings remain to be confirmed as reliable, cerebral-spinal fluid (CSF) biomarkers are relevant in indications such as AD where a decrease in Abeta 1-42, as well as an increase in Tau is observed (Ittner and Götz 2011). In addition, a reduction in total alpha-synuclein levels in PD, the increase of neurofilaments levels in amyotrophic lateral sclerosis (ALS), as well as the increase in various inflammatory markers in MS have been explored in many studies. The importance of these biomarkers and the need for their use to monitor drug effect warrants the integration of this procedure in clinical trials, while mitigating the invasive aspect of lumbar punctures (LPs).

In AD, for those sites who do not have access to PET scanning facilities and/or access to the PET ligands for amyloid detection, one option used by some pharmaceutical companies is to use CSF for confirmation of amyloid levels as an entry criterion for clinical trials that are utilizing an amyloid-targeting investigational product (ClinicalTrials.gov). As mentioned below, compliance is increasing with this modality and is generally higher in subjects with preclinical and/or prodromal AD than in subjects with mild to moderate disease (IQVIA proprietary database). LPs are not usually standard of care for AD subjects for most countries, and as such, it has taken a number of years to ensure compliance with this assessment (Hughes et al. 2014). However, in some countries, e.g., Sweden, LPs are routinely performed in the majority of subjects presenting with AD or possible AD. Recently, with the advent of the potential DMTs in PD, LPs are now being utilized in these early stage trials, leading to greater challenges than those facing AD trials more than 10 years ago. As CSF collection is not a standard of care in the PD population, other noninvasive procedures are being studied to avoid multiple LPs in these patients.

Neuroimaging

Multiple imaging techniques have contributed significantly to the advancement of drug development in Neurology. The demonstration of magnetic resonance imaging (MRI) as a highly reliable biomarker is most prominent in MS, allowing the approval of multiple drugs, which is without precedent in Neurosciences. The most frequently used MRI sequences are listed here:

MRI with gadolinium: T1 sequences with or without gadolinium enhancement are essential in monitoring the anti-inflammatory drug effect in MS, as gadolinium enhancement is a direct measure of the BBB disruption due to the inflammatory processes.

Structural and Volumetric MRI: Increasingly used to assess the changes of specific anatomical structures of the brain as a result of pathological disease processes. The measuring of medial temporal and hippocampal atrophy in AD is a reliable marker of disease progression (Riascher et al. 2009).

Other MRI techniques such as functional MRI (fMRI), magnetic resonance spectroscopy (MRS), and diffusion-weighted magnetic resonance imaging have been studied in AD and mild cognitive impairment (MCI), Huntington's disease (HD), PD, MS, and stroke.

Positron emission tomography (PET): PET scanning is widely used in a number of neurological indications including AD and MCI (reduced glucose uptake), PD, and HD. FDG PET has been shown to have a good sensitivity to detect brain dysfunction and early changes in AD and to follow its evolution over time (de Leon et al. 2001).

SPECT/DaTScan: This imaging technique uses a transporter to bind to the DAT in the striatum and then SPECT visualizes the amount of transporter present. It allows for the differential diagnosis between PD, essential tremor, multiple system atrophy (MSA), and other Parkinsonian syndromes.

Genomics

The advancement in genetic research has been instrumental in the understanding of multiple neuro-orphan diseases such as ALS (SOD1, C9orf), spinal muscular atrophy (SMA: SMN1 gene), and Duchenne muscular dystrophy (DMD: various exons). This has allowed the development of genome-targeted therapies, with recent significant therapeutic successes in SMA. Further research is critical to bring therapies to patients with rare disease.

Peripheral Biomarkers

Multiple biomarkers have been explored in peripheral fluid such as plasma, red blood cells, and saliva. Confirmation of their reliability is needed to avoid invasive CSF collection in patients in whom LPs are not part of the standard of care. An example of the relevance of peripheral biomarkers is Parkinson's disease as described next.

Clinical Biomarkers

The change in clinical biomarkers to monitor drug effect is a required primary objective of many pivotal trials. The mainstay of assessment for clinical trials in AD is still the use of a number of neurocognitive batteries, with changes in other biomarkers being used as secondary or exploratory endpoints. The use of the Expanded Disability Status Scale (EDSS) remains the gold standard in the observance of disease disability in MS as a hallmark of disease progression. In Parkinson's disease, multiple scales and corresponding subscales, of which most prominently the UPDRS, have been used to assess the effect of drugs in development, on motor function, and for drug-induced complications. Functional rating scales have also been widely used in a variety of neurological clinical trials, as a primary or secondary endpoint.

Electrophysiology

Electroencephalography (EEG), a noninvasive and cost-effective procedure, has proven to be a reliable biomarker in diseases such as AD, epilepsy, pain, PD (REM behavior disorder), and autism. The progress in EEG technologies will provide superior evaluation of central drug effect (Jobert et al. 2012). Although not widely performed in pivotal studies, its use is cost-effective and noninvasive.

Electromyography and electrical impedance myography (Rutkove et al. 2012), the motor-unit number estimation (MUNE), allows the monitoring of disease progression in ALS (Bowser et al. 2006).

Biomarkers in Alzheimer's Disease

There have not been any successful new drugs developed for AD for more than a decade. From 2002 to 2013, 96.4% of the 244 agents assessed in more than 1,000 clinical trials in AD subjects failed to achieve their primary endpoints (Ousset et al. 2001). These trials have been increasingly complex, requiring increased number of sites and assessments and a lower number of subjects overall per site, and subsequently have resulted in higher screen failure rates and, oftentimes, drop-out rates (Ousset et al. 2001). In addition, these trials are becoming more costly, and, in particular, the cost for screen failure rates is prohibitive for many companies.

The combination of biomarkers has been used not only for secondary or exploratory endpoints but also to enrich the clinical trial population, allowing for the most appropriate subjects to be exposed to the clinical trial investigational products and thus maximizing the probability of seeing a signal in the trial data.

The use of biomarkers as entry criteria increased following the data released in 2010 on Semagacestat (Doody et al. 2013) and on Bapineuzamab (Salloway et al. 2014). Both investigational products targeted amyloid, and the ultimate data analysis revealed that approximately 30% of the recruited suspected mild-moderate AD subjects did not, in fact, have amyloid

"confirmed" AD. This enrichment technique has proven to be necessary in large pivotal AD programs and has been found to be useful in successfully identifying AD patients who have a higher chance to benefit from an amyloid-targeting agent.

Subsequently, trials that had an investigational product targeting the amyloid pathway required that subjects have a baseline amyloid positivity before trial entry to ensure that the investigational product did have a target for engagement. However, for the mild to moderate population, this added a further 30% onto the screen failure rate of AD trials, in addition to the 20–30% already observed (Hughes et al. 2014). This effect of enrichment is even more pronounced in the earlier stages of AD with prodromal AD having, overall, a screen failure rate of 70–80%, and preclinical AD being even higher at 90% or more (Hughes et al. 2014). This all leads to an overall effect of decreasing recruitment rates at sites and increasing trial duration and costs.

Amyloid is assessed either via PET imaging with a fluorine-18 labeled ligand or via a lumbar puncture and collection and analysis of CSF. There are basically 3 amyloid ligands available: Avid (Florbetapir), GE (Flutametamol), and Piramel (Florbetaben). The geographical location of labeling centers dictates the site and country distribution for each imaging modality. Thus, there are many global sites and countries where the AD subjects are unable to participate in trials which utilize amyloid PET imaging as a screening tool as these countries either do not have access to a PET scanner and/or a cyclotron. The ligand is generally provided to the sites ready-labeled for immediate administration to the subject and the time from labeling, QC, and subject administration is approximately 4 h due to the half-life of the fluorine-18 (Jobert et al. 2012).

Alternatively, amyloid positivity can be accessed via LPs, which brings with it its own challenges. A review by IQVIA in 2013 of compliance with at least 2 LPs in 12 AD programs, which recruited in excess of 7500 subjects, showed a global compliance of 8% in 2007 increasing to 50% by 2012 (IQVIA proprietary database). This compliance rate is steadily increasing, but is still not ideal (Jobert et al. 2012).

These imaging and CSF biomarkers are being used primarily to ensure that the appropriate subjects are being recruited into the trials. They are not being used as primary endpoints as the regulators still request that a clinical trial in AD shows a positive clinical benefit to a subject and not just a change in nonclinical biomarkers.

An ongoing academic trial, however, utilizes a combination of biomarkers as primary endpoints. This trial is enrolling study participants who are biological adult children of a parent with a mutated gene known to cause dominantly inherited AD. Such individuals may or may not carry the gene themselves and may or may not have disease symptoms. However, it is known that the deposition of amyloid builds up some 10–20 years before clinical symptoms are manifest. Therefore, while utilizing potential disease modifying drugs which remove amyloid or prevent its deposition or build up is being used as surrogate endpoint as for many subjects, the neurocognitive decline is minimal at this very early stage (Bateman et al. 2017). This trial, which is assessing the impact of three investigational products versus placebo, is still recruiting and is being seen almost as a proof-of-concept trial to assess the utility of various AD biomarkers as a key outcome.

Other biomarkers used in AD trials include the assessment of Tau and p-Tau levels in the CSF, and recently, a number of Tau PET ligands have been developed and are now in clinical trials. Tau pathology appears to closely correlate with neurodegeneration and onset of clinical dementia, with dynamic changes in tau pathology being evident at the minimal cognitive impairment (MCI) stage (Lambracht-Washington and Rosenberg 2013). Therefore, this represents an interesting target for pharmaceutical companies, and there are a number of products in early stage clinical trials looking at the possibility of actively or passively immunizing against the various tau biomarkers (IQVIA proprietary database).

Recent advances in AD biomarkers include a race to develop a blood test for amyloid, as PET imaging with an amyloid ligand is not practical on

a global scale. If confirmed as a reliable biomarker, this will provide a simple, cost-effective assessment to be used for amyloid level assessment for detection of patients with either prodromal or very early disease and thus enable a physician to specifically target therapy with an appropriate drug.

Biomarkers in Parkinson's Disease

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder after Alzheimer's disease and affects approximately 1% of the population older than 60 (Tysnes and Storstein 2017). The main hallmark of the disease is loss of dopaminergic neurons in the mesencephalon resulting in the clinical presentation of the disease characterized by 3 cardinal symptoms: bradykinesia, extrapyramidal rigidity, and resting tremor. By the time the patient becomes symptomatic, over 60% of dopaminergic neurons would have been lost, hence, the importance of implementing disease modifying therapy strategies very early in the disease.

The treatment of PD remains symptomatic and suboptimal as a result of treatment complications. Developing disease-modifying therapies that would slow disease progression outside of isolated symptomatic relief is essential. For this effect, reliable biomarkers allowing early disease detection are fundamental. The ideal biomarker in PD would aim at early detection of disease at prodromal stage, confirm differential diagnosis with other diseases such as PSP or MSA, and allow monitoring of disease progression (Shapira 2013).

There have been important advances in biomarkers in PD in recent years. The Parkinson's Progression Markers Initiative (PPMI), a multicenter, observational study, is currently underway, and its main objective is identifying imaging, biological, clinical, and behavioral assessments biomarkers of disease progression in a standardized way (Parkinson Progression Marker Initiative 2011).

Multiple imaging techniques have found to be relevant biomarkers in PD, of which the most prominent is dopamine transporter (DAT) imaging. It allows the identification of significant reduction of dopaminergic neurons. The usefulness of this technique for disease progression remains however to be further confirmed (Wang et al. 2013). Alpha-synuclein, a protein found in Lewy bodies, the pathological hallmark of PD, has emerged as a very relevant biomarker. Levels of total alpha-synuclein have been found to be decreased in CSF, while oligomer forms of this protein have been reported to be increased. Recent findings seem to indicate that measuring the levels of alpha-synuclein might not only be diagnostic of PD but serve as a marker for disease progression (Wang et al. 2013). The assessment of peripheral alpha-synculein blood and Saliva levels could be utilized as biomarkers, avoiding invasive procedures (Devic et al. 2011). Significant advances have been made in the genetic understanding of familial forms of PD (mutations in LRKK2 and PINK genes) that might accelerate the development of effective and targeted therapies (Li and Yang 2011).

Biomarkers in Multiple Sclerosis

The discovery of MRI technology has contributed significantly to the accelerated drug development in multiple sclerosis (MS). Specific sequences such as T1 with and without gadolinium, T2, and FLAIR allow the detection of the severity of inflammation, demyelination, and axonal loss. Following the changes of these parameters provides information demonstrating drug effect on disease activity. Furthermore, whole brain and spinal volumetric measurement identifies the severity of atrophy in specific regions of the central nervous system and hence can be used as a surrogate marker of disease progression (Nair et al. 2013).

CSF, although an invasive procedure, is valuable for marker collection of inflammation (Oligoclonal bands), demyelination (Myelin Basic Protein, MBP), and axonal loss (Neurofilaments) (Giovannoni 2006). The levels of JCV in CSF are a safety marker to identify subjects at potential risk of developing

progressive multifocal leucoencephalopathy (PML) while on potent immunomodulators/immunosupressors.

Optic coherence tomography (OCT) measures retinal nerve fiber layer (RNFL) thickness and macular volume. This technique can be used as a marker of neuroprotective effect of certain drugs (Villoslada 2010).

The Expanded Disability Status Score (EDSS), the Multiple Sclerosis Functional Composite (MSFC) that includes the Paced Auditory Serial Addition Test (PASAT), and the Symbol Digit Modalities Test (SDMT) are recognized clinical biomarkers of both physical and cognitive disability in MS.

Biomarkers in Neuro-Orphan Indications

Biomarkers in Amyotrophic Lateral Sclerosis (ALS)

Multiple advances have been made in the past years in identifying biomarkers for amyotrophic lateral sclerosis (ALS). Significant progress has been observed in the genomic, neuroimaging (MRI, Diffusion Tensor Imaging, PET), and electrophysiologic marker field. Motor unit number estimation (MUNE), electrical impedance myography, and the neurophysiological index are the most prominent electrophysiological techniques that can be useful in disease progression monitoring (Bowser et al. 2011).

Recent findings have revealed increased neurofilament (NFL) levels in the CSF of ALS patients compared to controls as a marker of neurodegeneration, increased blood levels of creatinine, as well as NFL, and NOGO-A strong expression in muscles (Chen and Shang 2015).

Further validation of the different biomarkers and their utility, in isolation or in combination, in disease diagnosis and progression, is needed.

Biomarkers in Huntington's Disease

Huntington's disease is an autosomal dominant neurodegenerative condition caused by a CAG trinucleotide expansion in the Huntingtin gene (HTT) (The Huntington's Disease Collaborative Research Group 1993). A new wave of investigational products is currently under investigation in this disease arena, which is aimed at core disease mechanisms (Chen and Shang 2015). The development of new therapies depends on a robust and in depth understanding of the pathogenesis of this disease and the ability to translate this knowledge into pharmacodynamics biomarkers. Currently there is much interest in the quantification of mutant Huntingtin (mHTT) in the CSF and PET imaging targeting Huntingtin (HTT) as a potential biomarker (Mestra and Sampaio 2017). Although the actual goal of DMT is to lower the levels of the HTT protein, the molecular targets of the intervention are different products of the HTT gene – mostly RNAs or the gene itself. Ongoing trials in this arena include the assessment of various antisense oligonucleotides and RNA interference strategy agents (IQVIA proprietary database).

Summary

Significant progress has been made in the field of biomarkers for neurological disorders. These span from imaging, genetics, molecular, biochemical, clinical to electrophysiological methods. The use of biomarkers in isolation or combination, for diagnostic and predictive purposes in the clinic, while allowing the monitoring of drug effect in a clinical trial setting, has evolved rapidly in recent years. Greater investment and research in this field is still needed to allow for accelerated discovery of therapeutic interventions, in an area with important unmet medical needs.

References and Further Reading

Alavijeh MS, Chishty M, Qaiser MZ et al (2005) Drug metabolism and pharmacokinetics, the blood-brain barrier, and central nervous system drug discovery. NeuroRx 2:554–571

Bateman RJ, Benzinger TL, Berry S, DIAN-TU Pharma Consortium for the Dominantly Inherited Alzheimer Network et al (2017) The DIAN-TU next generation Alzheimer's prevention trial: adaptive design and disease progression model. Alzheimers Dement 13:8–19

Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69:89–95

Bowser R, Cudkowicz M, Kaddurah-Daouk R (2006) Biomarkers for amyotrophic lateral sclerosis. Expert Rev Mol Diagn 6:387–398

Bowser R, Turner M, Shefner J (2011) Biomarkers in amyotrophic lateral sclerosis: opportunities and limitations. Nat Rev Neurol 7:631–638

Chen X, Shang F (2015) New developments and future opportunities in biomarkers for amyotrophic lateral sclerosis. Transl Neurodegener 4:17

Calcagno A, Di Perri G, Bonora S (2014) Pharmacokinetics and pharmacodynamics of antiretrovirals in the central nervous system. Clin Pharmacokinet 53:891–906

ClinicalTrials.gov. https://clinicaltrials.gov/. Accessed 19 Sept 2017

Danhof M, de Jongh J, De Lange EC et al (2007) Mechanism-based pharmacokinetic-pharmacodynamic modeling: biophase distribution, receptor theory, and dynamical systems analysis. Annu Rev Pharmacol Toxicol 47:357–400

de Lange E (2013) The mastermind approach to CNS drug therapy: translational prediction of human brain distribution, target site kinetics, and therapeutic effects. Fluids Barriers CNS 10:12

de Leon MJ, Convit A, Wolf OT et al (2001) Prediction of cognitive decline in normal elderly subjects with 2-[(18)F]fluoro-2-deoxy-D-glucose/positron-emission tomography (FDG/PET). Proc Natl Acad Sci USA 98:10966–10971

Devic I, Hwang H, Edgar JS et al (2011) Salivary alpha-synuclein and DJ-1: potential biomarkers for Parkinson's disease. Brain 134:e178

Di L, Kerns EH (eds) (2015) Blood-brain barrier in drug discovery: optimizing brain exposure of CNS drugs and minimizing brain side effect for peripheral drugs. Wiley, Malden

Dingemanse J, Danhof M, Breimer DD (1998) Pharmacokinetic-pharmacodynamic modeling of CNS drug effects: an overview. Pharmacol Ther 38:1–52

Donovan MD, Boylan GB, Murray DM et al (2015) Treating disorders of the neonatal central nervous system: pharmacokinetic and pharmacodynamic considerations with a focus on antiepileptics. Br J Clin Pharmacol 8:62–77

Doody RS, Raman R, Farlow M, Alzheimer's Disease Cooperative Study Steering Committee, Siemers E, Sethuraman G, Mohs R, Semagacestat Study Group (2013) A phase 3 trial of semagacestat for treatment of Alzheimer's disease. N Engl J Med 369:341–350

Durham TB, Blanco MJ (2015) Target engagement in lead generation. Bioorg Med Chem Lett 25:998–1008

FDA (2014) Guidance for industry and FDA staff qualification process for drug development tools. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research. Silver Spring, MD, USA

Frank R, Hargreaves R (2003) Clinical biomarkers in drug discovery and development. Nat Rev Drug Discov 2:566–580

Giacomelli C, Daniele S, Martini C (2017) Potential biomarkers and novel pharmacological targets in protein aggregation-related neurodegenerative diseases. Biochem Pharmacol 131:1–15

Giovannoni G (2006) Multiple sclerosis cerebrospinal fluid biomarkers. Dis Markers 22:187–196

Hebert L, Weuve J, Scherr P et al (2013) Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology 80:1778–1783

Henley S, Bates G, Tabrizi S (2005) Biomarkers for neurodegenerative diseases. Curr Opin Neurol 18:698–705

Hughes L, Hayduk R, Vanbelle C (2014) Integrating biomarkers in Alzheimer's disease trials: review of compliance with biomarker assessments in Alzheimer's disease clinical trials. Available from the author upon request

ICH (2017). Pediatric guideline. Available at: http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/addendumclinical-investigation-of-medicinal-products-in-the-pediatric-population.html (accessed 30 November 2017)

Ittner LM, Götz J (2011) Amyloid-β and tau – a toxic *pas de deux* in Alzheimer's disease. Nat Rev Neurosci 12:67–72

Jobert M, Wilson F, Ruigt G, The IPEG Pharmaco-EEG Guidelines Committee et al (2012) Guidelines for the recording and evaluation of pharmaco-EEG data in man: the international Pharmaco-EEG society (IPEG). Neuropsychobiology 66:201–220

Johnson KA, Minoshima S, Bohnen NI et al (2013) Update on appropriate use criteria for amyloid PET imaging: dementia experts, mild cognitive impairment, and education. J Nucl Med 54:1011–1013

Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–716

Lambracht-Washington D, Rosenberg RN (2013) Anti-amyloid beta to tau-based immunization: developments in immunotherapy for Alzheimer's disease. Immunotargets Ther 2:105–114

Li T, Yang D, Sushchky S et al (2011) Models for LRRK2-linked Parkinsonism. Parkinsons Dis 2011:942412

Maharaj AR, Barrett JS, Edginton AN (2013) A workflow example of PBPK modeling to support pediatric research and development: case study with lorazepam. AAPS J 15:455–464

Melham M (2013) Translation of central nervous system occupancy from animal models: application of pharmacokinetic/pharmacodynamic modeling. J Pharmacol Exp Ther 347:2–6

Mestra TA, Sampaio C (2017) Huntington's disease: linking pathogenesis to the development of experimental therapeutics. Curr Neurol Neurosci Rep 17:18

Nair G, Shea C, Crainiceanu C et al (2013) Quantification of multiple-sclerosis-related brain atrophy in two

heterogeneous MRI datasets using mixed-effects modeling. Neuroimage Clin 3:171–179
Nielsen JC, Tolbert D, Patel M et al (2014) Vigabatrin pediatric dosing information for refractory complex partial seizures: results from a population dose–response analysis. Epilepsia 55:e134–e138
Ousset P-J, Cummings J, Delrieu J et al (2001) Is Alzheimer's disease drug development broken? What must be improved. JPAD 1(1), 2014 Proc Natl Acad Sci USA, pp 1–7
Pangalos M, Schechter L, Hurko O (2007) Drug development for CNS disorders: strategies for balancing risk and reducing attrition. Nat Rev Drug Discov 6:521–532
Parkinson Progression Marker Initiative (2011) The Parkinson progression marker initiative (PPMI). Prog Neurobiol 95:629–635
Pellock JM, Carman WH, Thyagarajan V et al (2012) Efficacy of antiepileptic drugs in adults predicts efficacy in children: a systematic review. Neurology 79:1482–1489
Pellock JM, Arzimanoglou A, D'Cruz, Pediatric Epilepsy Consortium for Extrapolation et al (2017) Extrapolating evidence of antiepileptic drug efficacy in adults to children $\geq$2 years of age with focal seizures: the case for disease similarity. Epilepsia 58:1686
Politis M, Piccini P (2012) Positron emission tomography imaging in neurological disorders. J Neurol 259:1769–1780
Quintiles proprietary database
Reichel A (2009) Addressing central nervous system (CNS) penetration in drug discovery: basics and implications of the evolving new concept. Chem Biodivers 6:2030–2049
Reichel A (2015) Pharmacokinetics of CNS penetration. In: Di L, Kerns EH (eds) Blood-brain barrier in drug discovery: optimizing brain exposure of CNS drugs and minimizing brain side effect for peripheral drugs. Wiley, Malden, pp 7–36
Riascher SL, Saykin AJ, West JP et al (2009) Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Curr Alzheimer Res 6:347–361
Rizk ML, Zou L, Savic RM et al (2017) Importance of drug pharmacokinetics at the site of action. Clin Transl Sci 10:133–142
Rutkove S, Caress J, Cartwright M et al (2012) Electrical impedance myography as a biomarker to assess ALS progression. Amyotroph Lateral Scler 13:439–445
Salloway S, Sperling R, Fox NC, Bapineuzumab 301 and 302 Clinical Trial Investigators et al (2014) Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med 370:322–333
Seifert KD, Wiener JI (2013) The impact of DaTscan on the diagnosis and management of movement disorders: a retrospective study. Am J Neurodegener Dis 2:29–34
Sethi S, Brietzke E (2016) Omics-based biomarkers: application of metabolomics in neuropsychiatric disorders. Int J Neuropsychopharmacol 19:pyv096
Shapira A (2013) Recent developments in biomarkers in Parkinson disease. Curr Opin Neurol 26:395–400
Silva G, Furie K (2009) Biomarkers in neurology. In: Woodbury-Harris KM, Coull BM (eds) Clinical trials in the neurosciences. Karger, Basel, pp 55–61
The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell 72:971–983
Tuntland H, Espehaug B, Forland O et al (2014) Reablement in community-dwelling adults: study protocol for a randomised controlled trial. BMC Geriatr 14:139
Tysnes O, Storstein A (2017) The epidemiology of Parkinson's disease. J Neural Transm (Vienna) 124:901–905
Villoslada P (2010) Focus on biomarkers: biomarkers for multiple sclerosis. Drug News Perspect 23:585–595
Wang J, Hoekstra J, Zuo C et al (2013) Biomarkers of Parkinson's disease: current status and future. Drug Discov Today 18:155–162
Wiltshire HR, Kilpatrick GJ, Tilbrook GS (2012) A placebo and midazolam-controlled phase I single ascending dose study evaluating the safety, pharmacokinetics, and pharmacodynamics of remimazolam. Anesth Analg 115:284–296
Yamamoto S, Karashima M, Arai Y et al (2017) Prediction of human pharmacokinetic profile after transdermal drug application using excised human skin. J Pharm Sci 106:2787–2794

Pharmacodynamic Evaluation: Pain Methodologies

5

Pieter Siebenga, Pieter Okkerse, Guido van Amerongen, Robert Jan Doll, Alex Mentink, Justin Hay, and Geert Jan Groeneveld

Contents

Abstract

Despite many advances in the last decades in understanding pain, the development of new analgesic compounds has not followed the same pace. The development of more targeted analgesic compounds with fewer side effects is therefore essential. With an increased demand to demonstrate pharmacodynamic effects of new analgesic compounds, the importance of human evoked pain models is now higher than ever.

Pharmacodynamic evaluation with human evoked pain models offers the possibility to determine the dose ranges at which new analgesics exert their pharmacological effect. Pain models may also aid in the choice of target population, determine which modality of pain a new drug is expected to be most suitable, help to differentiate between a central or more peripheral mode of action of new drugs, and help determine which other effects contribute to its mode of action, e.g., sedation.

Human evoked pain models are conducted in standardized laboratories where factors like

P. Siebenga · P. Okkerse · G. van Amerongen · R. J. Doll · A. Mentink · J. Hay · G. J. Groeneveld (✉)
Centre for Human Drug Research, Leiden, The Netherlands
e-mail: psiebenga@chdr.nl; pokkerse@chdr.nl; gvamerongen@chdr.nl; rjdoll@chdr.nl; amentink@chdr.nl; jhay@chdr.nl; ggroeneveld@chdr.nl

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_56

stimulus intensity, frequency, duration, and location can be controlled. Using pain models in healthy volunteers has important advantages over assessing the effects of new drugs in patients with pain; the pain elicited in human pain models is predictable in its intensity while clinical pain will naturally fluctuate. Analgesic properties can be investigated with pain models without the influence of accompanying symptoms that are often seen in patients with pain.

General Introduction

Pain is intended as a warning to the body that a noxious stimulus can (potentially) harm the body. The International Association for the Study of Pain (IASP) defines pain as an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage (Bonica 1979). Prevalence studies show that in Western countries 19–31% of the adult population suffers from a form of chronic pain (Macfarlane et al. 2013; Moore et al. 2015). Despite the availability of potent analgesics such as opioids, chronic pain remains a high unmet medical need as many effective analgesics have important side effects and chronic treatment with opioids leads to tolerance and addiction. The development of better and more specific analgesic compounds therefore remains essential. With an increased demand to demonstrate pharmacodynamic effects of new compounds as early as possible in clinical drug studies, the importance of human evoked pain models is now more than ever.

In a pure neurophysiological sense, nociceptive pain occurs when nociceptors are stimulated by noxious stimuli (e.g., mechanical, thermal, electrical, or chemical stimuli). After a threshold has been reached, the nociceptive nerve fiber transmits the pain signal to the spinal cord. The signal is modulated at several locations along ascending pathways through the dorsal horn and spinal cord. From the spinal cord, the pain signal is projected to supraspinal centers where the brain can modulate the excitatory activity via descending control (Olesen et al. 2012). Perception of pain is even more complex where more than one sensory system is responsible for transmission of the painful stimulus (Aguggia 2003).

From a more neuropsychosocial point of view, pain is a complex experience influenced by many factors such as emotion, fear, anxiety, but also cultural background, sex, genetics, and educational background. Due to its complexity, it can be difficult to assess the effects of analgesic drugs on pain in patients, and animal pain models demonstrate low predictability for clinical efficacy in humans. Several explanations are receptor dissimilarity between species, differences in pharmacokinetics, and morphological and functional differences between the brains of animals and humans (Olesen et al. 2012). Human evoked pain models can control some of these influencing factors. Therefore, these models are an important step in the translation of animal research to pain patients.

Pharmacodynamic evaluation through human evoked pain models offers the possibility to differentiate between a centrally or peripherally acting drug, for which modality of pain a new drug will be most suitable (nociceptive, neuropathic, or inflammatory), and which other effects contribute to its mode of action (e.g., sedation, tolerance) (Oertel and Lötsch 2013; Okkerse et al. 2017; Olesen et al. 2012; Staahl et al. 2009a). This can be done in early clinical trials with healthy volunteers, which is not only cost-reducing but also time saving. Other advantages of using human evoked pain models are (1) stimulus intensity, duration, and modality are controlled and do not vary over time; (2) differentiated responses to different standardized stimulus modalities; (3) the response can be assessed quantitatively and compared over time; (4) pain sensitivity can be compared quantitatively between various normal/affected/treated regions; (5) models of pathological conditions can be studied and the effects of drugs on such mechanisms quantified; and (6) pain models can be easily performed in healthy subjects, who are easier to recruit into clinical studies (Arendt-Nielsen et al. 2007a).

Evoked pain is mostly short-lasting, with most stimuli being applied exogenously and are

generally focused on cutaneous nociceptor activation. Arguably, evoked pain models are limited in intensity due to ethical constraints related to the risk of tissue damage. In contrast to natural occurring pain which is mostly caused by endogenous factors, longer lasting and influenced by complex emotions (Moore et al. 2013). Since clinical pain is a complex sensation involving psychological, physiological, and cognitive factors, no single pain model is able to replicate all aspects of clinical pain (Okkerse et al. 2017).

Several methods exist for evoking pain in humans, such as mechanical, thermal, electrical, and chemical stimulation. A stimulus can be either phasic or tonic. Stimuli can be applied to different tissue types for instance skin, muscles, or viscera (Arendt-Nielsen et al. 2007a). This chapter focuses on the different pain models that are used to induce evoked pain in humans and the means to assess the evoked pain. Human evoked pain models are divided into the energy domain (i.e., mechanical, thermal, electrical, and chemical induction) and are further subdivided into area of stimulation (i.e., skin, muscle, and viscera) that is stimulated.

Pain Assessment Techniques

An evoked pain model consists of two elements; a stimulus needs to be applied to evoke pain and the related pain response needs to be measured (Gracely 2013). Possible assessment techniques for evoked pain responses can be divided into several categories: psychophysical, electrophysiological, and imaging.

Psychophysical Methods

Psychophysical methods aim to describe the relationship between physical stimuli (section "Pain Stimulation Techniques") and corresponding subjective responses in a quantitative manner. These methods are widely used to study stimulus processing in various research fields, including pain processing. The methods to quantify the relationship between stimulus and response can be subdivided into unidimensional methods and multidimensional methods.

Unidimensional Measures

Purpose and Rationale

Unidimensional questionnaires such as the Numeric Rating Scale (NRS), the Verbal Rating Scale (VRS), or the Visual Analogue Scale (VAS) provide a single subjective measure to the intensity of a stimulus. More quantitative unidimensional measures are thresholds: the intensity where a stimulus is first perceived as painful (pain detection threshold), and the intensity where the stimulus is no longer tolerable (pain tolerance threshold).

Procedure and Evaluation

The NRS consists of a horizontal line with evenly spaced perpendicular lines. The patient or subject is asked to rate his or her pain from 0 to 10 (11 point scale) or from 0 to 100 (101 point scale) where 0 equals "no pain" and 10 or 100 equals "worst possible pain." The VRS consists of a list of descriptive terms to express the different levels of pain. It at least should include the extremes, in this case "no pain" and "worst possible pain." Additional terms usually used are "very mild pain," "mild pain," "moderate pain," "severe pain," and "very severe pain." The VAS is a horizontal line consisting of the two extremes and subjects are asked to indicate on the horizontal line indicating the perceived intensity.

For the determination of pain thresholds, the stimulus intensity increases until a certain pain threshold is reached (Gracely 2013). The pain detection threshold (PDT) is reached when a change in sensation from nonpainful to painful is felt by the subject. The pain tolerance threshold (PTT) is the stimulus intensity at which the pain is no longer tolerable. Depending on the method used, the stimulus is ceased before or when reaching the pain tolerance threshold.

Several stimulus modalities (such as thermal, mechanical, or electrical) can be used to quantify pain perception and to assess sensory function (Arendt-Nielsen and Yarnitsky 2009). These can then also be used as an indicator of the current

state of the pain system. For example, it can be used to detect hyperalgesia, which is a condition in which an enhanced pain response to noxious stimuli is observed. Hyperalgesia is indicated by a decrease in the pain thresholds and an increase in pain to supra-threshold stimuli.

Critical Assessment of the Method

The main disadvantage of NRS, VRS, and VAS scales is that they measure a single qualitative aspect of pain, namely, intensity or unpleasantness, while pain consists of more qualities (section "Multidimensional Measures"). Moreover, even though these (subjective) unidimensional scales can be used as a coarse measure of the level of pain a subject experiences, they cannot be used to distinguish individual contributions of nociceptive and pain-related mechanisms.

Modification of the Method

To obtain a more detailed description of pain perception and qualities, multidimensional questionnaires can be used (section "Multidimensional Measures").

More advanced psychophysical procedures are being developed to relate stimulus properties (e.g., pulse width, number of pulses, and inter-pulse interval) and perceptions to nociceptive processes (Doll et al. 2016; Yang et al. 2015). Also, combining unidimensional measures with neurophysiological measures may provide more information on underlying processes.

Multidimensional Measures

Purpose and Rationale

Unidimensional questionnaires are often found to be limited in their capabilities in describing pain perceptions. Therefore, multidimensional measures provide means to describe perception using multiple sensory and affective qualities of pain.

Procedure

In contrast to the unidimensional measures, subjects are asked to fill in several answers to various questions, or have to choose from a large range of adjectives to describe their pain perception. One of the most widely used multidimensional tool is the McGill Pain Questionnaire (Melzack 1975). This questionnaire not only measures the pain intensity but also measures the sensory and affective qualities of pain. The McGill Pain Questionnaire has been used in a large number of studies and has been translated and validated in multiple languages (Melzack 2005; Melzack and Katz 2013).

Critical Assessment of the Method

Multidimensional questionnaires often take more time to complete than simpler unidimensional questionnaires. Particularly in a clinical setting, multidimensional questionnaires either need to be compressed (such as the short-form McGill Pain Questionnaire) or replaced by unidimensional questionnaires. Additionally, in the experimental setting of evoked pain models not all components of a questionnaire may be applicable, e.g., items related to affective aspects of pain sensation.

Modification of the Method

Numerous pain questionnaires exist that measure different qualities of neuropathic and non-neuropathic pain. These include the Pain Quality Assessment Scale (PQAS), Leeds Assessment of Neuropathic Symptoms and Signs (LANSS), and PainDETECT. Moreover, questionnaires targeting specific patients groups exist as well; the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and Knee Society Score (KSS) assign pain, stiffness, and functional scores to patients suffering from osteoarthritis.

Electrophysiological and Imaging Methods

Electrophysiological readouts include evoked potentials via electroencephalography (EEG). Imaging readouts include functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Electrophysiological and imaging readouts provide a more objective measurement of pain. However, they have a larger variation in outcome measurements, are more expensive, and are technically more difficult to

perform in a large group of subjects (Arendt-Nielsen et al. 2007a; Kakigi et al. 2005; Wager et al. 2013).

Functional Magnetic Resonance Imaging

Purpose and Rationale

Neuroimaging has identified several cortical regions in the brain that are typically active when a painful stimulus is applied to the body. With fMRI these pain responses in the brain can be observed. Structures that are active during pain perception are the primary and secondary somatosensory, the cingulate, and the insular cortices which together are called the pain matrix (Ingvar 1999; Peyron et al. 2000; Porro 2003; Rainville 2002; Tracey and Mantyh 2007). Activation of the pain matrix due to nociceptive stimuli is the functional imaging analogue of conscious pain perception. Measuring the activities in the pain matrix during a painful stimulus can be used as an objective measure for pain perception (Borsook et al. 2010).

Procedure and Evaluation

A subject is given a painful stimulus while being scanned in the MRI. fMRI measures brain activity by detecting changes in blood flow (hemodynamic response) associated with neuronal activation (Huettel et al. 2014). It uses the relative abundance of deoxyhemoglobin in blood that changes the proton signal from water molecules surrounding a blood vessel, producing blood oxygenation level-dependent contrasts (BOLD) (Ogawa et al. 1990). Via the BOLD signal an indirect index of neural activity is provided. Several fMRI methods are used in pain research, which reveal the neural correlation of pain perception and modulation by characterizing the brain response to evoked stimuli (e.g., pain, allodynia), task-driven responses, or drugs (phMRI) (Borsook et al. 2010).

Critical Assessment of the Method

This noninvasive method of measuring pain can provide a measure of cerebral perfusion that correlates with an acute painful stimulus in healthy volunteers. Detecting chronic pain is more complex due to confounding factors like disease and treatment. Brain systems like emotion, memory, and motivation are also active during measurement in these patients (Borsook et al. 2010). These systems can also be triggered by placebo analgesia, which adds to the complexity (Morton et al. 2016). Additionally, the resolution of fMRI is inferior compared to the EEG, which means that it is not suitable to investigate the primary neuronal activity directly related to the pain stimulus and less suitable to investigate the deeper structures of the brain (e.g., brainstem and thalamus) (Olesen et al. 2012).

Modifications of the Method

When including fMRI in a study, a distinction can be made in the type of pain model that will be used during the study. Mechanical (Baron et al. 1999), thermal (Lapotka et al. 2017; Shukla et al. 2011), electrical (Kocyigit et al. 2012), and chemical (Baron et al. 1999) induced pain models can be used during an fMRI scan, with the exception of models that entail the use of water and metal. Modifications can be made in the type of MRI, protocol used for scanning, and analyzing protocols.

Electroencephalography and Evoked Potentials

Purpose and Rationale

EEG is a noninvasive technique which records (spontaneous) synchronized postsynaptic neuronal activity of the human cortex. In contrast to brain imaging techniques, EEG has a high temporal resolution. This high resolution makes EEG an effective method of observing (rapid) changes in brain activity. Additionally, EEG is a suitable method for recording evoked potentials (EP) to painful stimuli (e.g., thermal, mechanical, or electrical) and may provide important information on (central) pain processing (Mouraux and Iannetti 2008).

Procedure and Evaluation

To record EEG, several electrodes are placed on the scalp, either using individual Ag/AgCl electrodes or specialized caps. The impedance should

be similar among all electrodes and is advised to be held under 50 kΩ. Quantifying resting state EEG can be done by means of spectral analysis. Due to its nonstationary behavior, recorded data is divided in short epochs ranging between 2 s and 10 s (Jobert et al. 2013). Each epoch is then transformed into the frequency domain and after removing or correcting epochs affected by artifacts (e.g., ocular or muscular activity) averaged. The frequency range is then subdivided into bands (i.e., delta, theta, alpha, beta, and gamma) and then integrated over frequency bands to obtain the total power per frequency band.

EPs are monophasic deflections of spontaneous EEG and are time and phase locked on the onset of the stimuli (Mouraux and Iannetti 2008). These waveforms are typically characterized by their polarity, latency, amplitude, and position on the scalp. As the signal-to-noise ratio (SNR) is relatively low when recording EPs, repeatedly stimulating and recording cortical responses is required. This allows improving the SNR by means of averaging and allows the characterization of the evoked response in terms of the amplitudes and latencies. The data collected generally includes the peaks and latencies of the N1, N2, P2, and P3, or the top-top amplitude between N2 and P2 (Treede et al. 2003). Moreover, the reaction time after the presentation can also be recorded. The N1 is most prominent at the contralateral temporal side (i.e., T3 or T4) referenced to the frontal Fz. The N2 and P2 are most prominently visible at the vertex Cz referenced to the (linked) earlobes A1 and A2. The P3, which is not thought to be nociceptive specific, has a relatively long latency and is best observed at Pz referenced to the earlobes.

Critical Assessment of the Method

It is important to know that EEG recordings will be contaminated with artifacts. Common sources of artifacts are ocular movements and blinks, muscle contraction, and cardiac activity. Each of these artifacts have their own characteristics in both time domain and frequency domain and must be dealt with prior to analyzing EEG. Possibilities for dealing with these artifacts are either excluding parts of the recording for analysis or correcting for them (e.g., by using regression techniques, filtering, or blind source separation techniques).

EPs have been shown to be sensitive to various changes in pain pathways. EPs are sensitive to changes induced by analgesics (Schaffler et al. 2017). It must be taken into account, however, that the amplitudes of the EPs are dependent on the attention of the subject; reduced attention results in significantly lower peak amplitudes. Moreover, the across trial variability is relatively high making comparisons between groups and trials complicating.

Modification of the Method

When several electrodes are recorded during a resting state EEG, multichannel topography allows observing the activity recorded at several locations. Time-dependent changes in power spectra can therefore also be visualized. Even though EEG has a relatively low spatial distribution in contrast to techniques such as fMRI, source localization techniques are used to find brain sources of the recorded potentials (Grech et al. 2008). However, high-density electrode placement is required for more reliable source localization (Song et al. 2015).

As a result of averaging EPs, non-phase-locked information is lost. Time-frequency analysis of epochs does provide the means to study non-phase-locked information (Hu et al. 2015; Mouraux and Iannetti 2008).

Nociceptive Spinal Flexion Reflex

Purpose and Rationale

The nociceptive spinal flexion reflex (NFR), also called the RIII reflex, is a physiological, polysynaptic reflex allowing for painful stimuli to activate an appropriate withdrawal response (Skljarevski and Ramadan 2002). It is one of the available tools for objective quantification of spinal nociception in humans. The NFR reflex can be elicited in all four limbs. Here, the most standard procedure, with stimulation of the lower limb (sural nerve), is described.

Procedure and Evaluation

Electrical constant current stimulation is delivered to the retromalleolar pathway of the sural nerve. Each stimulus consists of five pulses of 1 ms duration, separated by 4 ms, resulting in a total duration of 21 ms. Electromyographic responses are recorded from the ipsilateral biceps femoris (short head) via surface electrodes placed 4–5 cm apart over the muscle belly. The RIII reflex is identified as a polyphasic muscle response appearing with an onset latency between 90 ms and 130 ms after stimulation (Willer 1977). Following stimulation of the sural nerve, three responses are sequentially recorded: the tactile reflex (also known as RII), the nociceptive flexion reflex (RIII), and an involuntary movement signal (Skljarevski and Ramadan 2002). For quantification of the RIII reflex response, the reflex area is obtained by integrating the rectified signal within a 50 ms analysis window starting between 90 ms and 120 ms after stimulation. Stimulus–response curves are recorded by increasing stimulation intensity in 0.5 mA steps starting from 0.5 mA. Participants can rate the pain intensity of each stimulus using an NRS or VAS. The pain threshold is determined as the stimulus intensity that first evokes a painful sensation (defined as an NRS rating ≥ 1 or VAS > 0). The RIII threshold is defined as the stimulus intensity that first evokes a reflex response exceeding a raw area of 100 μV × ms (Ruscheweyh et al. 2015).

Critical Assessment of the Method

Reduction of the NFR by a pharmaceutical compound does not necessarily imply reduction of ascending nociception, but may also indicate modulation of other components that play a role in the RIII reflex, such as deep dorsal horn interneurons (Schouenborg et al. 1995) or motor neurons (Ruscheweyh et al. 2015).

Modifications of the Method

NFR is affected by demographic factors. It is indicated that female subjects and children have lower NFR thresholds (Page and France 1997; Sandrini et al. 1989). Other factors affecting the outcome include cardiac cycle, baroreceptors, stimulation site, and even diurnal rhythm (Skljarevski and Ramadan 2002). All should be addressed before the start of the trial.

Conditioned Pain Modulation

Purpose and Rationale

Conditioned pain modulation (CPM) (also known as descending noxious inhibitory control (DNIC) for animal studies or heterotopic noxious counter-stimulation) is a paradigm that uses a conditioning stimulus to influence a test stimulus and can be used to assess the endogenous analgesic capacity of both healthy subjects and patients (Nir and Yarnitsky 2015). The assessment of CPM provides an indication of the balance between descending facilitation and inhibition. It is hypothesized that malfunction of pain modulation network may be the cause, rather than the effect, of chronic pain development (Yarnitsky et al. 2010).

The endogenous network is mediated via descending serotonergic, noradrenergic, and dopaminergic pathways, with the conditioning stimulus activating decreasing the activity of on-cells in the rostral ventromedial medulla (RVM) (Hernández et al. 1994).

Procedure and Evaluation

CPM is assessed by using a conditioning stimulus which is generally a tonic nociceptive stimulus; however, nonpainful conditioning stimuli have also been reported (Bouhassira et al. 1998; Lautenbacher et al. 2002). Various methods can be used as the conditioning stimulus. The method most commonly used is the cold pressor test (Pud et al. 2009). However, other stimuli have also been reported including hot water, ischemic pain, heat, chemically induced pain, electrical induced pain, and physically induced muscle pain (Popescu et al. 2010).

Various methods are used as the test stimulus including nociceptive flexion reflex, electrical, heat, and pressure stimulation, among others. CPM is considered to be a systemic experience, and as such, heterotopic stimulation is used for the assessment of CPM. CPM is assessed by

comparing the endpoint from the test stimulus before and after administration of the conditioning stimulus. The duration of effect from the conditioned stimuli may be assessed in parallel or soon after administration of the conditioning stimulation; however, the duration of effect is paradigm-dependent with reports of CPM effects lasting 5 min, 30 min, and up to 60 min after application of the conditioning stimulus (Fujii et al. 2006; Graven-Nielsen et al. 1998; Tuveson et al. 2006).

Critical Assessment of the Method

CPM paradigms are reported to have both reasonable to high intra- and inter-individual variability especially due to the endpoint being derived from subjective pain reports (Nir and Yarnitsky 2015). Furthermore, divergent terminology used in the literature for the same phenomena complicate comparisons. Factors affecting CPM are use of medication, psychological state of subjects (anxiety, depression, emotional status, and attention span), and even ethnic origin (Goubert et al. 2015).

Modifications of the Method

As highlighted above and as with other pain models, there are numerous ways to assess CPM including different methods for the conditioning and test stimulus including using different application area. Furthermore, the endpoints vary considerably between research groups, with some groups assessing change in the test stimuli endpoint before and after the conditioning stimuli while others reporting changes in the endpoint during administration of the conditioning stimulus (Doll et al. 2014).

Pain Stimulation Techniques

Mechanical Stimulation

Human evoked pain models date back to the late nineteenth century, in which mechanical pressure was used to induce pain (Hardy et al. 1940). Over time, mechanical stimulation techniques became more accurate and are used to stimulate the skin, muscle, or viscera. The skin is the most used organ, because of its practical implementation. However, muscles can also be targeted both endogenous (post-exercise or ischemic) or exogenous (saline injection). Balloon distention in the viscera can be used to investigate new analgesic compounds in healthy volunteers but also as a diagnostic tool in patients (e.g., gut disorders). Mechanical stimulation can be divided into touch/pinprick, pressure, or pinching methods. This chapter focuses on the different techniques related to mechanical stimulation on the skin, muscles, and viscera using touch/pinprick, pressure, or pinching methods.

Mechanical Skin Stimulation

Touch and Pinprick

Purpose and Rationale

Mechanical stimulation via touch can be done with a cotton swab or a brushstroke (light mechanical stimulation) or pinprick. Light mechanical stimulation will not induce pain, but can be used to assess allodynia (lowered activation threshold for a nonpainful stimuli) by other pain stimuli.

A common method of applying pinprick stimulation is the use of von Frey filaments. These filaments are used to quantify touch as well as the PDT and the PTT. In animal and patient experiments, it is mostly used to determine functional recovery of Aδ- or Aβ-fibers. In human pain models, it is also used to determine hyperalgesia effects (e.g., in the capsaicin/UV-B model).

Procedure and Evaluation

Von Frey filaments are calibrated filaments, originally made of human or animal hair, and later with acrylic or synthetic fibers or optical glass fibers. The filaments bend at a certain designated force (Fruhstorfer et al. 2001). The applied force is dependent on the stiffness, which is in turn dependent on the diameter of each filament. During stimulation, the filament is placed perpendicular to the skin and pressed down with a constant increase in force until it bends for 1 s. Subsequently, the filament is removed with a constant

decrease in force and the subject is given some time to evaluate the stimulation.

Critical Assessment of the Method

The von Frey test is characterized by its simplicity, thereby showing advantages in clinical settings for rough quantification of functional nerve regeneration. In addition, it can be used to classify responders in the evaluation of hyperalgesia.

Although the von Frey method is still commonly used to evaluate peripheral nerve function in patients, it has several disadvantages as a human pain model. Pinprick stimulation always coactivates non-nociceptive Aβ-fibers. This may contaminate the signal and influence the results affecting the specificity. Additionally, intra- and inter-observer variability of this method is high. The skill of the observer will have an important impact on the results. Furthermore, each filament is produced for one constant standardized level of pressure, which means that many filaments are needed to determine sensory or pain thresholds with a high precision. Environmental changes may affect the calibration of the filaments. Lastly, it is difficult to combine von Frey stimulation with quantification methods that require time-locked responses, such as evoked potentials or time-frequency analyses in EEG.

Modifications of the Method

Von Frey filaments are inherently limited for method modification, because each filament is calibrated for one designated force application. Depending on the calibrated force, the filaments activate mainly Aβ-fibers (0.5–128 mN) or Aδ-fibers (128–512 mN) (Curatolo et al. 2000). Both conventional as electronic pinprick paradigms have been described (Möller et al. 1998). To eliminate the intra- and inter-observer variation one can consider electronic pinprick devices.

Impactometers/Pinch Interdigital Web/Joint

Purpose and Rationale

Pressure algometers are usually applied to the muscle or bone, but other sites including the interdigital web, skinflap, earlobe, or a finger or toe joint can be used as an area of investigation (Brennum et al. 1989; Curatolo et al. 1997; Staahl et al. 2006). Handheld or computer controlled pressure algometers are clinically similar to palpation (Olesen et al. 2012). For methods using a pinch methodology, the pain is due to a combination of mechanical stimulation and local ischemia, while when pressure is applied to muscle the pain is related to muscle strain (Olesen et al. 2012).

Procedure and Evaluation

A pressure algometer is applied to the area of interest. In a controlled manner, the pressure increases at a constant rate until a psychophysical endpoint of interest is attained. Many commercial handheld pressure algometers provide user feedback to ensure pressure is applied at a constant rate. For repeated applications, the pressure algometer should be applied to the same area to ensure intra-individual variability is minimized as variation in location can lead to different thresholds. PDT and PTT are the usual primary outcome measures.

Critical Assessment of the Method

For handheld pressure algometers, control of the rate of onset, muscle contraction, and examiner expectations are the primary limitations (Woolf and Max 2001). Differences in the size and shape of the probe limit comparisons between research groups.

Modifications of the Method

The model can be used in combination with other methods that induce hyperalgesia or sensitization. Stimulus-response curves can be compared with nonsensitized locations.

Rather than using a punctate pressure algometer, cuff algometry can be used with the advantage that the model is generally computer-controlled thereby controlling the rate of application of the pressure pain (Polianskis et al. 2001). Pain induced by cuff algometry is primarily related to muscle pain with minimal contribution from skin nociceptors.

Mechanical Muscle Stimulation

Post-exercise Muscle Soreness

Purpose and Rationale

Delayed onset muscle soreness (DOMS) is believed to be mediated by a combination of lactic acid, muscle spasm, connective tissue damage, muscle damage, inflammation, and endogenous substances (e.g., bradykinin and prostaglandins) (Nie et al. 2006). This pain model is thought to mimic clinical pain by inducing central sensitization while having no spontaneous pain at rest compared with exogenous induced pain models (Olesen et al. 2012). Central sensitization is an increase in general excitability of the dorsal horn neurons which can be caused by nerve injury (Laird and Bennett 1993). Secondary hyperalgesia is thought of as a transient state of central sensitization (Torebjörk et al. 1992).

Procedure and Evaluation

Subjects perform an eccentric exercise that they are unaccustomed to with insufficient rest periods. Peak allodynia/hyperalgesia occurs 24–48 h post-exercise. The affected muscles are assessed with a VAS to evaluate pain intensity and pressure algometry to evaluate allodynia/hyperalgesia.

Critical Assessment of the Method

The method is somewhat nonspecific, with an inflammatory component but these may be site specific (Staahl and Drewes 2004).

Modifications of the Method

The method can be modified by using various locations/muscle groups with development of hyperalgesia dependent on the size of the muscle (Svenson and Arendt-Nielsen 1995).

Ischemic Tourniquet

Purpose and Rationale

The tourniquet model is a tonic pain model with nociceptive contributions from the muscle, skin, and periosteum (the vascular connective tissue enveloping the bones). Clinically, tourniquets are used to perform intravenous regional anesthesia or to provide a bloodless operating field. The tourniquet leads to metabolic changes, primarily acidosis, and compression which leads to the release of prostaglandins. Neuropathic pain induced by nerve compression may also contribute to the pain felt (Kumar et al. 2016). The method can also be used as a conditioning stimuli for the CPM paradigm (section "Conditioned Pain Modulation").

The pain is thought to be mediated by the unmyelinated, slow conducting C-fibers that are usually inhibited by the Aδ-fibers (Kumar et al. 2016). The Aδ-fibers are blocked by mechanical compression after about 30 min, while the C-fibers continue to function (Casale et al. 1992). Tourniquet compression leads to release of prostaglandins by the injured cells (Kumar et al. 2016). These prostaglandins increase pain perception by sensitizing and exciting pain receptors. Also, limb ischemia causes central sensitization (Kumar et al. 2016).

Procedure and Evaluation

A pneumatic tourniquet is applied to an extremity, generally the thigh, following exsanguination of the leg. The cuff is inflated above the systolic blood pressure with ranges of 100–600 mmHg above the systolic pressure having been reported (Smith et al. 1966).

Critical Assessment of the Method

For healthy subjects, the tourniquet can be left for up to 2 h; however, pain is nonspecific with pain being felt under the tourniquet and/or in the lower limb. The method can also lead to temporary hypoesthesia and lower limb paralysis. Following reperfusion, the subject may experience hyperalgesia/allodynia and muscle cramps in the affected limb.

Modifications of the Method

The method can also be used with combination with voluntary muscle contractions or exsanguination of the leg by gravity or esmarch bandage (Hagenouw et al. 1986; Olesen et al. 2012). The width and type of tourniquet and the maximum inflation pressure can be used to modify the method.

Hypertonic Saline Injection

Purpose and Rationale

Intramuscular injection of hypertonic sodium chloride (4–6%) results in a tonic, deep, diffuse pain. The pain usually lasts for several minutes following infection and leads to both local and referred pain. The pain is primarily mediated via direct excitation of C-fibers than caused by saline-induced tissue injury (Schulte et al. 2006; Svendsen et al. 2005). The procedure can also be used to induce local, cutaneous hypoesthesia (Graven-Nielsen et al. 1997).

Procedure and Evaluation

A bolus saline solution is injected intramuscularly using computer-controlled infusion pump. Earlier models used manual bolus injections (Graven-Nielsen and Mense 2001). Pain intensity is measured by using a VAS with peak pain or area under the curve (AUC) being the primary outcome measures. Referred pain patterns and changes in the pressure pain thresholds of the local and referred pain areas can also be assessed (Ge et al. 2006).

Critical Assessment of the Method

A limitation of the model is that hypertonic injection may elicit excitation from both non-nociceptive and nociceptive nerve fibers (Korotkov et al. 2002).

Modifications of the Method

Potassium chloride is occasionally used instead of sodium chloride. Various sites can be used for injection with the most common being the musculus trapezius and the musculus tibialis anterior (Ge et al. 2006; Schulte et al. 2006).

Mechanical Visceral Stimulation

Barostat/Esophagal Distention/Bladder Distention

Purpose and Rationale

Induction of pain in viscera is difficult to perform due to the location of the organs associated with visceral pain as pain originates from the internal thoracic, pelvic, or abdominal organs (Johnson and Greenwood-van Meerveld 2016). Evoked pain models assessing mechanical visceral pain are generally limited to different accessible areas of the gastrointestinal (GI) tract, the urinary tract (Maggi 1990), and the uterine cervix (Drewes et al. 2003a).

Procedure and Evaluation

Mechanical stimulation of the viscera is generally performed using a balloon placed in the GI tract with the preferred locations being the esophagus or rectum. The most common method used is the barostat method whereby changes in air volume within a balloon while maintaining constant pressure are measured (Drewes et al. 2002). Generally, the balloon/bag is inflated until moderate pain is reported (up to approximately 7 on a VAS) and the corresponding balloon volume is reported (Staahl et al. 2006).

Critical Assessment of the Method

One of the main limitations of organ distention is the distortion of the balloon. However, calculation of strain by impedance planimetry or calculation of balloon radius may overcome these limitations (Drewes et al. 2003a; Staahl et al. 2006). Assessment of visceral pain is difficult due to the diffuse, referred, vague, and deep nature of the pain associated. Furthermore, autonomic reactions and the risk of perforation may limit the use of visceral pain models (Ness and Gebhart 1990).

Modifications of the Method

Perfusion of the GI tract with chemical substances (e.g., acid or capsaicin) can be used to sensitize the organs and nervous system and generally mimics the clinical situation where the organs are hypersensitive (Johnson and Greenwood-van Meerveld 2016).

Thermal Stimulation

Thermal stimuli can be used to induce pain. These can be roughly subdivided in cold and heat stimuli. Cold stimulation can be induced by emerging a body part in a cold water bath (cold pressor test), by using a cooling thermode,

by inducing a freezing lesion, or by applying ice directly to the skin. Heat pain tasks are among the most widely used pain models in human volunteers to investigate nociception, due to the relative ease of application and robustness. Nociceptive nerve fibers are activated by changes in temperature detected by a range of thermal receptors, of which Transient Receptor Potential ion channel subfamily vanilloid (TRPV) is most responsible. When a sharp increase in temperature (>43°) is detected, Aδ-fibers are activated, whereas C-fibers are activated by slower or more dull temperature changes. Different clinical pain states can lead to sensitization to heat pain stimuli and consequently lower activation thresholds, which can be replicated using hyperalgesia models.

Thermal Skin Stimulation

Cold Stimulation

Cold Pressor

Purpose and Rationale

The cold pressor test can be used to investigate nociception, but also cardiovascular responses (sympathetic functions) and can be used as to induce CPM. The cold sensation and pain induced by this method is mediated by activity of Aδ-fibers (cold sensation) and C-fibers (cold pain) (Olesen et al. 2012).

Procedure and Evaluation

One of the methods of induction of cold pressor pain is based on methods previously described by Eckhardt et al. and Jones et al. (Eckhardt et al. 1998; Jones et al. 1988). Here, subjects place their nondominant hand into a warm water bath for 2 min. At 1 min 45 s, a blood pressure cuff on the upper-arm is inflated to 20 mmHg below resting diastolic pressure. At 2 min, the subject moves their hand from the warm water bath, directly placing their hand into a cold water bath. The baths are two thermostat-controlled, circulating water baths set at 35.0 ± 0.5 °C and 1.0 ± 0.5 °C, respectively. The subject rates their pain intensity using a rating scale (e.g., VAS or NRS). When pain tolerance is reached, or when a time limit is reached, subjects are instructed to remove their arm from the water, at which point the blood pressure cuff is deflated. Typically, PTT expressed in seconds after immersion in the cold water is recorded as primary outcome measure.

Critical Assessment of the Method

Many different methodologies have been described which negatively influences the extent to which different studies can be compared. Small variations in water temperature can result in significant changes in pain intensity and tolerance times (Mitchell et al. 2004). Therefore, it is important to use water baths that are able to circulate that water to prevent warming of the water around the hand.

Modification of the Method

Different temperature settings of the water baths can be used and different body parts can be immersed in the baths. Instead of a cold water bath, also a cool gel substance has been reported to induce pain which makes it suitable for fMRI testing (Lapotka et al. 2017). Also a blood pressure cuff can be used to prevent compensatory blood flow to the hand.

Cooling Thermode

Procedure and Evaluation

The cold pressor test (section "Cold Pressor") is the most commonly used method to induce cold pain. However, a cooling thermode can also be used to induce cold pain. A contact thermode is attached to a part of the human body. The method consists of administering a temperature with an intensity that gradually decreases at a constant rate, usually 1 °C per second. The subject halts the stimulus when the cold pain threshold is reached or when a set lower cut off temperature is reached. Most studies using this methodology only report the cold pain detection threshold; the temperature at which the sensation has "just become painful."

Critical Assessment of the Method

For none of these compounds this methodology provided evidence for analgesic efficacy, while other pain induction methods were able to provide this evidence (Staahl et al. 2009a, b). The added value of this methodology in clinical pharmacology studies is limited.

Modification of the Method

Several adjustments can be made to this paradigm. The temperature at which the test starts, the rate at which the temperature decreases, and the temperature at which the test ends can be modified. Cold hyperalgesia can be induced by applying menthol to the skin prior to testing (Andersen et al. 2015). A mean threshold of multiple measurements can be taken to minimize subject variability.

Thermal Grill

Purpose and Rationale

The thermal grill is based on the phenomenon in which simultaneous application of innocuous cutaneous warm and cold stimuli can induce a sensation of burning pain, the so-called "thermal grill illusion." The illusion is thought to be caused by the central integration of ascending pain and temperature sensory channels, where the inhibition that is usually exerted by the cold afferents on the nociceptive system is reduced (Bouhassira et al. 2005; Craig and Bushnell 1994).

Procedure and Evaluation

The thermal grill consists of a number of juxtaposed bars of cold and warm nonpainful temperatures (e.g., 18 °C and 42 °C) on which the subject places a body part for a certain period of time. Possible outcome measures, during and after the test, can be cold and hot sensation, pain intensity, and sensation of unpleasantness.

Critical Assessment of the Method

Studies in which the thermal grill has been used applied a range of combinations of warm and cold stimuli to assess relationships between painful and nonpainful sensations (Adam et al. 2014; Kern et al. 2008; Okkerse et al. 2017). The occurrence of paradoxical pain elicited by the thermal grill illusion can be variable. A study by Bouhassira and colleagues reported a large subpopulation of subjects who only reported paradoxical pain when large cold-warm differentials were applied (Bouhassira et al. 2005). Due to the apparent necessity to tailor this method to each individual subject, it is difficult to standardize this method.

Modification of the Method

For the induction of the thermal grill illusion, different temperature combinations of the cold and warm bars can be used. Also a different number and width of the bars and a different distance between the bars can be used.

Skin Freezing

Purpose and Rationale

Skin freezing is an induction method of hyperalgesia. Hyperalgesia can be experimentally induced with chemical and electrical stimulation or by injuring tissue using UVB or freeze lesions (Kilo et al. 1994; Lötsch and Angst 2003).

Procedure and Evaluation

A copper cylinder is cooled to −28 °C and placed to a part of the skin for a brief period of time. For better thermal contact, a filter paper soaked with saline can be placed between the skin and the copper cylinder. This freezing induces cutaneous inflammation and hyperalgesia. Approximately 24 h after induction sensory testing can be performed.

After induction of the freeze lesion, pain and sensation testing can be performed via mechanical stimulation with stroking brushes (subjects have to indicate if a stroke with a certain load is painful), von Frey filaments (with increasing strength, subject have to indicate when the punctuation becomes painful) and blunt pressure using a pressure algometer (threshold in N/cm^2), or electrical stimulation using a constant current device (thresholds in mA).

Critical Assessment of the Method

There are only a handful of studies reporting using freeze lesions. An advantage of this method is the extent to which this methodology can be standardized. The temperature, pressure, and exposure time for induction of the lesion can be controlled. Furthermore, the lesion provides stable test conditions 1 day after induction. The lesion ceases over a period of days (Lötsch and Angst 2003). The freeze lesion may cause hyperpigmentation, which can be visible for several months (Kilo et al. 1994).

Modification of the Method

The location, temperature, pressure, and exposure time for induction of the lesion can be varied. Furthermore, the time window between induction of the lesion and testing can be changed.

Heat Stimulation

Heating Thermode

Purpose and Rationale

Heat pain thresholds can be determined by applying a peltier element to the skin, where the increase in temperature activates nociceptors via TRPV and TRPM channels (Caterina et al. 1997; Li 2017).

Procedure and Evaluation

A contact heat thermode probe, typically with a surface of 9.0–12.5 cm^2, is placed on the skin at a standardized nonpainful baseline temperature between 30 °C and 39 °C. Temperature is subsequently increased in a tonic or phasic fashion at a predetermined rate up to a temperature of 50–52 °C. After a subject has indicated its pain detection or tolerance threshold, the probe is rapidly cooled to the baseline temperature. To reduce variability, the test can be repeated consecutively three times, and the average of these measurements is considered the pain threshold (Bishop et al. 2009). Outcome measures consist of pain thresholds as well as subjective pain scores (NRS, VAS).

Critical Assessment of the Method

Heat pain thresholds are considered to be robust and reproducible endpoints, due to their clear physiological relationship with nociceptor activation thresholds. This method is widely used and contact heat thermodes are commercially available. Limitations to using a contact heating thermode is the relatively slow heating and cooling rate of the thermode, and the fact that the thermode touches the skin compared to, for example, laser, making it less suitable for investigating temporal summation or specific activation of Aδ-fibers.

Modifications of the Method

In addition to investigating pain sensation in healthy skin, this method is often used to quantify sensitization, by comparing pain sensation of normal skin to an area of sensitized (UVB, capsaicin, menthol, cinnamaldehyde) skin (Roberts et al. 2011; Schaffler et al. 2017). The contact heat thermode can be used in conjunction with an EEG or fMRI modality, together known as CHEPS (Contact Heat Evoked Potentials). (Roberts et al. 2008, 2011).

UVB Erythema

Purpose and Rationale

Inflammation is the biological response to any type of bodily injury and is recognized by increased blood flow, elevated cellular metabolism, vasodilatation and the release of soluble mediators, and extravasation of fluids and cellular influx. Many different neuro-active factors are released during inflammation which stimulate nociceptors itself or by lowering the depolarization threshold of afferent nerves. The UVB (or "sunburn") model is regarded as a model for inflammatory pain and as such it is most sensitive to the effects of NSAIDs (Bishop et al. 2009; van Amerongen et al. 2016); in this model, hyperalgesia is evoked by exposing an area of skin to an individualized dose of UVB on the skin.

Procedure and Evaluation

Prior to the start of the study, the minimal erythema dose (MED) for a subject is determined (Sayre et al. 1981). Subsequently, a one-, two-, or threefold multiple of this dose is applied to the skin. Over the course of 2–96 h, a clearly discernible dose-related area of erythema becomes

apparent, where allodynia and hyperalgesia is observed. Maximum hyperalgesia is reached at 24 h after irradiation. Typically, no background pain is observed. UVB induced hyperalgesia or allodynia can be quantified using a thermal (heat or cold) or mechanical (stroking, pinprick, pressure algometry) challenge. Pain thresholds or a subjective pain score can be used as endpoints. Mechanical allodynia to pinpricks or a pressure algometer can be expressed as a PDT, when ascending strengths of von Frey filaments are used. Moreover, the area of allodynia is measured using a fixed von Frey filament or brush.

Critical Assessment of the Method

The UVB model has been proven to be valuable tool to induce hyperalgesia and allodynia associated with inflammatory pain. One caveat, however, is the risk of postinflammatory hyperpigmentation (PIH) (Brenner et al. 2009). PIH is a harmless condition in which areas of skin become darker in color compared to the surrounding skin. PIH can occur at any age and any skin type; however, it is more common in patients with darker skin (Fitzpatrick skin type 4–6) (Fitzpatrick 1988).

Modifications of the Method

In general, there are three degrees of freedom to modify the UVB method: (1) the dose can be altered between estimated 1 to 3 MED (Bauer et al. 2015; Gustorff et al. 2004; Ing Lorenzini et al. 2012); (2) the location can be varied between leg, arm, and back; and (3) the time between UVB exposure and hyperalgesia assessment may vary between 12 h and 36 h.

Heat Burn Model

Purpose and Rationale

A first-degree burn, comparable to a slight sunburn, resulting from a heat stimulus is used to initiate a local inflammatory response which results in reduced pain sensation thresholds (Thalhammer and LaMotte 1982). Additionally, the intense nociceptive excitation is thought to induce central sensitization (Pedersen and Kehlet 1998; Woolf 1983), rendering the burn model a model for both peripheral and central neuronal sensitization.

Procedure and Evaluation

A superficial cutaneous burn is induced using a thermode at a fixed temperature of 45–47 °C, for a period of 2–7 min, which is applied at a standardized pressure on the skin. The leg is predominantly selected as the location, but the arm is also used. The acceptable timeframe for detectable hyperalgesia and allodynia is typically up to 4 h after exposure to the heat stimulus. A distinction in sensitization can be made when investigating responses in the primary (exposed) area and the secondary (adjacent, nonexposed) area. Hyperalgesia resulting from the heat burn model is most distinctly quantified using a thermal or mechanical stimulus, due to locally reduced pain sensation thresholds in the primary area (van Amerongen et al. 2016). The PDT is predominantly selected as an outcome measure. Furthermore, the area of secondary hyperalgesia can be quantified using mechanical (pinprick, stroking) stimuli.

Critical Assessment of the Method

The heat burn model in combination with a mechanical (pinprick) assessment of sensitization is moderately sensitive to the effects of NMDA receptor antagonists (Ilkjaer et al. 1996; Mikkelsen et al. 1999). Analgesic effects of other treatments are less conclusive. As an evoked pain model, its principle is founded in controlled tissue damage, by inducing a first-degree burn, with reports of blistering in up to 20% of the studies conducted with this paradigm (van Amerongen et al. 2016). This may be considered to be an advantage in terms of external validity. However, from an ethical perspective a more short-lasting model without actual tissue damage may be preferred.

Modifications of the Method

The execution can vary from using a contact heat thermode with a short and intense stimulus (100 s at 50 °C), to the more commonly used prolonged exposure at lower temperature (7 min at 47 °C). Other heat sources, including laser stimulation or heat radiation, can be used. The arm or leg can be used as location of exposure.

Lasers

Purpose and Rationale

Laser stimulation (LS) uses the energy to heat up the epidermis and parts of the dermis with very brief (range of ms) and powerful (8–200 mJ/mm^2) stimuli directed at the skin (Plaghki and Mouraux 2003). This type of stimulation causes a characteristic double pain sensation, consisting of an initial sharp pinpricking-like pain (Aδ-fibers) and a second longer burning pain (C-fibers) (Price 1996, 2000). Most commonly used laser stimulators are based on CO_2, Argon, and the YAG (yttrium-aluminum-garnet).

Procedure and Evaluation

LS can stimulate the skin in a well-reproducible manner making it useful as a tool to elicit evoked potentials. Evoked potentials via EEG can easily be registered due to the brief nature of the stimulus of which the timing can be controlled as well. EEG has been used in combination with laser stimulation to distinguish between Aδ- and C-nociceptive activities (Mouraux et al. 2003).

Critical Assessment of the Method

Importantly, during LS no (Aβ-fibers) tactile mechanoreceptors are activated, making LS a useful tool for investigating the nociceptive system without the interference of non-nociceptive input. However, due to the fast rise in temperature of the skin, overstimulation may cause nociceptors to become fatigued over repetitive stimuli (Hüllemann et al. 2015). This in turn has an unwanted effect on quantification of the nociceptive system, as laser evoked potential habituation may occur (Hüllemann et al. 2015; Treede et al. 2003). It is therefore advised to vary the stimulation location slightly after each stimulus and use a randomized inter-stimulus interval. Additionally to habituation, precise settings are necessary to prevent damage to the skin. Hence, power, duration, and surface area must be properly set up.

Modifications of the Method

LS stimulates both Aδ- and C-fibers by thermal activation. However, adjusting the stimulation method may shift the preferential activation of either nociceptor. Preferential C-fiber stimulation is based on a characteristic difference in heat threshold (Aδ-fiber ±46 °C, C-fiber ±40 °C) and distribution density in the upper skin (Ochoa and Mair 1969). Shifting between Aδ and C-fiber activation using LS is possible by choosing the right pulse width, stimulation area in combination with keeping track of the skin temperature, and reaction time.

Thermal Muscle Stimulation

Heated Saline

Purpose and Rationale

Thermosensitive receptors located on muscle tissue afferents are thought to be involved in thermoregulation (Hertel et al. 1976). As such, these have been identified as potential targets to investigate nociception of deep muscle tissue. This is investigated by exposing muscular tissue to a high intensity thermal stimulation (Graven-Nielsen et al. 2002). Only a single study was found using this method.

Procedure and Evaluation

An intramuscular injection of sterile isotonic (1.5 ml) heated saline is injected over 20 s (270 ml h_1) into the musculus tibialis anterior. Hyperalgesia can be quantified using a thermal and mechanical stimulus.

Critical Assessment of the Method

Compared to hypertonic saline in the same study, peak pain score resulting from intramuscular injection of isotonic saline at different temperatures was significantly lower. Mechanical sensitization appeared to be largest after injection at the highest temperature (48 °C). To avoid cutaneous sensations, the injection site was anesthetized with intradermal injections of 0.2 ml lidocaine before the intramuscular injection (Graven-Nielsen et al. 2002).

Modifications of the Method

In the single study using this method, intramuscular injections of isotonic saline at different

temperatures were investigated, ranging from 8 °C to 48 °C. Different muscles can be used for injection.

Thermal Visceral Stimulation

Esophageal

Purpose and Rationale

Thermal stimulation of the GI tract activates specific nociceptive afferents selectively through TRPV1. This is in contrast to mechanical and electrical stimulation, which activate afferents both superficial and deeper in the layers of the viscera (Sengupta and Gebhart 1994). This makes thermal stimulation of the GI tract a useful technique for specific activation of nonmyelinated afferents in the mucosa.

Procedure and Evaluation

In several studies, a model was used to thermally stimulate the esophagus (Arendt-Nielsen et al. 2009; Drewes et al. 2002, 2003b; Krarup et al. 2013). In these experiments, thermal stimuli were performed by changing the temperature (5–60 °C) of recirculating water in a bag that was placed in the lower part of the esophagus. Temperatures were continuously measured inside the bag to control the thermal stimulation in the esophagus. Both for cold as for heat pain, a linear stimulus-response (°C-VAS) can be observed.

Critical Assessment of the Method

The upper GI tract (esophagus) is able to differentiate between thermal stimuli in the temperature range that can be used without chronic damage. Quantification of the visceral pain is more difficult to distinguish.

Modifications of the Method

It has been demonstrated that fast increases in temperature (1.5 °C/min) affect the precision of the response (Olesen et al. 2010). Therefore, the experimental esophageal model can be modified by using slower temperature increases to ascertain better results in pain assessments.

Electrical Stimulation

Electrical stimulation is used extensively for testing the sensitivity of the pain system in studies activating cutaneous structures, muscular structures, and in visceral structures (Andersen et al. 1994; Arendt-Nielsen et al. 1997; Laursen et al. 1997). Electrical stimulation initiates activity in nerve fibers directly without activating receptors. The stimulus intensity determines the size of the current field in the tissue and thereby the number of fibers activated (Andersen et al. 2001). In case a rectangular pulse is applied to the skin, thick fibers mediating mechanoreceptive input are activated at the lowest stimulus intensities. Increasing the stimulus intensity leads to concurrent activation of thin myelinated fibers (Aδ-fibers) and eventually C-fibers.

Electrical Skin Stimulation

Stimulation can be done cutaneous or intracutaneous with various stimulation paradigms with diverse waveforms, frequencies, and durations to selectively activate different afferents and nervous structures and thereby evoke various pain sensations. In addition, summated neural activity, as a result of the stimuli, can activate central mechanisms (Koppert et al. 2001), which is described further in paragraph "Electrical Single Stimulation."

Electrical Single Stimulation

Purpose and Rationale

This electrical stimulation paradigm leads to a nociceptive, Aδ- and C-fiber mediated type of pain, which is well controllable. The electrical current stimulates nerve fibers directly because the intensity is far below that required to stimulate the actual receptors in the skin (Dotson 1997).

Procedure and Evaluation

For cutaneous electrical pain, two electrodes (Ag-AgCl) are placed on clean (scrubbed) skin, e.g., the skin overlying the tibial bone. Electrical resistance between electrodes should be less than 2 kΩ. Each stimulus (10 Hz tetanic pulse with a

duration of 0.2 ms) is controlled by a computer-controlled constant current stimulator. Current intensity increases from 0 mA in steps of 0.5 mA/s (cutoff 50 mA). The pain intensity after each stimulation is measured using an (electronic) VAS, until pain tolerance level is reached or a maximum of 50 mA is reached (Olofsen et al. 2005).

Critical Assessment of the Method

Electrical stimulation is easily controlled. Electrical stimulation of the skin to induce pain has several shortcomings: (1) they excite the afferent pathways in an unnatural synchronized manner; (2) they excite the full spectrum of peripheral nerve fibers (Aβ-, Aδ-, and C-fibers); and (3) stimulation bypasses the receptors on the sensory nerve endings, and therefore, all information on specific activation and transduction processes is lost (Handwerker and Kobal 1993).

Modifications of the Method

Aδ- and C-fibers are activated at different stimulus intensities where C-fibers have a higher activation threshold. Modeling approach can be useful for certain drug trials (Handwerker and Kobal 1993; Lee et al. 2007). The nonspecificity toward nociceptive specific stimulation is thought to be overcome by using small specialized needle-like electrodes. These electrodes slightly protrude through the epidermis and can preferentially stimulate nociceptive Aδ-fibers (Bromm and Lorenz 1984; Inui and Kakigi 2012; Mouraux et al. 2010). Intracutaneous stimulation can be chosen to mimic more a stinging/burning sensation and less throbbing (Bromm et al. 1984b).

Electrical Burst (Temporal Summation)

Purpose and Rationale

Increasing pain in response to a series of stimuli (temporal summation) reflects the first phase of "wind-up" in animal studies. Temporal summation can be induced with mechanical, thermal, and electrical stimulation (Arendt-Nielsen et al. 2000; Granot et al. 2006; Mauderli et al. 2003; Nie et al. 2006). Temporal and spatial summation evoked in the skin reflects a central nervous system modulation of the response, and it is believed to mimic neuropathic pain conditions because a likely contribution of central sensitization to neuropathic pain has been demonstrated (Woolf 2011). Application of transcutaneous electrical stimuli, with variation in electrical burst frequency, has been shown to be a reliable model to induce temporal summation in human subjects (Arendt-Nielsen et al. 2000).

Procedure and Evaluation

For burst stimulus, each single stimulus is repeated 5 times with a frequency of 2 Hz. Pain threshold is taken as the value (mA) whereby a subject indicates either: all 5 stimuli are painful, or the train of 5 stimuli started feeling nonpainful but ends feeling painful (VAS $>$ 0) (Arendt-Nielsen et al. 2000; Hay et al. 2016).

Critical Assessment of the Method

As facilitated temporal summation is a feature in neuropathic pain patients, it has been hypothesized that induction of temporal summation using electrical stimulation can be used as a biomarker of drug effects on neuropathic pain (Arendt-Nielsen et al. 2007b). In a recent study, drug effects of analgesic compounds, including several used in the treatment of neuropathic pain, could not be established using this evoked pain paradigm, while other evoked pain paradigms manage to demonstrate pharmacological effects convincingly. This appeared to be related to a higher intra-subject variability that may necessitate larger subject groups (Okkerse et al. 2017).

High-Frequency Electrical Stimulation

Purpose and Rationale

High-frequency electrical stimulation (HFS) of the human skin induces increased pain sensitivity in the surrounding unconditioned skin (Van den Broeke et al. 2014). It has been shown that sustained nociceptive input can induce activity-dependent changes in synaptic strength within nociceptive pathways, leading to an amplification of nociceptive signals (Ikeda et al. 2003). This is thought to play a key role in the development and maintenance of chronic pain and in particular some forms of hyperalgesia (Latremoliere and

Woolf 2009; Sandkühler 2009). HFS-induced hyperalgesia within the surrounding unconditioned skin mimics the phenomenon of secondary hyperalgesia (Meyer and Treede 2004). As such, it constitutes a suitable model to study the mechanisms underlying central sensitization of nociceptive pathways (Klein et al. 2008).

Procedure and Evaluation

HFS is delivered to the test site, e.g., the volar forearm, and consists of 5 trains of 100 Hz pulses lasting 1 s, (10 s interstimulus interval; 2 ms single pulse duration) at 10 times the detection threshold (Pfau et al. 2011). The electrical stimulation is generated by a constant-current electrical stimulator and delivered to the skin using a specifically designed electrode that has been demonstrated to activate peptidergic nociceptive afferents in the skin (Klein et al. 2004). The heterotopical effect of HFS is usually characterized using mechanical punctate stimuli. The test stimuli are applied to the skin surrounding the area onto which HFS is applied as well as to the same skin area on the contralateral arm, which serves as control to take into account a possible time-dependent habituation (van den Broeke et al. 2014).

The intensity of perception elicited by the three types of test stimuli is assessed using a numerical rating scale (NRS). After approximately 1 h, the level of heterotopical hyperalgesia starts to diminish, however is still measurable and significant from baseline up to 8 h after HFS (Pfau et al. 2011).

Critical Assessment of the Method

HFS offers an alternative to other models that lead to secondary hyperalgesia, such as the capsaicin model or the UVB model with some important advantages. The major advantage versus the UVB model is that the mechanism underlying the secondary hyperalgesia is thought to involve heterosynaptic facilitation and, hence, to constitute a suitable model of central sensitization of nociceptive pathways (Klein et al. 2008), while the secondary hyperalgesia in the UVB model is thought to be due to a more peripheral sensitization of nociceptors, induced by inflammation (Bishop et al. 2009). The interval during which the secondary hyperalgesia is measurable can be carefully regulated and is relatively stable over the first hour after application of HFS (Pfau et al. 2011). The major disadvantage of this method is that is has not been used to demonstrate pharmacodynamic effects of analgesic drugs, although one recent study did show that the model can be combined with the assessment of drug effects (Vo et al. 2016).

Modifications of the Method

Modifications of the method are primarily related to the type of sensory stimulus to determine the heterotopical hyperalgesic effect and to the quantification. Heterotopical hyperalgesia can be demonstrated for mechanical punctate stimuli, but also for thermonociceptive stimuli induced by heat probes or laser stimulation. Van den Broeke et al. used the model in conjunction to event related potentials to objectively demonstrate the hyperalgesic phenomena (Van den Broeke et al. 2014).

Electrical Muscle Stimulation

Purpose and Rationale

Electrical stimulation of muscle tissue can be used to elicit both local and referred muscle pain. It possesses the ability to generate referred muscle pain in an "on and off" manner, and it is capable of maintaining referred pain for at least 10 min (Laursen et al. 1997). Intramuscular electrical stimulation appears to be used more often to study the nature of muscle pain than as a model to determine the pharmacodynamic effects of new analgesic compounds.

Procedure and Evaluation

In the intramuscular electrical stimulation paradigm, two needle electrodes with uninsulated tips are inserted into a muscle (e.g., the musculus tibialis anterior). A computer-controlled constant current stimulator is used to induce referred pain in the ventral part of the ankle by stimulating the muscle (Laursen et al. 1997). Each stimulation consists of five constant current rectangular pulses (1 ms) delivered at 200 Hz. The referred pain threshold is defined as the lowest stimulus intensity required for the subject to notice a "just barely

painful" sensation in the referred pain area. Referred pain thresholds are determined by a staircase regime consisting of five ascending and four descending series of stimuli (Gracely 1994; Laursen et al. 1997).

Critical Assessment of the Method

Electrical muscle and skin stimulation can use the same modalities which makes it possible to compare both models. A disadvantage of the model is that referred pain due to intramuscular electrical stimulation does not occur in all subjects; approximately, three quarters of patients experience it (Laursen et al. 1997). The referred pain typically arises approximately 40 s after the onset of electrical stimulation, which may mean that temporal summation is involved (Laursen et al. 1997).

Modifications of the Method

Modifications can be made with the stimulation settings. Pulse range of 100–200 Hz has been described, as well as a pulse width of 1–2 s (Laursen et al. 1997; van den Broeke et al. 2014).

Electrical Visceral Stimulation

In the viscera, it is difficult to determine the pain threshold to a single stimulus, whereas the pain threshold is easily determined if a train of stimuli is used. Furthermore, the referred pain area gradually expands if stimulation is continued for 120 s (Arendt-Nielsen et al. 1997).

Chemical Stimulation

Administration of algogenic substances to the skin, muscle, or viscera is believed to be a close resemblance of clinical inflammation. Various substances have been used to induce cutaneous hyperalgesia. The most commonly used are capsaicin, nerve growth factor (NGF), glutamate, mustard oil, and menthol, but other chemical stimulation models exist as well. Intramuscular injection of chemical substances is less common and harder to control in a clinical trial. The esophagus is the target organ when it comes to chemical viscera stimulation because of its easy access.

Chemical Skin Stimulation

Capsaicin

Purpose and Rationale

Capsaicin is a highly selective agonist for TRPV1, notorious for its pungent property in red chili peppers. TRPV1 channels are major transducers of physically and chemically evoked sensations (Hauck et al. 2015). The vanilloid 1 subtype is activated by noxious heat ($\geq$ 43 °C) (Frølund and Frølund 1986) and is expressed on C-fibers, and on a subset of Aδ-fibers (Le Bars et al. 1979). The direct effects of applying topical capsaicin are burning sensations, hyperalgesia, allodynia, and erythema. In addition, it triggers the release of proinflammatory agents at peripheral terminals, such as substance P and calcitonin gene-related peptide (CGRP) (Kakigi 1994; Yarnitsky et al. 2010).

Procedure and Evaluation

Capsaicin can be administered topically and intradermal. Intradermal injection with capsaicin 0.1 mg can cause hyperalgesia, but a dose of 100 mg or higher is needed to produce hyperalgesia for an hour (Simone et al. 1987). A dose of 100 mg is most frequently used (Baron et al. 1999; Serra et al. 1998; Torebjörk et al. 1992).

Topical administration of capsaicin in low concentrations (up to 3%) can cause temporary mechanical and heat hyperalgesia. Sensitization can be induced by preheating the skin to 45 °C for 5 min with a thermode directly before capsaicin application. Sensitization can be rekindled throughout a study by reheating the skin up to 40 °C for 5 min. Application of the capsaicin is most commonly done on the forearm or the back, but can be done on any area of the skin. Topical application of capsaicin can induce peripheral and central sensitization shown respectively by primary mechanical/thermal hyperalgesia and by secondary mechanical hyperalgesia/allodynia. This pain model can therefore be used to study novel analgesic compounds targeting these typical symptoms of neuropathic pain. Peripheral sensitization is caused by modulation of peripheral

afferents and is therefore restricted to the site of injury, i.e., primary hyperalgesia. Central sensitization is caused by modulation of the nociceptive processing in the central nervous system. To quantify the effects of this pain model, laser stimulation (LS) can be used in combination with electro-encephalogram (EEG).

Critical Assessment of the Method

Peripheral sensitization is closely linked to primary hyperalgesia, and central sensitization is partly explained by hyperalgesia in the surrounding area, i.e., secondary hyperalgesia. Moreover, nociceptive integration at spinal cord level may include non-nociceptive mechanoreceptors. Therefore, central sensitization may also cause Aβ-fiber mediated pain (allodynia). Higher concentrations (capsaicin 8%) initially causes increased sensitivity but is then followed by a decrease in sensitivity due to a reduced TRPV1 expression (Messeguer et al. 2006; van Amerongen et al. 2016). High concentration capsaicin is indicated in postherpetic neuralgia. Besides, capsaicin may also have a neurolytic property, where it (partly) eliminates epidermal nerve fibers (ENFs) in treated areas over time (Dworkin et al. 2010). Re-innervation occurs over time (Hüllemann et al. 2015).

Modifications of the Method

There are several variations that need to be addressed when designing a study utilizing capsaicin, e.g., concentration of the capsaicin, dose administration (intradermal or topical), vehicle of the capsaicin (alcohol or cream), duration of the application, location of administration, and pre-/rekindling.

Nerve Growth Factor Injection

Purpose and Rationale

NGF is a member of the neurotrophin family, which also includes brain derived neurotrophic factor (BDNF), neurotrophin-3 (NT3) and neurotrophin-4/5 (NT4/5). NGF binds to both a high affinity tyrosine kinase receptor trkA and a low affinity receptor p75. NGF can sensitize nociceptors so that they show an increased response to thermal and chemical stimuli (Bennett 2007). Administration of NGF to human skin evokes mechanical sensitization and profound hyperalgesia to thermal stimuli that develops within 3 h postinjection and peaks between day 1 and 7 (Dyck et al. 1997). Sensitization to heat and hyperalgesia to cold develops within days after injection and lasts up to 21 days, while hypersensitivity to mechanical impact stimuli develops over a longer period and persists for at least 49 days (Rukwied et al. 2010). Intradermal NGF administration provokes a pattern of sensitization that can be used as experimental model for neuropathic pain (Rukwied et al. 2010).

Procedure and Evaluation

One microgram of human recombinant lyophilized NGF is dissolved in 50 μL saline and injected intradermally into the central volar forearm. The same volume of saline is administered into the contra-lateral site as vehicle control (Rukwied et al. 2010). Vasodilatation upon NGF- and saline-injection can be recorded by laser Doppler imaging. Nociceptor sensitization can be explored to mechanical (touch, pinprick, pressure), thermal (cold, heat), and electrical (current pulses) stimuli. Stimuli for investigating static and dynamic allodynia and pinprick hyperalgesia are administered 5–7 cm distal from the injection site and continued in steps of 1 cm until the subject reports a definite increase of pinprick pain or switch from touch to an aversive sensation (Rukwied et al. 2010). The point where this starts is marked on the skin and the distance to the injection site measured. Pain thresholds and subjective scores with NRS/VAS can be used to evaluate the mechanical, thermal, or electrical stimulation.

Critical Assessment of the Method

Increased levels of NGF have been reported in human painful disorders including arthritis (Kidd and Urban 2001). Injection of NGF therefore appears to mimic processes found in clinical disease (Olesen et al. 2012). Even though NGF may also be upregulated in the UVB burn (Bishop et al.

2007), anti-NGF has been shown to only partially reduce UVB induced hyperalgesia (Bishop et al. 2007). Apparently, the NGF induced mechanism of mechanical sensitization is different to UVB evoked primary hyperalgesia. NGF induces a particularly long lasting mechanical sensitization including static allodynia and cold hyperalgesia without any visible signs of inflammation and therefore adds to the spectrum of human evoked pain models (Rukwied et al. 2010). The long-lasting local allodynia and hyperalgesia after subcutaneous or intradermal injection, up to 49 days after injection, form the most important disadvantage of the model. Even though considered a model for neuropathic pain, it is unlikely that central sensitization plays a role.

Modifications of the Method

Systemic administration of NGF 1 μg/kg i.v. has been shown to lead to mild to moderate muscle pain mainly in the bulbar and truncal musculature that lasted 2–8 days (Petty et al. 1994).

NGF has been injected into the musculus masseter to induce allodynia and hyperalgesia and as a model of myofascial temporomandibular disorder pain (Svensson et al. 2003).

Other Chemical Mediated Models

Mustard Oil

Mustard oil is a plant-derived irritant. The noxious effects of mustard oil are currently ascribed to specific activation of the cation channel transient receptor potential, subfamily A, member 1 (TRPA1) in nociceptive neurons (Olesen et al. 2012). Topical administration leads to a burning pain in the area exposed to mustard oil as well as secondary allodynia and hyperalgesia in the surrounding unaffected area, similar to the topical capsaicin model (Koltzenburg et al. 1992).

Menthol

Menthol acts as an agonist on the transient receptor potential cation channel subfamily M member 8 (TRPM8) receptor. The topical application of high concentration (40%) menthol is thought to activate and sensitize cold-sensitive TRPM8-expressing C-nociceptors and activates cold-specific Aδ-fibers (Binder et al. 2011). Topical application has been used as an evoked pain model of cold hyperalgesia, which is a clinical symptom that occurs frequently in patients with peripheral or central nervous system lesions (Hatem et al. 2006). In addition to cold hyperalgesia, the model elicits primary and secondary mechanical (pinprick) hyperalgesia combined with the sensation of burning (Binder et al. 2011). The menthol model has been shown to be sensitive to a range of analgesics (Altis et al. 2009).

Chemical Muscle Stimulation

Nerve Growth Factor

Intramuscular injection with NGF is most commonly done in the musculus tibialis anterior or musculus masseter (Andersen et al. 2008; Svensson et al. 2008). It induced a long-lasting hyperalgesia and lower pressure pain threshold can be observed, lasting up to 4 days in the musculus tibialis anterior and up to 14 days in the musculus masseter (Andersen et al. 2008; Svensson et al. 2008). An advantage of the intramuscular NGF paradigm is the long-lasting hyperalgesia which can simulate clinical pain more than most other paradigms, but this is also the disadvantage where ethical consideration may play a role. The paradigm is difficult to control where hyperalgesia is dependent on the dose and the size of the muscle (Andersen et al. 2008).

Chemical Visceral Stimulation

Esophagal (Gut) Perfusion with Acid, Alcohol, Glycerol, Capsaicin, and Hypertonic Saline

Purpose and Rationale

Chemical stimulation of the GI tract may be used to stimulate C-fibers selectively via TRPV1 receptors and modulate the visceral pain system due to their sensitization effects. Having a model of central sensitization of the viscera can be helpful in the development of new analgesics, as this is

thought to be an important element of chronic visceral pain.

Procedure and Evaluation

Using acid to stimulate the esophagus is the most used method to sensitize the gut (Bernstein and Baker 1958; Demedts and Tack 1998; Drewes et al. 2005; Reddy et al. 2005). However, other chemicals such as alcohol, glycerol, capsaicin, and hypertonic saline are used to stimulate the gut as well (Louvel et al. 1996; Drewes et al. 2003a, b). The chemical compound is usually infused into a container/bag residing in the esophagus with a small perfusion hole to release the compound into the esophagus. Chemical stimulation is able to modulate the visceral pain system by selectively activating non-myelinated C-fibers for a longer amount of time. This tonic activation may result in central sensitization effects, which can be quantified by subsequent thermal, electrical, or mechanical stimulation.

Critical Assessment of the Method

A high variation in the outcome measures is seen with this model. The reproducibility is challenging because several factors are hard to control, like exposure time to the chemical stimulus, size of the treated area, and latency time to onset of effects. Furthermore, tissue injury results in the release of multiple molecules working together, and to mimic this situation it may be necessary to use a mixture of chemical substances (Reddy et al. 2005). Blinding this procedure is difficult, since subjects are able to taste the compound. Therefore, the experimental setup requires that both subject and assessor are ignorant of the possible influence of the compound on the pain threshold (Drewes et al. 2003b).

Modifications of the Method

Each chemical substance will have an impact on the results. For example, the motility may interfere with the results when glycerol is used. Other stimuli, such as injection of hypertonic saline and application of capsaicin, the pain is elicited shortly after the chemical comes into contact with the mucosa, and the motility has minor impact on the results.

Discussion

Healthy Subjects Versus Patients

Despite many advances in the last decades in understanding pain, the development of new analgesic compounds lacked behind. In almost 60 years, only 59 compounds were registered for the treatment of pain, of which two thirds were specifically developed as analgesics (Kissin 2010). Historically, pain states have been classified and investigated on the basis of a disease state. Based on preclinical animal models, target patient populations were selected. In patient studies, efficacy is then reported as change in the patient's response to pain (McQuay and Moore 2013). Unfortunately, several promising compounds have failed in this late-stage development where pharmacotherapy only provides meaningful pain relief in less than 50% of patients with neuropathic pain (Finnerup et al. 2010, 2015). But a negative outcome does not automatically mean inefficacy of the compound. Pathophysiological mechanisms of pain vary between individuals with the disease state. Selecting and clustering the patients in groups of pathophysiology rather than disease might be necessary to obtain meaningful results. The use of human evoked pain models can provide more information.

Multimodal testing in healthy volunteers can provide information about the analgesic activity of the compound and possibly find the active dose level range. In a way, by using different pain modalities, the results will create a certain pain profile of the compounds (Okkerse et al. 2017). These results may reflect effects of analgesic drugs on mechanisms involved in clinical pain. Thus, multimodal pain testing may aid in determining the optimal target population for new analgesic compounds based on their profile of effects on a diversity of pain mechanisms and depending on the contribution of each of these mechanisms in clinical pain phenotypes. In several chronic pain populations, such as chronic whiplash,

rheumatoid arthritis, vulvodynia, and fibromyalgia, changes in pain tolerance levels, pain modulation, and augmented brain responses and altered responses to analgesics have been found (Daenen et al. 2014; Hampson et al. 2013; van Laarhoven et al. 2013). Using evoked pain in these patients can provide insight into the analgesic mechanisms – or lack thereof – in these altered pain states (Olesen et al. 2012). In patients with chronic (neuropathic) pain, different sensory profiles exist. These profiles possibly match with different neurobiological mechanism of pain (Baron et al. 2017).

Predictive Value of Models for Drug Development

Human evoked pain models in healthy volunteers can be conducted in standardized laboratories. Factors like stimulus intensity, frequency, duration, and location can be controlled, and when a model is stable and reproducible, it can be regarded as suitable for pharmacodynamic evaluation of new analgesic drugs. Using pain models in healthy volunteers has important advantages over assessing the effects of new drugs in patients with pain; the pain elicited in human pain models is predictable in its intensity while clinical pain will naturally fluctuate, and in pain models analgesic properties can be investigated without the influence of accompanying symptoms that are often seen in patients with pain. However, it should always be asked whether a pain model at all resembles naturally occurring pain. Clinical pain is a subjective perception, influenced by cognitive processes, by emotions, social context, and even cultural background, while pain models are solely based on the infliction of a noxious stimulus and its response. An important question is whether or not a positive result in a certain evoked model is also predictive of clinical efficacy.

Two approaches have been used to investigate this. Moore and colleagues investigated which naturally occurring pain was physiologically most in agreement with evoking a pain response causing the same type of pain. For instance, they concluded that intramuscular electrical stimulation closely matched clinical acute musculoskeletal pain (Moore et al. 2013). Oertel and Lötsch evaluated the differences between human pain models and clinical efficacy. First they looked at which drugs were effective in different pain conditions (e.g., NSAIDs were effective for inflammatory arthritis), then they investigated which drugs were effective in which pain model (e.g., NSAIDs influence pain response in laser evoked pain). If a certain drug was effective both in the model and in the particular clinical setting, the model was concluded to possibly be predictive for the type of clinical pain. Some level of agreement could be observed for a large number of pain models with many different clinical forms of pain (Oertel and Lötsch 2013). In another review, the mutual agreement between pain models and clinical efficacy was statistically assessed. It was observed that a small set of pain models seemed predictive for efficacy in the clinic, for example, capsaicin induced hyperalgesia with mechanical stimulation is associated with trigeminal neuralgia and renal colic, and UVB induced hyperalgesia in combination with heat stimulation can be linked to burn injuries or postoperative pain (Lötsch et al. 2014).

Several reviews investigated which evoked pain models were sensitive to the analgesic effects of different classes of analgesics in healthy subjects (Oertel and Lötsch 2013; Okkerse et al. 2017; Staahl et al. 2009a, b). With the aid of these studies, well-considered decisions can be made on which evoked pain models to include in studies investigating potentially analgesic compounds.

Multi-model Assessment of Pain

Pain comes in various types and can originate in many different tissues. It is obvious that different analgesics will influence different types of pain according to their respective mechanism of action. If an analgesic drug with a novel mechanism of action is studied, it can occur that a single pain model, thought to relate to a specific clinical pain syndrome, demonstrates lack of efficacy of the new compound. In these cases, a combination of human evoked pain models can be used to screen for possible analgesic effects of these compounds. For instance, a combination of a mechanical,

thermal, and electrical pain models: pressure stimulation assesses the nociception generated from within the muscle (Polianskis et al. 2001); cold pain induced by the cold pressor test mainly activates C-fibers in the skin (Olesen et al. 2012); heat stimulation initially activates Aδ-fibers in the skin, followed by C-fiber activation; induction of inflammation via sunburn or UVB induces the production of cytokines that lead to sensitization of cutaneous nociceptors (Bishop et al. 2009); and electrical stimulation directly stimulates sensory nerve endings of both Aδ and C-fibers in the skin (Handwerker and Kobal 1993). This multimodal testing with a battery of different pain models has been performed by multiple study groups (Enggaard et al. 2001; Okkerse et al. 2017; Olesen et al. 2014; Staahl et al. 2006). The batteries have in common that they induce pain via different modalities and in different tissues and mimics clinical pain better than a single pain model can. The multimodal batteries can be used to profile the analgesic effects of new drugs, to obtain the optimal dose of new analgesics, and to benchmark new drugs against profiles of well-known analgesics (Okkerse et al. 2017).

References and Further Reading

Adam F, Alfonsi P, Kern D, Bouhassira D (2014) Relationships between the paradoxical painful and nonpainful sensations induced by a thermal grill. Pain 155:2612–2617

Aguggia M (2003) Neurophysiology of pain. Neurol Sci 24:57–60

Altis K, Schmidtko A, Angioni C, Kuczka K et al (2009) Analgesic efficacy of tramadol, pregabalin and ibuprofen in menthol-evoked cold hyperalgesia. Pain 147:116–121

van Amerongen G, de Boer MW, Groeneveld GJ, Hay JL (2016) A literature review on the pharmacological sensitivity of human evoked hyperalgesia pain models. Br J Clin Pharmacol 82:903–922

Andersen OK, Jensen LM, Brennum J, Arendt-Nielsen L (1994) Evidence for central summation of C and A delta nociceptive activity in man. Pain 59:273–280

Andersen OK, Sonnenborg FA, Arendt-Nielsen L (2001) Reflex receptive fields for human withdrawal reflexes elicited by non-painful and painful electrical stimulation of the foot sole. Clin Neurophysiol 112:641–649

Andersen H, Arendt-Nielsen L, Svensson P, Danneskiold-Samsøe B et al (2008) Spatial and temporal aspects of muscle hyperalgesia induced by nerve growth factor in humans. Exp Brain Res 191:371–382

Andersen HH, Poulsen JN, Uchida Y, Nikbakht A et al (2015) Cold and L-menthol-induced sensitization in healthy volunteers—a cold hypersensitivity analogue to the heat/capsaicin model. Pain 156:880–889

Arendt-Nielsen L, Yarnitsky D (2009) Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. J Pain 10:556–572

Arendt-Nielsen L, Drewes AM, Hansen JB, Tage-Jensen U (1997) Gut pain reactions in man: an experimental investigation using short and long duration transmucosal electrical stimulation. Pain 69:255–262

Arendt-Nielsen L, Sonnenborg FA, Andersen OK (2000) Facilitation of the withdrawal reflex by repeated transcutaneous electrical stimulation: an experimental study on central integration in humans. Eur J Appl Physiol 81:165–173

Arendt-Nielsen L, Curatolo M, Drewes A (2007a) Human experimental pain models in drug development: translational pain research. Curr Opin Investig Drugs 8:41–53

Arendt-Nielsen L, Frøkjaer JB, Staahl C, Graven-Nielsen T et al (2007b) Effects of gabapentin on experimental somatic pain and temporal summation. Reg Anesth Pain Med 32:382–388

Arendt-Nielsen L, Olesen AE, Staahl C, Menzaghi F et al (2009) Analgesic efficacy of peripheral K-opioid receptor agonist CR665 compared to oxycodone in a multimodal, multi-tissue experimental human pain model. Anesthesiology 111:616–624

Baron R, Baron Y, Disbrow E, Roberts TP (1999) Brain processing of capsaicin induced secondary hyperalgesia: a functional MRI study. Neurology 53:548–557

Baron R, Maier C, Attal N, Binder A et al (2017) Peripheral neuropathic pain: a mechanism-related organizing principle based on sensory profiles. Pain 158:261–272

Bauer M, Schwameis R, Scherzer T, Lang-Zwosta I et al (2015) A double-blind, randomized clinical study to determine the efficacy of benzocaine 10% on histamine-induced pruritus and UVB-light induced slight sunburn pain. J Dermatolog Treat 26:367–372

Bennett D (2007) Sensitization of nociceptors. In: Schmidt RF, Willis WD (eds) Encyclopedia of pain, 1st edn. Springer-Verlag, Berlin/Heidelberg, pp 1338–1342

Bernstein LM, Baker LA (1958) A clinical test for esophagitis. Gastroenterology 34:60–81

Binder A, Stengel M, Klebe O, Wasner G et al (2011) Topical high-concentration (40%) menthol-somatosensory profile of a human surrogate pain model. J Pain 12:764–773

Bishop T, Hewson DW, Yip PK, Fahey MS et al (2007) Characterisation of ultraviolet-B-induced inflammation as a model of hyperalgesia in the rat. Pain 131:70–82

Bishop T, Ballard A, Holmes H, Young AR et al (2009) Ultraviolet-B induced inflammation of human skin: characterisation and comparison with traditional models of hyperalgesia. Eur J Pain 13:524–532

Bonica JJ (1979) The need of a taxonomy. Pain 6:247–252

Borsook D, Sava S, Becerra L (2010) The pain imaging revolution: advancing pain into the 21st century. Neuroscientist 16:171–185

Bouhassira D, Sabaté JM, Coffin B, Le Bars D et al (1998) Effects of rectal distensions on nociceptive flexion reflexes in humans. Am J Phys 275:410–417

Bouhassira D, Kern D, Rouaud J, Pelle-Lancien E et al (2005) Investigation of the paradoxical painful sensation ('illusion of pain') produced by a thermal grill. Pain 114:160–167

Brenner M, Coelho SG, Beer JZ, Miller SA et al (2009) Long-lasting molecular changes in human skin after repetitive in situ UV irradiation. J Invest Dermatol 129:1002–1011

Brennum J, Kjeldsen M, Jensen K, Jensen TS (1989) Measurements of human pressure-pain thresholds on fingers and toes. Pain 38:211–217

van den Broeke EN, Geene N, van Rijn CM, Wilder-Smith OH et al (2014) Negative expectations facilitate mechanical hyperalgesia after high-frequency electrical stimulation of human skin. Eur J Pain 18:86–91

Bromm B, Lorenz J (1984) Neurophysiological evaluation of pain. Electroencephalogr Clin Neurophysiol 107:227–253

Bromm B, Jahnke MT, Treede RD (1984) Responses of human cutaneous afferents to CO2 laser stimuli causing pain. Exp Brain Res 55:158–166

Burke D, Mackenzie RA, Skuse NF, Lethlean AK (1975) Cutaneous afferent activity in median and radial nerve fascicles: a microelectrode study. J Neurol Neurosurg Psychiatry 38:855–864

Casale R, Glynn C, Buonocore M (1992) The role of ischaemia in the analgesia which follows Bier's block technique. Pain 50:169–175

Caterina MJ, Schumacher MA, Tominaga M, Rosen TA et al (1997) The capsaicin receptor: a heat-activated ion channel in the pain pathway. Nature 389:816–824

Craig AD, Bushnell MC (1994) The thermal grill illusion: unmasking the burn of cold pain. Science 265:252–255

Curatolo M, Petersen-Felix S, Arendt-Nielsen L, Zbinden AM (1997) Epidural epinephrine and clonidine: segmental analgesia and effects on different pain modalities. Anesthesiology 87:785–794

Curatolo M, Petersen-Felix S, Arendt-Nielsen L (2000) Sensory assessment of regional analgesia in humans: a review of methods and applications. Anesthesiology 93:1517–1530

Daenen L, Nijs J, Cras P, Wouters K et al (2014) Changes in pain modulation occur soon after whiplash trauma but are not related to altered perception of distorted visual feedback. Pain Pract 14:588–598

Demedts I, Tack J (1998) Chest pain of esophageal origin. Curr Opin Gastroenterol 14:340–344

Doll RJ, Buitenweg JR, Meijer HG, Veltink PH (2014) Tracking of nociceptive thresholds using adaptive psychophysical methods. Behav Res Methods 46:55–66

Doll RJ, van Amerongen G, Hay JL, Groeneveld GJ et al (2016) Responsiveness of electrical nociceptive detection thresholds to capsaicin (8%)-induced changes in nociceptive processing. Exp Brain Res 234:2505–2514

Dotson RM (1997) Clinical neurophysiology laboratory tests to assess the nociceptive system in humans. J Clin Neurophysiol 14:32–45

Drewes AM, Schipper KP, Dimcevski G, Petersen P et al (2002) Multimodal assessment of pain in the esophagus: a new experimental model. Am J Physiol Gastrointest Liver Physiol 283:95–103

Drewes AM, Schipper KP, Dimcevski G, Petersen P et al (2003a) Gut pain and hyperalgesia induced by capsaicin: a human experimental model. Pain 104:333–341

Drewes AM, Schipper KP, Dimcevski G, Petersen P et al (2003b) Multi-modal induction and assessment of allodynia and hyperalgesia in the human oesophagus. Eur J Pain 7:539–549

Drewes AM, Reddy H, Staahl C, Pedersen J et al (2005) Sensory-motor responses to mechanical stimulation of the esophagus after sensitization with acid. World J Gastroenterol 11:4367–4374

Dworkin RH, O'Connor AB, Audette J, Baron R et al (2010) Recommendations for the pharmacological management of neuropathic pain: an overview and literature update. Mayo Clin Proc 85:3–14

Dyck PJ, Peroutka S, Rask C, Burton E et al (1997) Intradermal recombinant human nerve growth factor induces pressure allodynia and lowered heat-pain threshold in humans. Neurology 48:501–505

Eckhardt K, Li S, Ammon S, Schanzle G et al (1998) Same incidence of adverse drug events after codeine administration irrespective of the genetically determined differences in morphine formation. Pain 76:27–33

Enggaard TP, Poulsen L, Arendt-Nielsen L, Hansen SH et al (2001) The analgesic effect of codeine as compared to imipramine in different human experimental pain models. Pain 92:277–282

Finnerup NB, Sindrup SH, Jensen TS (2010) The evidence for pharmacological treatment of neuropathic pain. Pain 150:573–581

Finnerup NB, Attal N, Haroutounian S, Kamerman P et al (2015) Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. Lancet Neurol 14:162–173

Fitzpatrick TB (1988) The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 124:869–871

Frølund F, Frølund C (1986) Pain in general practice: pain as a cause of patient-doctor contact. Scand J Prim Health Care 4:97–100

Fruhstorfer H, Gross W, Selbmann O (2001) Von Frey hairs: new materials for a new design. Eur J Pain 5:341–342

Fujii K, Motohashi K, Umino M (2006) Heterotopic ischemic pain attenuates somatosensory evoked potentials induced by electrical tooth stimulation: diffuse noxious inhibitory controls in the trigeminal nerve territory. Eur J Pain 10:495–504

Ge HY, Madeleine P, Cairns BE, Arendt-Nielsen L (2006) Hypoalgesia in the referred pain areas after bilateral injections of hypertonic saline into the trapezius

muscles of men and women: a potential experimental model of gender-specific differences. Clin J Pain 22:37–44
Goubert D, Danneels L, Cagnie B, van Oosterwijck J et al (2015) Effect of pain induction or pain reduction on conditioned pain modulation in adults: a systematic review. Pain Pract 8:765–777
Gracely RH (1994) Studies of pain in normal man. In: Wall PD, Melzack R (eds) Textbook of pain, 3rd edn. Churchill Livingstone, Edinburgh, pp 315–336
Gracely RH (2013) Studies of pain in human subjects. In: McMahon SB, Koltzenburg M, Tracey I, Turk DC (eds) Wall and Melzack's textbook of pain, 6th edn. Elsevier, Philadelphia, pp 283–300
Granot M, Granovsky Y, Sprecher E, Nir RR, Yarnitsky D (2006) Contact heat-evoked temporal summation: tonic versus repetitive-phasic stimulation. Pain 122:295–305
Graven-Nielsen T, Mense S (2001) The peripheral apparatus of muscle pain: evidence from animal and human studies. Clin J Pain 17:2–10
Graven-Nielsen T, Arendt-Nielsen L, Svensson P, Staehelln Jensen T (1997) Quantification of local and referred muscle pain in humans after sequential i.m. injections of hypertonic saline. Pain 69:111–117
Graven-Nielsen T, Babenko V, Svensson P, Arendt-Nielsen L (1998) Experimentally induced muscle pain induces hypoalgesia in heterotopic deep tissues, but not in homotopic deep tissues. Brain Res 787:203–210
Graven-Nielsen T, Arendt-Nielsen L, Mense S (2002) Thermosensitivity of muscle: high-intensity thermal stimulation of muscle tissue induces muscle pain in humans. J Physiol 540:647–656
Grech R, Cassar T, Muscat J, Camilleri KP et al (2008) Review on solving the inverse problem in EEG source analysis. J Neuroeng Rehabil 5:25
Gustorff B, Anzenhofer S, Sycha T, Lehr S et al (2004) The sunburn pain model: the stability of primary and secondary hyperalgesia over 10 hours in a crossover setting. Anesth Analg 98:173–177
Hagenouw RR, Bridenbaugh PO, van Egmond J, Stuebing R (1986) Tourniquet pain: a volunteer study. Anesth Analg 65:1175–1180
Hampson JP, Reed BD, Clauw DJ, Bhavsar R et al (2013) Augmented central pain processing in vulvodynia. J Pain 14:579–589
Handwerker HO, Kobal G (1993) Psychophysiology of experimentally induced pain. Physiol Rev 73:639–671
Hardy JD, Wolff HG, Goodell H (1940) Studies on pain. A new method for measuring pain threshold: observations on spatial summation of pain. J Clin Invest 19:649–657
Hatem S, Attal N, Willer JC, Bouhassira D (2006) Psychophysical study of the effects of topical application of menthol in healthy volunteers. Pain 122:190–196
Hauck M, Domnick C, Lorenz J, Gerloff C et al (2015) Top-down and bottom-up modulation of pain-induced oscillations. Front Hum Neurosci 9:375
Hay JL, Okkerse P, van Amerongen G, Groeneveld GJ (2016) Determining pain detection and tolerance thresholds using an integrated, multi-modal pain task battery. J Vis Exp 14:110
Heinricher MM, Fields L (2013) Central nervous system mechanisms of pain modulation. In: McMahon SB, Koltzenburg M, Tracey I, Turk DC (eds) Wall and Melzack's textbook of pain, 6th edn. Elsevier, Philidelphia, pp 129–142
Hernández N, Dmitrieva N, Vanegas H (1994) Medullary on-cell activity during tail-flick inhibition produced by heterotopic noxious stimulation. Pain 58:393–401
Hertel HC, Howaldt B, Mense S (1976) Responses of group IV and group III muscle afferents to thermal stimuli. Brain Res 113:201–205
Hu L, Zhang ZG, Mouraux A, Iannetti GD (2015) Multiple linear regression to estimate time-frequency electrophysiological responses in single trials. NeuroImage 111:442–453
Huettel SA, Song AW, McCarthy G (eds) (2014) Functional magnetic resonance imaging, 2nd edn. Sinauer Associates Inc, Sunderland
Hüllemann P, Watfeh R, Shao YQ, Nerdal A (2015) Peripheral sensitization reduces laser-evoked potential habituation. Neurophysiol Clin 45:457–467
Ikeda H, Heinke B, Ruscheweyh R, Sandkühler J (2003) Synaptic plasticity in spinal lamina I projection neurons that mediate hyperalgesia. Science 299:1237–1240
Ilkjaer S, Petersen KL, Brennum J, Wernberg M et al (1996) Effect of systemic N-methyl-D-aspartate receptor antagonist (ketamine) on primary and secondary hyperalgesia in humans. Br J Anaesth 76:829–834
Ing Lorenzini K, Besson M, Daali Y, Salomon D et al (2012) Validation of the simplified UVB model to assess the pharmacodynamics of analgesics in healthy human volunteers. Chimia (Aarau) 66:296–299
Ingvar M (1999) Pain and functional imaging. Philos Trans R Soc Lond Ser B Biol Sci 354:1347–1358
Inui K, Kakigi R (2012) Pain perception in humans: use of intraepidermal electrical stimulation: figure 1. J Neurol Neurosurg Psychiatry 83:551–556
Jobert M, Wilson FJ, Roth T, Ruigt GSF (2013) Guidelines for the recording and evaluation of pharmaco-sleep studies in man: the international pharmaco-EEG society (IPEG). Neuropsychobiology 67:127–167
Johnson AC, Greenwood-Van Meerveld B (2016) The pharmacology of visceral pain. Adv Pharmacol 75:273–301
Jones SF, McQuay HJ, Moore RA, Hand CW (1988) Morphine and ibuprofen compared using the cold pressor test. Pain 34:117–122
Kakigi R (1994) Diffuse noxious inhibitory control. Reappraisal by pain-related somatosensory evoked potentials following CO2 laser stimulation. J Neurol Sci 125:198–205
Kakigi R, Inui K, Tamura Y (2005) Electrophysiological studies on human pain perception. Clin Neurophysiol 116:743–763

Kern D, Plantevin F, Bouhassira D (2008) Effects of morphine on the experimental illusion of pain produced by a thermal grill. Pain 139:653–659

Kidd BL, Urban LA (2001) Mechanisms of inflammatory pain. Br J Anaesth 87:3–11

Kilo S, Schmelz M, Koltzenburg M, Handwerker HO (1994) Different patterns of hyperalgesia induced by experimental inflammation in human skin. Brain 117:385–396

Kissin I (2010) The development of new analgesics over the past 50 years: a lack of real breakthrough drugs. Anesth Analg 110:780–789

Klein T, Magerl W, Hopf HC, Sandkühler J (2004) Perceptual correlates of nociceptive long-term potentiation and long-term depression in humans. J Neurosci 24:964–971

Klein T, Stahn S, Magerl W, Treede RD (2008) The role of heterosynaptic facilitation in long-term potentiation (LTP) of human pain sensation. Pain 139:507–519

Kocyigit F, Akalin E, Gezer NS, Orbay O et al (2012) Functional magnetic resonance imaging of the effects of low-frequency transcutaneous electrical nerve stimulation on central pain modulation. A double-blind, placebo-controlled trial. Clin J Pain 28:581–588

Koltzenburg M, Lundberg LE, Torebjörk HE (1992) Dynamic and static components of mechanical hyperalgesia in human hairy skin. Pain 51:207–219

Koppert W, Dern SK, Sittl R, Albrecht S et al (2001) A new model of electrically evoked pain and hyperalgesia in human skin: the effects of intravenous alfentanil, S(+)-ketamine, and lidocaine. Anesthesiology 95:395–402

Korotkov A, Ljubisavljevic M, Thunberg J, Kataeva G et al (2002) Changes in human regional cerebral blood flow following hypertonic saline induced experimental muscle pain: a positron emission tomography study. Neurosci Lett 335:119–123

Krarup AL, Gunnarsson J, Brun J, Poulakis A et al (2013) Exploration of the effects of gender and mild esophagitis on esophageal pain thresholds in the normal and sensitized state of asymptomatic young volunteers. Neurogastroenterol Motil 25:766–e580

Kumar K, Railton C, Tawfic Q (2016) Tourniquet application during anesthesia: "what we need to know?". J Anaesthesiol Clin Pharmacol 32:424–430

van Laarhoven AI, Kraaimaat FW, Wilder-Smith OH, van Riel PL, van de Kerkhof PC, Evers AW (2013) Sensitivity to itch and pain in patients with psoriasis and rheumatoid arthritis. Exp Dermatol 22:530–534

Laird JM, Bennett GJ (1993) An electrophysiological study of dorsal horn neurons in the spinal cord of rats with an experimental peripheral neuropathy. J Neurophysiol 69:2072–2085

Lapotka M, Ruz M, Salamanca Ballesteros A, Ocon Hernandez O (2017) Cold pressor gel test: a safe alternative to the cold pressor test in fMRI. Magn Reson Med 78:1464–1468

Latremoliere A, Woolf CJ (2009) Central sensitization: a generator of pain hypersensitivity by central neural plasticity. J Pain 10:895–926

Laursen RJ, Graven-Nielsen T, Jensen TS, Arendt-Nielsen L (1997) Referred pain is dependent on sensory input from the periphery: a psychophysical study. Eur J Pain 1:261–269

Lautenbacher S, Roscher S, Strian F (2002) Inhibitory effects do not depend on the subjective experience of pain during heterotopic noxious conditioning stimulation (HNCS): a contribution to the psychophysics of pain inhibition. Eur J Pain 6:365–374

Le Bars D, Dickenson AH, Besson JM (1979) Diffuse noxious inhibitory controls (DNIC). I. Effects on dorsal horn convergent neurones in the rat. Pain 6:283–304

Lee YS, Kho HS, Kim YK, Chung SC (2007) Influence of topical capsaicin on facial sensitivity in response to experimental pain. J Oral Rehabil 34:9–14

Li H (2017) TRP channel classification. Adv Exp Med Biol 976:1–8

Lötsch J, Angst MS (2003) The μ-opioid agonist remifentanil attenuates hyperalgesia evoked by blunt and punctuated stimuli with different potency: a pharmacological evaluation of the freeze lesion in humans. Pain 102:151–161

Lötsch J, Oertel BG, Ultsch A (2014) Human models of pain for the prediction of clinical analgesia. Pain 155:2014–2021

Louvel D, Delvaux M, Staumont G, Camman F et al (1996) Intracolonic injection of glycerol: a model for abdominal pain in irritable bowel syndrome? Gastroenterology 110:351–361

Macfarlane GJ, McBetch J, Jones GT (2013) Epidemiology of pain. In: McMahon SB, Koltzenburg M, Tracey I, Turk DC (eds) Wall and Melzack's textbook of pain, 6th edn. Elsevier, Philidelphia, pp 232–247

Maggi CA (1990) The dual function of capsaicin-sensitive sensory nerves in the bladder and urethra. Ciba Found Symp 151:77–90

Mauderli AP, Vierck CJ Jr, Cannon RL, Rodrigues A et al (2003) Relationships between skin temperature and temporal summation of heat and cold pain. J Neurophysiol 90:100–109

McQuay HJ, Moore A (2013) Methods of therapeutic trials. In: McMahon SB, Koltzenburg M, Tracey I, Turk DC (eds) Wall and Melzack's textbook of pain, 6th edn. Elsevier, Philidelphia, pp 402–412

Melzack R (1975) The McGill pain questionnaire: major properties and scoring methods. Pain 1:277–299

Melzack R (2005) The McGill pain questionnaire: from description to measurement. Anesthesiology 103:199–202

Melzack R, Katz J (2013) Pain measurement in adult patients. In: McMahon SB, Koltzenburg M, Tracey I, Turk DC (eds) Wall and Melzack's textbook of pain, 6th edn. Elsevier, Philidelphia, pp 301–314

Messeguer A, Plannells-Cases P, Ferrer-Montiel A (2006) Physiology and pharmacology of the vanilloid receptor. Curr Neuropharmacol 4:1–15

Meyer RA, Treede RD (2004) Mechanisms of secondary hyperalgesia: a role for myelinated nociceptors in punctate hyperalgesia. In: Brune K, Handwerker HO (eds) Hyperalgesia: molecular mechanisms and clinical implications, 1st edn. IASP Press, Seatlle, pp 143–155

Mikkelsen S, Ilkjaer S, Brennum J, Borgbjerg FM et al (1999) The effect of naloxone on ketamine-induced effects on hyperalgesia and ketamine-induced side effects in humans. Anesthesiology 90:1539–1545

Mitchell LA, MacDonald RA, Brodie EE (2004) Temperature and the cold pressor test. J Pain 5:233–237

Möller KA, Johansson B, Berge OG (1998) Assessing mechanical allodynia in the rat paw with a new electronic algometer. J Neurosci Methods 84:41–47

Moore DJ, Keogh E, Crombez G, Eccleston C (2013) Methods for studying naturally occurring human pain and their analogues. Pain 154:190–199

Moore RA, Wiffen PJ, Derry S, Maguire T (2015) Non-prescription (OTC) oral analgesics for acute pain – an overview of Cochrane reviews. Cochrane Database Syst Rev 11:CD010794

Morton DL, Sandhu JS, Jones AK (2016) Brain imaging of pain: state of the art. J Pain Res 9:613–624

Mouraux A, Iannetti GD (2008) Across-trial averaging of event-related EEG responses and beyond. Magn Reson Imaging 26:1041–1054

Mouraux A, Guérit JM, Plaghki L (2003) Non-phase locked electroencephalogram (EEG) responses to CO2 laser skin stimulations may reflect central interactions between a partial differential- and C-fibre afferent volleys. Clin Neurophysiol 114:710–722

Mouraux A, Iannetti GD, Plaghki L (2010) Low intensity intra-epidermal electrical stimulation can activate Aδ-nociceptors selectively. Pain 150:199–207

Ness TJ, Gebhart GF (1990) Visceral pain: a review of experimental studies. Pain 41:167–234

Nie H, Arendt-Nielsen L, Madeleine P, Graven-Nielsen T (2006) Enhanced temporal summation of pressure pain in the trapezius muscle after delayed onset muscle soreness. Exp Brain Res 170:182–190

Niesters M, Dahan A, Swartjes M, Noppers I et al (2011) Effect of ketamine on endogenous pain modulation in healthy volunteers. Pain 152:656–663

Nir RR, Yarnitsky D (2015) Conditioned pain modulation. Curr Opin Support Palliat Care 9:131–137

Ochoa J, Mair WG (1969) The normal sural nerve in man. I. Ultrastructure and numbers of fibres and cells. Acta Neuropathol 13:197–216

Oertel BG, Lötsch J (2013) Clinical pharmacology of analgesics assessed with human experimental pain models: bridging basic and clinical research. Br J Pharmacol 168:534–553

Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci U S A 87:9868–9872

Okkerse P, van Amerongen G, de Kam ML, Stevens J et al (2017) The use of a battery of pain models to detect analgesic properties of compounds: a two-part four-way crossover study. Br J Clin Pharmacol 83:976–990

Olesen AE, Staahl C, Arendt-Nielsen L, Drewes AM (2010) Different effects of morphine and oxycodone in experimentally evoked hyperalgesia: a human translational study. Br J Clin Pharmacol 70:189–200

Olesen AE, Andresen T, Staahl C, Drewes AM (2012) Human experimental pain models for assessing the therapeutic efficacy of analgesic drugs. Pharmacol Rev 64:722–779

Olesen AE, Brock C, Sverrisdóttir E, Larsen IM et al (2014) Sensitivity of quantitative sensory models to morphine analgesia in humans. J Pain Res 7:717–726

Olofsen E, Romberg R, Bijl H, Mooren R et al (2005) Alfentanil and placebo analgesia: no sex differences detected in models of experimental pain. Anesthesiology 103:130–139

Page GD, France CR (1997) Objective evidence of decreased pain perception in normotensives at risk for hypertension. Pain 73:173–180

Pedersen JL, Kehlet H (1998) Hyperalgesia in a human model of acute inflammatory pain: a methodological study. Pain 74:139–151

Petty BG, Cornblath DR, Adornato BT, Chaudhry V et al (1994) The effect of systemically administered recombinant human nerve growth factor in healthy human subjects. Ann Neurol 36:244–224

Peyron R, Laurent B, García-Larrea L (2000) Functional imaging of brain responses to pain. A review and meta-analysis (2000). Neurophysiol Clin 30:263–288

Pfau DB, Klein T, Putzer D, Pogatzki-Zahn EM et al (2011) Analysis of hyperalgesia time courses in humans after painful electrical high-frequency stimulation identifies a possible transition from early to late LTP-like pain plasticity. Pain 152:1532–1539

Plaghki L, Mouraux A (2003) How do we selectively activate skin nociceptors with a high power infrared laser? Physiology and biophysics of laser stimulation. Neurophysiol Clin 33:269–277

Polianskis R, Graven-Nielsen T, Arendt-Nielsen L (2001) Computer-controlled pneumatic pressure algometry–a new technique for quantitative sensory testing. Eur J Pain 5:267–277

Popescu A, LeResche L, Truelove EL, Drangsholt MT (2010) Gender differences in pain modulation by diffuse noxious inhibitory controls: a systematic review. Pain 150:309–318

Porro CA (2003) Functional imaging and pain: behavior, perception, and modulation. Neuroscientist 9:354–369

Price DD (1996) Selective activation of A-delta and C nociceptive afferents by different parameters of nociceptive heat stimulation: a tool for analysis of central mechanisms of pain. Pain 68:1–3

Price DD (2000) Psychological and neural mechanisms of the affective dimension of pain. Science 288:1769–1772

Pud D, Granovsky Y, Yarnitsky D (2009) The methodology of experimentally induced diffuse noxious inhibitory control (DNIC)-like effect in humans. Pain 144:16–19

Rainville P (2002) Brain mechanisms of pain affect and pain modulation. Curr Opin Neurobiol 12:195–204

Reddy H, Arendt-Nielsen L, Staahl C, Pedersen J et al (2005) Gender differences in pain and biomechanical responses after acid sensitization of the human esophagus. Dig Dis Sci 50:2050–2058

Roberts K, Papadaki A, Goncalves C, Tighe M et al (2008) Contact heat evoked potentials using simultaneous Eeg

and Fmri and their correlation with evoked pain. BMC Anesthesiol 8:8

Roberts K, Shenoy R, Anand P (2011) A novel human volunteer pain model using contact heat evoked potentials (CHEP) following topical skin application of transient receptor potential agonists capsaicin, menthol and cinnamaldehyde. J Clin Neurosci 18:926–932

Rosendal L, Larsson B, Kristiansen J, Peolsson M (2004) Increase in muscle nociceptive substances and anaerobic metabolism in patients with trapezius myalgia: microdialysis in rest and during exercise. Pain 112:324–334

Rukwied R, Mayer A, Kluschina O, Obreja O (2010) NGF induces non-inflammatory localized and lasting mechanical and thermal hypersensitivity in human skin. Pain 148:407–413

Ruscheweyh R, Weinges F, Schiffer M, Bäumler M (2015) Control over spinal nociception as quantified by the nociceptive flexor reflex (RIII reflex) can be achieved under feedback of the RIII reflex. Eur J Pain 19:480–489

Sandkühler J (2009) Models and mechanisms of hyperalgesia and allodynia. Physiol Rev 89:707–758

Sandrini G, Alfonsi E, Ruiz L, Livieri C et al (1989) Age-related changes in excitability of nociceptive flexion reflex. An electrophysiological study in school-age children and young adults. Funct Neurol 4:53–58

Sayre RM, Desrochers DL, Wilson CJ, Marlowe E (1981) Skin type, minimal erythema dose (MED), and sunlight acclimatization. J Am Acad Dermatol 5:439–443

Schaffler K, Nicolas LB, Borta A, Brand T et al (2017) Investigation of the predictive validity of laser-EPs in normal, UVB-inflamed and capsaicin-irritated skin with four analgesic compounds in healthy volunteers. Br J Clin Pharmacol 83:1424–1435

Schouenborg J, Weng HR, Kalliomäki J, Holmberg H (1995) A survey of spinal dorsal horn neurones encoding the spatial organization of withdrawal reflexes in the rat. Exp Brain Res 106:19–27

Schulte H, Segerdahl M, Graven-Nielsen T, Grass S (2006) Reduction of human experimental muscle pain by alfentanil and morphine. Eur J Pain 10:733–741

Sengupta J, Gebhart G (1994) Gastrointestinal afferent fibers and sensation. In: Johnson LR (ed.) Physiology of the Gastrointestinal Tract, 3rd edn. Raven, New York, pp 483–519

Serra J, Campero M, Ochoa J (1998) Flare and hyperalgesia after intradermal capsaicin injection in human skin. J Neurophysiol 80:2801–2810

Shukla S, Torossain A, Duann JR, Keung A (2011) The analgesic effect of electroacupunture on acute thermal pain perception – a central neural correlate study with fMRI. Mol Pain 7:45–56

Simone DA, Ngeow JYF, Putterman GJ, LaMotte RH (1987) Hyperalgesia to heat after intradermal injection of capsaicin. Brain Res 418:201–203

Skljarevski V, Ramadan NM (2002) The nociceptive flexion reflex in humans – review article. Pain 96:3–8

Smith GM, Egbert LD, Markowitz RA, Mosteller F (1966) A VAS is used to assess the subject's pain intensity. J Pharmacol Exp Ther 154:324–332

Song J, Davey C, Poulsen C, Luu P et al (2015) EEG source localization: sensor density and head surface coverage. J Neurosci Methods 256:9–12

Staahl C, Drewes AM (2004) Experimental human pain models: a review of standardised methods for preclinical testing of analgesics. Basic Clin Pharmacol Toxicol 95:97–111

Staahl C, Christrup LL, Andersen SD, Arendt-Nielsen L et al (2006) A comparative study of oxycodone and morphine in a multi-modal, tissue-differentiated experimental pain model. Pain 123:28–36

Staahl C, Olesen AE, Andresen T, Arendt-Nielsen L et al (2009a) Assessing analgesic actions of opioids by experimental pain models in healthy volunteers – an updated review. Br J Clin Pharmacol 68:149–168

Staahl C, Olesen AE, Andresen T, Arendt-Nielsen L et al (2009b) Assessing efficacy of non-opioid analgesics in experimental pain models in healthy volunteers: an updated review. Br J Clin Pharmacol 68:322–341

Svendsen O, Edwards CN, Lauritzen B, Rasmussen AD (2005) Intramuscular injection of hypertonic saline: in vitro and in vivo muscle tissue toxicity and spinal neurone c-fos expression. Basic Clin Pharmacol Toxicol 97:52–57

Svenson P, Arendt-Nielsen L (1995) Induction and assessment of experimental muscle pain. J Electromyogr Kinesiol 5:131–140

Svensson P, Cairns BE, Wang K, Arendt-Nielsen L (2003) Injection of nerve growth factor into human masseter muscle evokes long-lasting mechanical allodynia and hyperalgesia. Pain 104:241–247

Svensson P, Wang K, Arendt-Nielsen L, Cairns BE (2008) Effects of NGF-induced muscle sensitization on proprioception and nociception. Exp Brain Res 189:1–10

Thalhammer JG, LaMotte RH (1982) Spatial properties of nociceptor sensitization following heat injury of the skin. Brain Res 231:257–265

Torebjörk HE, Lundberg LE, LaMotte RH (1992) Central changes in processing of mechanoreceptive input in capsaicin-induced secondary hyperalgesia in humans. J Physiol 448:765–780

Tracey I, Mantyh PW (2007) The cerebral signature for pain perception and its modulation. Neuron 55:377–391

Treede RD, Lorenz J, Baumgärtner U (2003) Clinical usefulness of laser-evoked potentials. Neurophysiol Clin 33:303–314

Tuveson B, Leffler AS, Hansson P (2006) Time dependent differences in pain sensitivity during unilateral ischemic pain provocation in healthy volunteers. Eur J Pain 10:225–232

Vo L, Hood S, Drummond PD (2016) Involvement of opioid receptors and α2-adrenoceptors in inhibitory pain modulation processes: a double-blind placebo-controlled crossover study. J Pain 17:1164–1173

Wager TD, Atlas LY, Lindquist MA, Roy M et al (2013) An fMRI-based neurologic signature of physical pain. N Engl J Med 368:1388–1397

Willer JC (1977) Comparative study of perceived pain and nociceptive flexion reflex in man. Pain 3:69–80
Woolf CJ (1983) Evidence for a central component of post-injury pain hypersensitivity. Nature 306:686–688
Woolf CJ (2011) Central sensitization: implications for the diagnosis and treatment of pain. Pain 152:S2–15
Woolf CJ, Max MB (2001) Mechanism-based pain diagnosis: issues for analgesic drug development. Anesthesiology 95:241–249
Yang H, Meijer HGE, Doll RJ, Buitenweg JR (2015) Computational modeling of Adelta-fiber-mediated nociceptive detection of electrocutaneous stimulation. Biol Cybern 109:479–491
Yarnitsky D, Arendt-Nielsen L, Bouhassira D, Edwards RR et al (2010) Recommendations on terminology and practice of psychophysical DNIC testing. Eur J Pain 14:339

Pharmacodynamic Evaluation: Drug Dependency and Addiction

6

V. Tenev and M. Nikolova

Contents

V. Tenev (✉)
Department of Psychiatry, University Hospital, University of Iowa, Iowa city, IA, USA
e-mail: vesselin-tenev@uiowa.edu

M. Nikolova
University Hospital Alexandrovska, Clinic of Nephrology, Medical University, Sofia, Bulgaria
e-mail: milena_i_dani@abv.bg

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_49

Abstract

The intake of psychoactive substances has accompanied mankind since the dawn of human race. Different psychoactive substances have been extracted and synthesized in human history, and their pharmacodynamic properties have been studied thoroughly. This chapter presents the current knowledge on the pharmacodynamic profile of the most common illicit and recreational drugs and their influence on drug dependency and addiction. The substances with abuse potential grow exponentially. Our clinical and practical knowledge has still a long way to go in catching up with these realities. The dynamic interactions between different drugs observed in vitro cannot be fully replicated in vivo. We rely on randomized clinical trials, case reports, and own clinical experience with patients. Different clinical scenarios could provide further evidence and hypotheses regarding the sought and adverse effects of substances, their interactions with legal drugs and medications, and their impact on different stages of metabolism. In this chapter we attempted to summarize the available reliable data and suggest some ideas for future observation and research.

Introduction

Pharmacodynamics is a branch of pharmacology that studies the molecular, biochemical, and physiological effects and the mechanisms of action of a substance (i.e., a drug) on the human body (Campbell and Cohall 2017). All substances that affect the human body either influence normal biochemical or physiological processes or inhibit the vital processes of an "invader" to the body – microorganism or parasite. On molecular level, overall seven major mechanisms of drug action have been described in the human body:

- Stimulation/activation of receptor systems (agonism, e.g., beta-agonists in asthma)
- Inhibition/depression of receptor systems (antagonism, e.g., calcium channel blockers in hypertension)
- Blocking of receptors without further activation or inhibition ("silent" antagonism, e.g., naloxone in opioid intoxication)
- Stabilization of receptors (e.g., buprenorphine in opioid dependency)
- Exchange of substances (e.g., digitalis glycosides, anesthetics, etc.)
- Initiation/activation of beneficial chemical reactions (acetyl cysteine as initiator of free radical scavenging)
- Initiation/activation of harmful chemical reactions (e.g., cytotoxic treatment)

Pharmacodynamics of a substance comprises of three major types of processes: (1) binding to structures (receptors) in the body and resulting in desired and undesirable (adverse) effects, (2) post-receptor effects, and (3) interactions with other substances within the body. The binding to certain structures within the body (receptors, membrane structures, proteins, enzymes, ion pimps, etc.) leads to further molecular, biochemical, and physiological effects. The difference between the doses leading to desired effects and the one that leads to adverse events is the therapeutic window of a drug. The duration of action of a drug is the length of time that the drug remains effective. From pharmacodynamic point of view, the receptor binding, the therapeutic and adverse effects, the therapeutic window, and the duration of action in drug dependency and addiction depend mainly on the receptor target, the properties of the drug itself and the dose taken, and the concentration of the drug at the receptor site and on certain physiological changes in the body (aging, intake of other substances/drugs, genetic polymorphisms

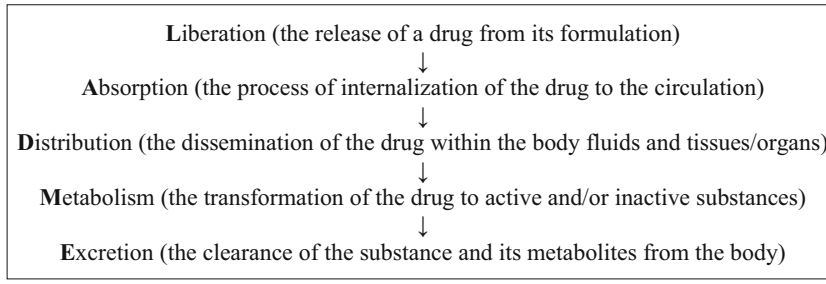


Fig. 1 LADME scheme of a pharmacokinetic profile of a drug

and conditions, metabolic disturbances [thyrotoxicosis, malnutrition, renal or hepatic failure, dyselectrolytemia, etc.]). Moreover, the concomitant abuse of several substances, including alcohol and prescription medications, could potentiate the effects of illicit drugs due to pharmacodynamic (and pharmacokinetic) interactions. All these factors alter the pharmacodynamics of illicit and recreational drugs and modify the profile of dependency, addiction, and withdrawal.

On the other hand, pharmacokinetics studies the effects of the body on the drug: drug absorption, distribution, metabolism, and excretion (Ruiz-Garcia et al. 2008). The pharmacokinetic profile of a drug can be presented schematically in the so-called LADME sequence (Ruiz-Garcia et al. 2008) (Fig. 1). Sometimes the terms metabolism and excretion are grouped together in the term "elimination."

In terms of drug abuse and dependence, pharmacodynamics and pharmacokinetics are often referred to as toxodynamics and toxokinetics because of the toxic effects of illicit drugs and the high rate of adverse and toxic reactions in this patients' population.

The pharmacodynamic and the pharmacokinetic profiles of illicit and recreational drugs define their effects, adverse reactions, addiction, and withdrawal symptoms. In general, the illicit drugs with more rapid absorption and entry into the circulation and the central nervous system, higher bioavailability, shorter half-life, high free drug levels, smaller volume of distribution, and higher clearance rate are more toxic and tend to cause higher rate of addiction and more severe withdrawal symptoms. Most of the drug users tend to adapt the route of administration, the dose and the additives, and/or the coadministered illicit drugs with additive/synergistic effects to their individual cravings in order to produce maximum drug effect for maximum time.

Before we start discussing the pharmacodynamics of addictive substances, we have to answer the following important questions: Which substances actually make us feel happy, and which parts of the brain give the signals of happiness? The substances that make us feel happy are physiological mediators in the brain that are secreted in response to a stimulus giving us the sensation of comfort or reward. These are the endorphins, serotonin, dopamine, and oxytocin. The physiological sites where these stimuli act are the limbic system (and particularly nucleus accumbens), the memory/experience part of the brain (hippocampus and amygdala), and the cortex (the frontal and prefrontal areas that supervise the first two parts stated) (Powledge 1999; Volkow and Morales 2015). This "reward pathway" in the brain is a very ancient dopaminergic pathway that was present long before humans in the brain of mammals. It plays crucial role for the motivation of behavior. It starts in the midbrain (in the ventral tegmental area) and extends to nucleus accumbens, hippocampus, amygdala, and the prefrontal and frontal cortical areas that are meant to inhibit all the structures before them in the pathway (Powledge 1999; Volkow and Morales 2015). After disinhibition of the subcortical structures, a vicious circle of constant "reward" stimulation is closed ("the reward cycle").Virtually all illicit drugs follow this pathway of addiction, along with nicotine, caffeine,

and alcohol (Powledge 1999). This neuromediator pathway is stimulated not only by natural stimuli (success, victory, self-content from the achieved, etc.) but by stimulant medications and addictive types of self-destructive behavior, including eating disorders and gambling (Powledge 1999; Volkow and Morales 2015).

All addictive substances tend to bind specific receptors in the brain that lead to liberation of serotonin, dopamine/norepinephrine, and/or oxytocin and therefore to imitate the state of comfort and/or the reward ensured by other natural stimuli in our everyday life, but the artificially induced state of happiness and/or excitement is more intense. Still, the administration of recreational drugs leads to structural and functional changes in neurons, called neuroplasticity. These adaptive changes alter the drug effect and metabolism and generate the need for more frequent administration in higher doses, which is called tolerance with further dependence and addiction. The sudden cease in drug intake leads to withdrawal symptoms. Withdrawal symptoms are often mediated by dopaminergic pathways and/or by extrahypothalamic corticotropin-releasing factor (CRF) system – release of CRF outside the hypothalamus (e.g., from the amygdala). This could explain the common withdrawal symptoms (sweating, changes in blood pressure and heart rate, abnormal peristaltics, joint pains, headache, etc.) for different illicit drugs, including opiates, stimulants, alcohol, nicotine, cannabinoids, etc.

In other words, the majority of psychoactive substances tend to bind specific receptor in the central nervous system (CNS) and to mimic the effects of endogenic substances with the effect being more potent and with longer duration: opioid, benzodiazepine/gamma-aminobutyric acid (GABA), serotonin, dopamine, and cannabinoid receptors (Quinn et al. 1997; Sharma et al. 2012). Moreover, these receptors are known to interact and to lead to a neurochemical correlation between substances abused within the brain (Quinn et al. 1997). Stimulant drugs (cocaine, amphetamines, and amphetamine derivatives) act by causing an increase in dopamine levels within the synaptic cleft – by facilitating dopamine release and inhibiting dopamine reuptake (Nestler 2005; Calipari and Ferris 2013). Cocaine also blocks sodium channels on cell membranes and has anesthetic effect (Volkow and Morales 2015). Nicotine activates *N*-cholinergic receptors and is known to stimulate catecholamine (dopamine, norepinephrine), glutamate, serotonin, acetylcholine, endorphin, and GABA release (Quinn et al. 1997; Benowitz 2009). Moreover, its primary metabolite, cotinine, is known to increase serotonin levels in the brain (Quinn et al. 1997). Only alcohol has no specific receptors in the brain and acts by altering the physiological properties of lipid membranes, modifying their fluidity and changing the receptor sensitivity to natural stimulant, and inhibiting neurotransmitters (Quinn et al. 1997).

As it was stated above, the changes in the human body in response to the intake of substances, especially in receptor systems and signaling pathways in the brain, are referred to as plasticity (respectively, neuroplasticity). These changes are responsible for the development of tolerance, i.e., the need for more frequent administration of the addictive substance and in higher doses. Once this need evolves to imperative urge, an addiction has developed. From pharmacodynamic point of view, the major mechanisms underlying neuroplasticity are as follows: changes in receptor structure, type, distribution, and functionality, changes in signaling pathways, development of tolerance and cross-tolerance, involvement of other receptor pathways due to the cross-reaction between receptor systems, and development of new pathways for signal channeling (Dumas and Pollack 2008).

Moreover, all addictive substances used for recreational purposes are known to cause permanent structural alteration in cells due to epigenetic and genetic effects, including microtubular toxicity, chromothripsis, genotoxicity, oncogenesis and embryo-/fetotoxicity, inhibition of tumor-suppressor genes (e.g., p53 by marijuana smoke) and activation of proto-oncogenes, etc. (Reece and Hulse 2016). The major mechanism behind the inheritable genetic abnormalities in illicit substance abuse is thought to be the process of chromothripsis – extensive genomic rearrangements and an oscillating pattern of DNA copy number levels due to microtubular

damage and changes in the mitotic spindle. These alterations are curiously restricted to one or a few chromosomes. All illicit drugs are known to cause changes in sister chromatid exchange levels, in the mitotic spindle, in DNA fragmentation, and in gene systems that regulate the cell processes (including growth and development) (Dumas and Pollack 2008; Reece and Hulse 2016; Li and Lin 1998; Reece 2009). Having in mind the wider prevalence of illicit drug intake worldwide, the possibilities for transgenerational genotoxicity (terato- and oncogenicity and toxicity) raise serious and increasing concern (Benowitz 2009).

A very important aspect of both pharmacodynamics and pharmacokinetics of illicit drugs are the drug interactions that determine the potentiation or inhibition of effect and the possibilities to influence withdrawal and cessation of illicit drug abuse by the administration of their analogues or medications that block their action or ameliorate abstinence symptoms.

Pharmacodynamic profiles of commonly abused drugs and their significance for the treatment of withdrawal and addiction.

Opioids

The term "opiate" refers to a substance derived from opium, i.e., the alkaloids found in the plant *Papaver somniferum* (opium poppy). Three main psychoactive compounds are isolated from this plant – morphine, codeine, and thebaine – along with several alkaloids that lack psychoactive properties and have only spasmogesic effect (papaverine, noscapine, and about 24 more substances). Other morphine-like substances that have been isolated in small amounts from the opium poppy are dihydrocodeine, metopon, oxycodone, and oxymorphol.

The term "opioids," on the other hand, includes a large group of substances that interact with the opioid receptors in a morphine-like way, producing analgesic, anesthetic, and psychoactive effects. Opioids have been familiar to humans for thousands of years for their analgesic and psychoactive properties. These substances are widely used for recreational purposes, including their euphoric, hallucinogenic, and other psychoactive effects. In 2013 up to 0.8% of the population aged 15–65 years worldwide reported using opioids for recreational purposes (Status and Trend Analysis of Illicit Drug Markets 2015). Their rewarding effects, explained by activation of dopaminergic pathways in the "reward cycle" of the brain (including parts of the limbic system), are the main cause of opioid abuse, addiction, and dependence. Severe withdrawal symptoms develop in abrupt drug cessation; therefore proactive treatment is needed.

According to their presence in nature, opioids are classified in several groups (Ghelardini et al. 2015; Koob and Le Moal 2006):

- Natural – morphine, codeine, thebaine, and salvinorin A (kappa-agonist)
- Morphine esters – morphine diacetate (heroin), morphine dinicotinate, morphine dipropionate, etc.
- Semisynthetic (created from natural opiates or their esters) – hydromorphone, hydrocodone, oxymorphone, oxycodone, buprenorphine, ethylmorphine, etc.
- Synthetic – fentanyl, methadone, tramadol, tapentadol, dextropropoxyphene, pethidine, levorphanol, etc.
- Endogenic – endorphins, enkephalins, dynorphins, and endomorphins

The adverse effects of opioid abuse include cognitive impairment, gastrointestinal symptoms (constipation, nausea, vomiting), hypotension, sexual dysfunction, and respiratory and cardiovascular center depression.

Three major types of opioid receptors have been identified and cloned: mu, delta, and kappa. An additional opioid substance binding type of receptor is the ORL-1 (opioid receptor-like 1, or nociceptin receptor). Three additional types of opioid ligand binding receptors have been discovered – zeta and epsilon opioid receptors and sigma receptors. All types of opioid receptors have different distributions and physiological roles (Ghelardini et al. 2015; Koob and Le Moal 2006; Stein et al. 2003; Gosnell et al. 2013):

- Mu: distributed in the brain (cortex, thalamus, striosomes, periaqueductal gray matter, rostral ventromedial medulla), in the spinal cord, in the peripheral sensory neurons, and in the gastrointestinal tract and other peripheral structures. Subtypes: mu1 (analgesia, dependence), mu2 (euphoria, miosis, respiratory and gastrointestinal motility depression, physical dependence), and mu3 (vasodilation?).
- Kappa: distributed in the brain (hypothalamus, periaqueductal gray matter, claustrum), spinal cord, and peripheral sensory neurons. Subtypes: kappa1, kappa2, and kappa3, responsible for analgesic, depressive, hallucinogenic, miotic, diuretic, dysphoric, neurodepressive, sedative, and neuroprotective effects.
- Delta: distributed in the brain (deep cortex, pontine nuclei, amygdala, olfactory bulb) and peripheral sensory neurons. Subtypes: delta 1 and delta 2, responsible for analgesic and antidepressant effects, convulsogenic properties, and dependence.
- ORL-1 (nociceptin receptor): distributed in the brain (amygdala, hypothalamus, hippocampus, cortex, septal nuclei) and the spinal cord. Responsible for anxiety, depression, appetite changes, and dependence to mu-agonists and affects both pain and reward signaling within the brain.
- Epsilon (binding beta-endorphin): distributed in the brain and peripheral sensory neurons, probably a splice variant or a heteromer of existing opioid receptors, antagonized by buprenorphine. Responsible for analgesic effect and for the release of met-enkephalin.
- Sigma: referred to as antitussive receptors, binding 4-phenyl-1-(4-phenylbutyl) piperidine and other substances (including dextromethorphan, phencyclidine, cocaine and methamphetamine, morphine and diacetyl morphine, fluvoxamine, dimethyltryptamine, etc.). Known to interact with kappa-opioid and NMDA glutamate receptors. Known two subtypes – sigma1 and sigma2 (sigma1 having no structural similarity to the opioid receptors). Their activations mimic acute stress reactions – tachycardia, mydriasis, overall stimulation, antitussive effect, and euphoria/dysphoria. Sigma-receptors bind to several hormones – dehydroepiandrosterone and gestagens.
- Zeta (opioid growth factor receptor): distributed in peripheral tissues (parenchymal organs – heart, liver, kidney, brain, pancreas, fat tissue, and skeletal muscles). Responsible for tissue growth, embryonic/fetal growth, wound healing, and development and cancer proliferation. The activation of these receptors decreases cell proliferation (i.e., acts as "negative" growth factor).

Opioid receptors are abundant in all tissues and organs, including the brain, peripheral nerves, gastrointestinal and immune system, endocrine glands, and skin, where they have different analgesic and non-analgesic physiological effects, as described above. All opioid receptors represent G protein-coupled receptors acting via changes (decrease) in adenylate cyclase activity and cAMP levels, protein kinase activity, CREB protein, and calcium and potassium ion transport. Moreover, the activation of opioid receptors leads to changes in substance P and GABAergic, glutamatergic, and dopaminergic transmission, leading to decrease in pain sensation and psychoactive properties, including activation of the reward cycle and euphoria (Quinn et al. 1997; Ghelardini et al. 2015; Koob and Le Moal 2006; Stein et al. 2003; Gosnell et al. 2013; Pasternak and Pan 2013). Tramadol and tapentadol also affect monoamine uptake (Quinn et al. 1997; Ghelardini et al. 2015; Koob and Le Moal 2006; Stein et al. 2003; Gosnell et al. 2013; Pasternak and Pan 2013). Opioid agonists (mu, kappa, and delta) are known to interact with oxytocin, neuropeptide Y, and melanocyte-stimulating hormone signaling systems (Gosnell et al. 2013; Pasternak and Pan 2013), which could explain their effects on feeding and appetite.

According to their effect on opioid receptors, the ligands can be classified as agonists, antagonists, partial agonists, and mixed agonists/antagonists:

- Agonists – bind strongly to opioid receptor and undergo strong conformational changes to exert effect: morphine, heroin, hydrocodone, hydromorphone, fentanyl, methadone, oxycodone, and oxymorphone.
- Partial agonists – bind less strongly and cause less conformational changes with less receptor activation and at low doses cause similar analgesic effects like full agonists; increasing the dose does not increase analgesic activity: buprenorphine, tramadol, and butorphanol.
- Mixed agonists/antagonists – agonists to some and antagonists to other opioid receptors and dose-dependent effect (i.e., agonists at some and antagonists on other doses): buprenorphine, butorphanol, nalbuphine, and pentazocine. For instance, buprenorphine is a partial mu-agonist and kappa-antagonist and weak delta-antagonist; butorphanol is a mu-antagonist and partial kappa-agonist, pentazocine is a partial mu-agonist and kappa-agonist, and nalbuphine is a mu-antagonist and kappa-agonist.
- Antagonists: naloxone and naltrexone.

To make the long story short, most psychoactive opioids are mu-agonists with different actions on kappa-receptors. As described above, the activation of mu-opioid receptors leads to G protein-mediated decrease of adenylate cyclase activity and inhibition of cAMP production with subsequent inhibition of calcium influx and potassium efflux with membrane hyperpolarization and analgesic effect. Moreover, these substances change the levels of substance P and GABAergic, glutamatergic, and dopaminergic transmission with suppression of pain signaling and activation of the reward cycle. In addition, many synthetic opioids have supplementary effects (i.e., inhibition of norepinephrine uptake and NMDA receptor inhibition with increased glutamate and GABA signaling), so other signal systems in the brain are also used to mediate their psychoactive effects (Koob and Le Moal 2006; Gosnell et al. 2013; Pasternak and Pan 2013).

Pharmacological Effects

Mu-receptor opioid agonists have the following pharmacological effects (Ghelardini et al. 2015; Koob and Le Moal 2006):

- Analgesic – mediated by mu-opioid receptors (at spinal and supraspinal levels). This effect is a result of complex ion- and mediator-induced changes in neuron interactions. At supraspinal level, it is a result of activation of mu-receptors on GABAergic neurons with subsequent activation of serotoninergic neurons. At spinal level, this effect is due to increase in the pain threshold and is mediated by inhibition of the release of mediators participating in the pain signaling – substance P and glutamate and nitric oxide from the nociceptive afferent neuron cells. Methadone also interacts with the mu-receptors on glutamatergic neurons and thus additionally decreases the transmission of the pain signal. Mesangial cells are also known to have opioid receptors which at least partially can explain the development of heroin-associated nephropathy.
- Psychotropic effects – these effects are mediated by the opioid receptors on structures of the limbic system, including the cortical areas, hypothalamus, locus coeruleus, and amygdala.
- Effects on respiratory functions – mediated by the opioid receptors in the brainstem, along with miosis.
- Gastrointestinal effects (decreased mobility, suppressed nausea) – via the opioid receptors on peripheral neurons and on the gastrointestinal tract.
- Respiratory effects – suppression of cough, in larger doses, and suppression of breathing.
- Endocrine effects (via hypothalamic mu-receptors with subsequent suppression of pituitary functions) – inhibition of pituitary function with decreased levels of LH, FSH, and ACTH.
- Paradoxical effects of morphine – at low doses morphine can increase the sensation of pain – hyperalgesia, probably due to activation of pronociceptive mediation via stimulation (not

inhibition) of adenylate cyclase and increase in neuron excitability. This effect is dose-dependent.

The development of tolerance and addiction is explained by several phenomena, including decrease in receptor number and affinity and internalization of receptors and changes on post-receptor level that decrease the ligand effect and lead to the need of more frequent administration of higher doses. The withdrawal symptoms of opioid dependence are very unpleasant and further increase the craving. They are mediated via changes in adrenergic and cholinergic mediation and neuropeptide Y changes in CRF receptor system (Koob and Le Moal 2006). According to their severity, these symptoms can be classified into 5 grades (from 0 to 4) (Koob and Le Moal 2006):

- Grade 0 – craving (for the drug) and anxiety
- Grade 1 – grade 0 plus yawning, increased perspiration, runny nose, and lacrimation
- Grade 2 – grade 0 and 1 with increased intensity plus sympathetic activation (mydriasis, gooseflesh with piloerection ("cold turkey detox"), marked tremor and twitches/spasms, hot and cold flushes); severe pain in the joints, bones, and muscles; and loss of appetite
- Grade 3 – all of the above, with increased intensity, plus insomnia, signs of sympathetic activation (increased blood pressure, body temperature, heart and respiratory rate, restlessness, muscle twitches), and nausea
- Grade 4 – all of the above, with increased intensity, plus vomiting, diarrhea, loss of appetite, weight loss, embryonic position, spontaneous ejaculation/orgasm, dehydration with hemoconcentration and eosinopenia, and high blood glucose

These symptoms can be alleviated with the administration of beta-blockers, sedatives, and antipsychotics, supportive treatment (hydration, parenteral feeding, gastroprotective agents, etc.), and addition of morphine analogues (Quinn et al. 1997; Pasternak and Pan 2013).

Natural Opioid-Like Substances

Several endogenic substances mimic the effects of opioid and are classified as endogenous opioids (Ghelardini et al. 2015; Pasternak and Pan 2013):

- Enkephalins (pentapeptides containing the sequence Tyr-Gly, linked to leucine or methionine and called, respectively, leu-enkephalin and met-enkephalin) – bind predominantly to kappa-receptors.
- Dynorphins A and B – bind mainly to delta-receptors.
- Endorphins – the beta-endorphins bind equally to mu- and delta-receptors.
- Endomorphin-1 and endomorphin-2 – bind mainly to mu-receptors.
- Endogenous morphine synthesis has been proven in humans and in animals, but the role of this "animal" morphine and its precursors and derivatives remains unclear.

All these substances take part in the pain and reward signaling, both central and peripheral, but their exact role in human physiology remains unclear.

Several peptides have been shown to modulate opioid action, including cholecystokinin and neuropeptide FF that reduce opioid effects (Mollereau et al. 2005) via changes in intracellular second messengers of nociception. The pronociceptive opioid analogues nociceptin and dynorphin (Mollereau et al. 2005) paradoxically are able not only to potentiate but also to attenuate the analgesic effects of opioids due to changes in pain circuit signaling.

Pharmacologic Interactions of Opioids

The epidemic of opioid prescription abuse makes it even more important to focus our clinical attention on their drug interactions. Methadone and buprenorphine, as with most of the psychoactive medications, are substrates of CYP450 3A4. The hepatic metabolism of opioids also goes through other isoenzymes from the CYP family, such as 2B6, CYP2C19, CYP2C9, and CYP2D67 for

methadone and 2C8 for buprenorphine. A classical interaction would be a strong 3A4 inhibitor (e.g., ketoconazole) increasing the plasma levels of methadone. Same interaction could be observed with an antifungal and an antibiotic – ciprofloxacin. Per FDA criteria, strong inhibition leads to fivefold and higher increase of the plasma concentration of the inhibited substrate. On the other hand, inducing strongly CYP450 system would lead to lower plasma level of its substrates. However, the correlations are not always that linear. One of the reasons is the influence of medications on the glucuronidation. Methadone inhibits glucuronidation of zidovudine, thus decreasing its elimination and increasing the risk of toxicity (McCance-Katz et al. 1998). There are several important class interactions: (1) opioids with medications treating infectious diseases (HIV, tuberculosis, Hep C, etc.), (2) opioids with psychopharmacologic agents (antidepressants, antipsychotics, benzodiazepines), and (3) opioids with alcohol or illicit substances.

Three types of consequences due to interactions:

1. **Toxicity, higher rate, and more severe side effects** – related to slowing the rate of metabolism/elimination, increasing the plasma levels of:
   (a) The concomitant drugs administered with opioids, e.g., zidovudine (lactic acidosis, transaminitis, myopathy, severe anemia or neutropenia, etc.)
   (b) Opioids, e.g., methadone (cognitive dysfunction, respiratory depression, QTc prolongation, arrhythmias) with cotreatment with azoles and ciprofloxacin *or* discontinuation of CYP450 inducers (such as carbamazepine, phenytoin, phenobarbital)
   (c) Synergistic and pharmacodynamic effects: Increased sedation, delirium, and respiratory drive (opioids with alcohol, benzodiazepines, antihistamine medications (diphenhydramine), dextromethorphan)
2. **Poor therapeutic response to concomitant drugs –** Related to increased rate of metabolism/elimination of antiretrovirals and poor efficacy. Complications: viral mutations, antiretroviral resistance, and increased risk for viral transmission (lower concentrations of didanosine and stavudine)
3. **Opioid withdrawal**
   (a) Related to increased rate of metabolism/elimination of opioids (methadone, buprenorphine), e.g., **HIV medications (**efavirenz, nelfinavir, lopinavir/ritonavir, etc.), **tuberculosis medications** (rifampin), **anticonvulsants** (carbamazepine, phenytoin, phenobarbital), and **stimulants (**cocaine (CYP3A4, P), glycoprotein inducer)
   (b) Discontinuation or lowering the dose of CYP450 inhibitors, medications which increase the plasma concentrations of opioids (fluvoxamine, fluoxetine) and antibiotics (including azoles)
   (c) Pharmacodynamic interactions – Cocaine during sublingual use of buprenorphine (vasoconstriction)

Drug Interactions for Specific Opioids

See Table 1

Interactions with Clinical Importance

Morphine delays the absorption of clopidogrel, prasugrel, and ticagrelor and enhances gabapentin pain tolerance in healthy volunteers. Quinidine can enhance the activity of opioids – morphine, fentanyl, oxycodone, codeine, dihydrocodeine, and methadone. Antimycotic medications increase the plasma concentrations of opioids – buprenorphine, fentanyl, morphine, oxycodone, methadone, tilidine, and tramadol. Protease inhibitors induce metabolism of opioids – oxycodone and fentanyl. Paroxetine inhibits the metabolism of hydrocodone, oxycodone, and tramadol. Escitalopram inhibits the metabolism of tramadol (Feng et al. 2017).

Stimulants: Cocaine and Amphetamine and Its Derivatives

All psychostimulants act by increasing monoamine (norepinephrine, dopamine, and serotonin) release in the synaptic space and by inhibiting

Table 1 Pharmacokinetic and pharmacodynamic interactions of opioids with other medications

Opioid	**Morphine**			
	↑ own concentration – **toxicity**	↑ the other agent's concentration toxicity	↓ the other agent's concentration	↓ own concentration – **withdrawal**
Absorption	Metoclopramide	P2Y12 inhibitors Gabapentin		Rifampin
Metabolism Elimination	Quinidine Itraconazole Other azoles Amantadine			Rifampin St. John's wort
	Methadone			
Absorption	Quinidine Voriconazole Ketoconazole Grapefruit juice	AZT (zidovudine) Desipramine		Rifampin
Metabolism	Delavirdine Amitriptyline Dextromethorphan Quetiapine Ciprofloxacin		Didanosine, stavudine	Darunavir Efavirenz Nelfinavir Nevirapine Lopinavir/ritonavir Carbamazepine Phenytoin Phenobarbital
	Buprenorphine			
Metabolism	Antimycotics			Carbamazepine Phenytoin Phenobarbital
	Oxycodone			
	Antimycotics Macrolides Ketolides Protease inhibitors Voriconazole Ketoconazole Grapefruit juice Paroxetine Quinidine			

their reuptake leading to increased neurotransmitter levels for a longer time in the synaptic cleft (Quinn et al. 1997).

Cocaine

Cocaine is the second most frequently used recreational drug worldwide after cannabis. It is a natural alkaloid extracted from the leaves of the coca plant (*Erythroxylum coca* var. *coca*, var. *ipadu*, var. *novogranatense*, and var. *truxillense*), growing in South America. Cocaine can be extracted from coca leaves or synthesized and used as a recreational substance, or further processed to crack cocaine – a freebase form of cocaine that can be smoked. Between 14 and 21 million people are estimated to have used this drug every year (Pomara et al. 2012).

Cocaine has been used for more than 1000 years by the indigenous South American people as a stimulant and for religious and recreational purposes in the form of *Erythroxylum coca* leaves that can be chewed or processed to extract the alkaloid. There are proofs that cocaine has been used as anesthetic in ancient times (Gay et al. 1975). In the seventeenth century when the Spanish arrived to the New World, they described

the stimulant, hunger-suppressing, anesthetic, and recreational effects of coca leaves. The alkaloid was first isolated by Friedrich Gaedcke in 1855 and was initially named erythroxyline. Approximately 40 years later, in 1989, the first synthetic cocaine appeared. The drug was initially used as a painkiller; subsequently its local anesthetic properties were used. Cocaine was found to be a unique anesthetic because unlike all other anesthetics it decreased bleeding due to its local vasoconstriction effect. In 1879 cocaine was introduced for the treatment of morphine addiction, and in the next few year, its use as a psychostimulant and appetite-suppressing drug started. In the beginning of the twentieth century, it was marketed as stimulant and was subsequently used in world wars as stimulant and anesthetic. Gradually, cocaine has become the second most abused illicit drug worldwide. It is used by all socioeconomic strata, age, and demographic, economic, social, political, and religious groups all over the world. Cocaine can be insufflated (snorted), taken orally (gingival administration and chewing coca leaves), smoked, administered rectally, and injected intravenously or intramuscularly. It can be taken alone or in combination with heroin (speedball). In modern medicine its use is limited as local/topical anesthetic, mainly in ophthalmology.

Cocaine has sympathomimetic effects, influencing serotonin receptor and membrane ion transport. Cocaine is also known to have long-term endocrine and genetic effects.

The pharmacodynamic effects of cocaine are determined by its three major actions (Quinn et al. 1997; Pomara et al. 2012; Gay et al. 1975):

- Increased release of catecholamines in the synaptic cleft due to stabilization of the dopamine transporter
- Decreased mediator reuptake via blockage of the presynaptic dopamine transporter
- Blockage of neuronal membrane sodium channels with local anesthetic effect

Additionally, cocaine interacts with serotonin 5-HT3 and 5-HT2 receptors, and these interactions explain its effect on appetite and locomotion. The effects on locomotion could also be explained by its interaction with dopamine levels in the substantia nigra.

Cocaine also interacts with kappa-opioid, sigma, D1, and NMDA receptors.

Unlike amphetamine, cocaine *does not* inhibit monoamine oxidase (MAO) (Quinn et al. 1997).

The net effect of these ligand-receptor interactions is sympathomimetic effect with buildup of dopamine in the limbic system structures (especially in the nucleus accumbens) and stimulation of pleasure and reward feeling that explains the addiction and dependence in long-term abuse (Nestler 2005). The increase of dopamine levels in the nucleus accumbens is a normal physiological process, part of the fight-or-flight response to stress, giving the body and the mind the assurance that the stress-inducing stimulus has been eliminated and generating the sensation of comfort and pleasure – i.e., when a thirsty person drinks water, or when a reward for achievement has been given (Nestler 2005). Thus, the external stimulation and the buildup of dopamine levels in the nucleus accumbens by cocaine are far more potent than the physiological effect and give the sensation of euphoria and stimulation. This is the underlying mechanism of addiction and dependence. Cocaine also exerts its dopamine buildup effects in other regions of the brain, associated with the limbic system, including memory centers (hippocampus and amygdala) and the frontal cortex. It is believed that the repeated exposure to cocaine with increase in dopamine availability in the hippocampus and amygdala leads to functional and organic changes that every memory of cocaine intake urges an almost compulsory craving for repeated intake (Volkow and Morales 2015). The repeated increase in dopamine levels in the frontal cortex by cocaine abuse is associated with changes in this region and decrease of its inhibitory effect over the urges generated in the nucleus accumbens, hippocampus, and amygdala and subsequent addictive pattern.

The interactions with serotonin receptors may explain the mood and appetite-suppressing effects of cocaine.

A more serious molecular effect that can explain the addiction in cocaine intake is the genetic impact of this alkaloid. Cocaine is known to change the amount of dopamine transporters and dopamine receptors on nerve cells via alteration of gene expression (genetic effects of cocaine). ΔFosB is a natural protein substance present in small amount in nerve cells, especially in the nucleus accumbens. It plays a role in the genetic mechanisms of the basic cell functions – the integrity and the interaction with other cells. In chronic cocaine intake, this protein accumulates in large quantities in the nucleus accumbens and is thought to be the part of the mechanisms explaining the addiction to cocaine. Changes in ΔFosB levels in the nucleus accumbens have been demonstrated in long-term cocaine intake in mouse models. It is known that one of the genes stimulated by ΔFosB, the enzyme cyclin-dependent kinase 5 (CDK5), promotes nerve cell growth. This factor also affects nuclear factor-kappa B and MEF2 (myocyte enhancer factor-2) expression. These effects are not well understood. It has been speculated that probably these transcriptional and epigenetic changes could be the genetic mechanism of the very long-term effects of cocaine. In a very long term, intake of cocaine increased dendrite growth and increases the number of the neurons in the nucleus accumbens that has been observed, i.e., increased cell contacts with other parts of the nervous system with altered information pathways and increased amount of signals coming to and originating from these cells with stable behavioral changes (Volkow and Morales 2015; Robison and Nestler 2011). These very long-term effects, based on genetic and epigenetic changes in the brain, probably make cocaine addiction very difficult to counteract. Another long-term effect of cocaine is dopamine depletion that is probably responsible for withdrawal symptoms (Quinn et al. 1997).

The main medical strategies to treat cocaine addiction and withdrawal are (Quinn et al. 1997) the following: the use of antidepressants (in order to inhibit neurotransmitter reuptake, particularly desipramine), dopamine agonists (to counteract dopamine depletion in the central nervous system), dopamine antagonists, anticonvulsants, and opioids (buprenorphine – probably through affecting the linkage between opioid and dopaminergic pathways).

The treatment of acute intoxication with cocaine and amphetamines is generally supportive: regulation of hydration and electrolyte disturbances; treatment of hypertension, rhythm, and conduction disturbances; use of vasodilators; gastroprotection; etc. Similar to heroin, cocaine and amphetamines are known to cause severe endothelial dysfunction and hemolytic-uremic syndrome (Kavannagh et al. 2006). Therefore, antithrombotic prophylaxis should be administered.

Cocaine has several metabolites: benzoylecgonine, ecgonine methyl ester, and norcocaine. Benzoylecgonine is a potent vasoconstrictor in vitro, but does not cross the blood-brain barrier in vivo. Ecgonine methyl ester (EME) is actually a vasodilator. It is produced by metabolization of cocaine by plasma cholinesterase (also known as pseudocholinesterase, or butyrylcholinesterase). "Pseudocholinesterase deficiency" due to BCHE gene mutations, is a specific condition that could render patients more vulnerable to severe intoxication with cocaine, to prolonged paralysis with succinylcholine and mivacurium.

Cocaine drug interactions could be examined in the light of three situations: cocaine intoxication, withdrawal, and long-term treatment and craving prevention. It seems there is scarce evidence of interaction between cocaine and CYP3A4 inhibitors, ketoconazole, erythromycin, and clarithromycin. There are other factors whose importance has to be established in the future, such as glutathione peroxidase-1 deficiency and microRNAs (Gallelli et al. 2017).

Cocaine intoxication leads to tachycardia, hypertension, and vasospasm. Treatment of these sometime fatal symptoms is done through the use of benzodiazepines, calcium channel blockers, and nitric oxide-mediated vasodilators. Nitroglycerine could induce reflex tachycardia through severe hypotension, so it should be used with extreme caution. Alpha-1 blockers had been tried with limited evidence. Alpha-2-adrenoceptor agonist trials had better results, especially with the

use of dexmedetomidine. There had been a widespread belief that beta-blockers could dangerously worsen hypertension during cocaine intoxication (Lange et al. 1990). However, this concept had been challenged recently. There were several Level I/II, Level III, and Level IV/V studies of β-blockers, with 1744 subjects, 7 adverse drug events, and 3 treatment failures. There were no adverse events reported for labetalol and carvedilol, mitigating hypertension and tachycardia (Richards et al. 2016). Antipsychotics have been used and studied for the treatment of hypertension and tachycardia, improving agitation and psychosis (paranoia), but there are significant risks with QTc prolongation and extrapyramidal adverse effects. Since second-generation antipsychotics have serotonergic effects, clinicians need to be aware of the potential risk of serotonin syndrome, by potentiating serotonergic effects of cocaine. Other medications include lidocaine, sodium bicarbonate, amiodarone, procainamide, propofol, intravenous lipid emulsion, and ketamine.

Cocaine withdrawal and cravings is a challenging condition due to several phenomena, including behavioral sensitization. Antipsychotics have been tried with mixed results. The biological mechanism of counteracting the effects of cocaine is thought to be due to presynaptic action on dopaminergic and serotonergic, while cocaine affects directly and indirectly the postsynaptic cascades. Data analysis shows that actually antipsychotics do not have advantages over placebo in regard to cocaine use and cocaine abstinence or craving. They could even cause more discomfort, even depression related to discontinuation (Kishi et al. 2013). Cochrane review did not support the notion of using antidepressants in the treatment of cocaine withdrawal (Pani et al. 2011). There had been some serious adverse reactions reported regarding the use of citalopram and cocaine – potentiation of serotoninergic vasoconstriction (Medicines and Healthcare products Regulatory Agency 2016). The more successful medication interaction is the one with GABAergic medications, topiramate, although there is still not enough conclusive unequivocal evidence for its efficiency.

Drug Interactions of Cocaine and Other Substances

Antipsychotics – increased risk of antipsychotic induced acute dystonias, both in intoxication and chronic treatment. Clozapine could lead to increased cocaine plasma concentrations and reduced psychotic and pressor effects.

Mood stabilizers (carbamazepine) – plasma concentrations of norcocaine increase – higher risk of hepatotoxic and cardiotoxic effects (Tenev 2008).

Benzodiazepines – oversedation and increased risk of benzodiazepine abuse.

Disulfiram – threefold increase of plasma levels of cocaine and increased risk of cardiotoxic complications.

β-blockers – very high risk of myocardial ischemia.

Nicotine – has a synergistic effect on dopamine release in the reward areas of the brain; lowers the oxygen supply, arterial pressure, and cardiac contractility; and increases the incidence of cardiac complications arising from cocaine use.

Alcohol – ethanol-induced metabolite, cocaethylene, of cocaine is more reinforcing than cocaine and is potentially more toxic.

Amphetamine and Its Derivatives

Amphetamine and its derivatives are not present in nature and represent purely synthetic substances. Amphetamine was first synthesized in 1887 and was initially used for the treatment of nasal congestion and subsequently as stimulant, athletic performance and cognitive enhancer, aphrodisiac, and euphoria inducer.

Amphetamine and its derivatives (methamphetamine and methoxy-substituted amphetamines; 3,4-methylenedioxyamphetamine (MDA); 3,4-methylenedioxy-methamphetamine (MDMA) or ecstasy; *N*-ethyl-3,4-methylenedioxyamphetamine (MDEA); 2,5-dimethoxy-4-methylamphetamine (DOM); p-hydroxydimethoxy-4-methylamphetamine (PMA)) are purely synthetic stimulants that act by increasing the monoamine levels in the synaptic

cleft (Quinn et al. 1997; Volkow and Morales 2015). The half-life of MDMA in humans is 8–10 h.

- Inhibition of monoamine uptake (competitive inhibition of dopamine uptake)
- Increase in neurotransmitter release (facilitation of dopamine release from the vesicles and increase in dopamine transporter-mediated reverse transport of the mediator into the synaptic cleft, independently from the action potential-induced vesicular release)
- Inhibition of monoamine oxidase (MAO)

The first and the second mechanisms are mediated by binding to trace amine-associated receptor 1 (TAAR1). Ecstasy is also known to increase serotonin liberation (Rudnik and Wall 1992) and the release of oxytocin.

The molecular and physiological effects of amphetamines are similar to those of cocaine, but they are known to inhibit MAO and to have virtually no local anesthetic effect. Methamphetamine has two enantiomers with the S-(+) being five times more active.

The physiological, psychological, and toxic effects of amphetamines are similar to those of cocaine and are mediated by their sympathomimetic and serotonin-mediated effects.

The underlying mechanisms of addiction, dependence, and withdrawal in amphetamine intake are associated with changes in gene expression (transcriptional and epigenetic changes) in the mesocorticolimbic projection. The major transcription factors responsible for these alterations are ΔFosB, CREB (cAMP response element-binding protein), and nuclear factor-kappa B (Rudnik and Wall 1992). The crucial role of ΔFosB overexpression in the development of drug addiction to many substances (including alcohol, cannabinoids, cocaine and amphetamines, nicotine, opioids, dissociative anesthetics, and hallucinogens) is demonstrated by the profound effect of ΔJunD in such cases. ΔJunD is an enzyme that blocks ΔFosB overexpression, and when brought to the nucleus accumbens by a viral vector, it could reverse the behavioral changes in chronic drug abuse and addiction.

The overexpression of ΔFosB in amphetamine abuse (like in cocaine abuse) leads to marked and long-standing functional effects and changes in dopaminergic neurons, especially in the nucleus accumbens, hippocampus, amygdala, and frontal cortex with the development of addiction. This addiction pattern is due to deep receptor, mediator, and structural changes in the neurons, and currently there are no known medications to counteract addiction in such patients.

Medical strategies have been developed to treat acute intoxication – i.e., for the treatment of cardiac (tachycardia, rhythm and conduction disturbances, hypertension) and vascular (vasoconstriction, endothelial dysfunction) symptoms, hyperthermia, dehydration, dyselectrolytemia, inadequate antidiuretic hormone secretion, intracranial complications (ischemic stroke and hemorrhage), respiratory failure and acute respiratory distress syndrome, and hepatic and liver failure. The hepatic failure is known to develop due to the oxidation of mitochondrial proteins and acute microsomal toxicity, combined with ischemia (vasoconstriction plus thrombosis), and renal failure is usually due to dehydration in combination with rhabdomyolysis and/or development of hemolytic-uremic syndrome (Kavannagh et al. 2006; Moon et al. 2008).

There are acute and long-term toxicity phenomena. There are several sources of data: in vitro experiments, animal models, and in vivo observations. There are still a lot of studies to be done to unequivocally prove the specific interactions and their clinical significance.

Acute toxicity	Chronic toxicity
Euphoria, well-being, happiness, stimulation, increased energy, extroversion, feeling close to others, increased empathy, increased sociability, enhanced mood, and mild perceptual disturbances. In addition, cardiovascular-	Neurotoxicity Impairment in serotonin function Neurodegeneration Phenocopying phenomenon – compromising the extensive metabolizer capability; developing

(continued)

Acute toxicity	Chronic toxicity
related somatic symptoms, autonomic effects (dry mouth, sweating, tremor, mydriasis tremor, jaw clenching, and restlessness), and moderate derealization have been observed (de la Torre et al. 2004) Hyponatremia – uncommon, associated with inappropriate antidiuretic hormone (SIADH) secretion and excessive water intake (also in polymorphic reduced COMT activity) Fulminant hepatitis and hepatic necrosis have been described too	low tolerance to methamphetamine after a short period of experiencing less toxicity of the substances (EM to PM status change)

The toxic effects are related to the metabolism of MDMA and methamphetamine and their metabolites. MDMA is a substrate to CYP2D6, but also a potent inhibitor through the so-called mechanism-based inhibition, by the phenocopying phenomenon. The effective enzyme amount decreases, so even genotypically active metabolizers become similar to poor metabolizers. Regardless of the genotype/phenotype, it could take up to 10 days to resynthesize CYP2D6 and restore it back to its baseline level of activity after even a single recreational dose. It was thought that there were sex differences, with 67% of males and 100% of females having such phenotyping effect, exposing them to the adverse effects of the drugs. Female subjects in the study setting would display more intense physiological (heart rate and oral temperature) and negative effects (dizziness, sedation, depression, and psychotic symptoms).

Currently it had been proven that the wide genotype allelic variations of CYP2D6 actually do not play the role they had assigned before. That could be due to the alternate pathways during the first phase of methamphetamine and MDMA metabolism: CYP1A2, CYP2B6, CYP2C19, and CYP3A4. They also have multiple genotype/ phenotype variations and could undergo the same phenocopying phenomenon, thus making the occurrence of acute and chronic adverse effects dose-independent and unpredictable.

The second phase of metabolism of MDMA is through COMT. It converts the catechol metabolites HHMA and HHA into HMMA and HMA. The same enzyme inactivates dopamine (DA) and noradrenaline (NE). It exists in two forms. MB-COMT is in the brain and S-COMT in the liver/kidneys. There are two basic functional polymorphisms – valine (val) to methionine (met) substitution at codon 108 in S-COMT and at codon 158 in MB-COMT. The latter variant, the Met allele, is associated with low enzymatic activity, while the former, val allele, has higher activity. Roughly one fourth of the population has low activity, and one fourth has high activity. The lower the activity, the higher the toxicity through the accumulation of the immediate active MDMA metabolites – HHMA and HHA. **This could subsequently increase the risk of clinical symptoms including hyperthermia, hypertension, tachycardia, seizures, serotonin syndrome, and rhabdomyolysis.**

HMMA plasma concentrations play significant role, regardless if these are linked to CYP2D6 genotype (higher with two functional alleles). Genotypes of COMT val158met or 5-HTTLPR with high functionality (val/val or l/*) determine greater cardiovascular effects and with low functionality (met/* or s/s) negative subjective effects, such as dizziness, anxiety, and sedation.

An important role is attributed to glutathione S-transferase (GST) in the detoxification of HMMA. There had been some data in vitro showing differences in toxicity related to GST polymorphism, which actually had not been observed in vivo. The conjugation during elimination process in phase II of metabolism goes through SULT system, leaving sulfate conjugated MDMA urinary metabolites and UGT system – glucuronide conjugate urinary metabolites. There are also some genetic variants, especially in UGT system, which could lead to decreased enzymatic activity, hence longer elimination and increased toxicity of MDMA (UGT2B15).

MDMA and amphetamine toxicity is dynamically related to individual differences in DAT expression both at pre- and postsynaptic levels. The dopamine transporter gene could modify

indirectly the receptor signaling done by the drugs. Reduced SERT potentiates self-administration of MDMA and cocaine (Brox and Ellenbroek 2018).

Types of consequences due to interactions or enzymatic polymorphisms

1) MDMA-induced toxicity

Enzymatic polymorphism	Interactions
Phase I – CYP2D6, CYP1A2, CYP2B6, CYP2C19, CYP3A4 Phase II (COMT, GST, SULT, UGT) – Less active isoenzyme genotype, or depletion of normal activity genotype (phenocopying) NT reuptake transporters: SERT, NET, DAT NT synthesis, breakdown: TH, TPH, COMT, MAO NT receptors	CYP2D6 inhibitors SERT inhibitors – Fluoxetine, paroxetine, citalopram NET inhibitors – Duloxetine, reboxetine DAT inhibitors – Bupropion, duloxetine 5-HT2 antagonists – Ketanserin, mirtazapine α-β-adrenergic antagonists – Carvedilol Antipsychotics (Rietjens et al. 2012)

The most dangerous toxic phenomena related to MDMA and methamphetamine are as follows:

Serotonin syndrome: (1) Mental status changes, (2) autonomic hyperactivity, and (3) neuromuscular abnormalities, all with varying signs from tremor and diarrhea to delirium, neuromuscular rigidity, and life-threatening hyperthermia. Death could occur due to increased serotonin levels via MDMA-induced 5-HT release and inhibition of 5-HT degradation via MAO inhibitors (Rietjens et al. 2012). The highest risk for this syndrome is in combination with antidepressants.

Hyperthermia: MDMA-induced cutaneous vasoconstriction and metabolic heat production.

Several dangerous reactions related to CYP2D6 inhibition had been described. Ritonavir and antiretroviral drugs have had life-threatening effects described. The unpredictability of these reactions is derived from the genetic polymorphism of CYP450 and alternate pathways for Phase I of MDMA metabolism.

Pharmacodynamic interactions that could lead to MDMA tolerance and increase the recreational dose, due to lack of the desired effect (no "high," less intense sensation of euphoria) This effect could be protective against neurotoxicity, exerted directly by MDMA or its toxic metabolites. Two mechanisms for that had been suggested. The first is the reduced 5-HT release and SSRI exerted prevention of MDMA to interact with SERT, blocking the efflux of serotonin through SERT. The second seems to be direct inhibition of CYP2D6 by such antidepressants like paroxetine and fluoxetine. Thus, MDMA metabolism is blocked. Concentration of toxic metabolites HHMA and HMA and their reactive quinones remains low. The risk of this type of interaction is that consumer could increase the dose.

Drug Interactions

It is important to clarify the timeline for assessing and predicting the drug interactions, since there is a differentiated response to MDMA and methamphetamine after infrequent one dose or seldom binges vs chronic daily use. It is possible that one and the same medication has different effect on the metabolism of these drugs after sporadic or chronic stimulation of the enzymatic activity and the genotype predisposition of CYP450, UGT, GST, COMT, and SULT. The level of affected NAT, DAT, and SERT transporters is also worth mentioning. In this context some varieties could be anticipated with regard to reaction and side effects of MDMA and methamphetamine with patients taking antidepressants, antipsychotics, or antiepileptic medications. Another very important issue is if there had been pretreatment, i.e., the person had been receiving a medication before using MDMA (or amphetamines). It could potentially change the reaction during acute intoxication, withdrawal, or maintenance of sobriety treatment. Further research needs to be conducted to elucidate individual differences (Table 2).

Of note: severe MDMA intoxication is addressed by cooling measures and use of benzodiazepines.

Antidepressants – Bupropion (CYP2D6 inhibitor) (could lower the pharmacological effects (both cardiovascular and euphoric) of

Table 2 Pharmacokinetic and pharmacodynamic interactions of MDMA, methamphetamine and other medications at different time points in treatment timeline

Timeline of medication intervention related to meth/ MDMA use	Pretreatment	During intoxication	During withdrawal	During maintenance of abstinence
Antidepressants				
Citalopram	Prevention of depletion of 5-HT (Schmidt and Taylor 1987) (in rat models)	Could ↑ locomotor activity, through ↑ D2 receptor expression (rat models)	Could exacerbate physiologic effects	Do not improve significantly depressive symptoms
Paroxetine Fluoxetine (FLX)	Inhibition of CYP2D6, low level of active, toxic metabolites	Limited data FLX reduces 5HT depletion, does not affect hyperthermia SSRIs could ↑ risk of serotonin syndrome Not suitable to start treatment, since FLX needs 6–8 h to reach therapeutic plasma concentration, same time for elimination of MDMA	Reverse reward deficits during amphetamine withdrawal (Harrison et al. 2001)	Nonconclusive reports regarding prolongation of abstinence or treating depressive symptoms
Duloxetine	↑ MDMA levels, through inhibiting SERT and NET, ↓ tachycardia, ↓ hypertension, weak DAT inhibitor	Could be used, not enough data	Not enough data	No adverse effects noted
Mirtazapine	Could ↓ consumption, ↓ erratic sexual behaviors	Not enough data	Further studies need to be done	Not enough data
Bupropion	↑ MDMA levels, through CYP2D6 block, ↓ adverse effects			Could prolong abstinence
Imipramine		↑ risk of cardiovascular, GI, anticholinergic effects		↑ abstinence time
Sertraline	Should not be administered to patients with methamphetamine-related disorders, due to adverse effects on abstinence, AWMF (Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften www.awmf.org)			
Antipsychotics				
Haloperidol	Reduces hyperthermia ↓ depletion of 5-HT	Change subjective MDMA effects from a pleasurable state of well-being and euphoria to a more dysphoric state with slightly increased anxiety, i.e., *akathisia* (Rietjens et al. 2012)		
Clozapine	Reversal of MDMA-induced cutaneous vasoconstriction (Blessing et al. 2003) and inhibition of MDMA-induced Increases in metabolic heat production			

methamphetamines by blocking its toxic metabolites). It had been tried in treating moderate and non-daily users. It could be administered for prolonging the abstinence period, although data is still not conclusive (Härtel-Petri et al. 2017b). From clinical perspective, one could make the case that increasing dopaminergic transmission could facilitate resolution of temporary depressive symptoms after "methamphetamine crash." Bupropion could also give false-positive urine drug screen for amphetamines, in 41% of cases (Casey et al. 2011).

Trazodone could yield false-positive urine drug screen for amphetamines, especially after pretreatment with phenothiazines.

Mood stabilizers – Lithium (dehydration more pronounced).

Antipsychotics – Quetiapine and risperidone for the treatment of depressive and psychotic symptoms in the chronic methamphetamine use syndrome. In acute phase antipsychotics and methamphetamine could reduce the efficacy of each other.

Antiretroviral drugs – Ritonavir (severe CYP2D6 inhibitor, increased level of MDMA).

α-β-adrenergic receptor antagonists – Carvedilol (could potentially decrease hyperthermia).

Alcohol – Slows down the effects of MDMA and increases nephrotoxicity, leading to high risk of lethal dehydration.

Urinary alkalinizers' (OTC medications) use leads to increased tubular reabsorption, via the increased amounts of non-ionized amphetamine. Thus, methamphetamine could have its half-life increased two- to threefold, while MDMA's half-life could increase by twofold.

Antihypertensive medications – MDMA and methamphetamine counteract their effects and could render the hypertension control more difficult to maintain in the long term.

Tobacco/nicotine – Smoking methamphetamine in combination with tobacco creates the pyrolysis product cyanomethylmethamphetamine. This metabolite has some stimulant properties (Dean 2006).

Methamphetamine-related, post-acute persistent or comorbid syndromes such as methamphetamine-associated psychosis (MAP), depressive syndromes, anxiety, and sleep disorders are usually treated in a symptom-oriented manner. The interactions could be unpredictable, could happen on many different levels, and could change dynamically. This makes it very important to use medications with the most available data for efficacy possible. Further research is warranted (Härtel-Petri et al. 2017a). Methamphetamine and MDMA could lower the seizure threshold.

Marijuana and Synthetic Cannabinoids

Cannabis is the most widely used illicit substance all over the world (Sharma et al. 2012). It has been used for centuries for recreational purposes. It contains more than 400 active substances, 61 of which are cannabinoids and have certain psychoactive properties. The main psychoactive substance is delta-9-tetrahydrocannabinol (THC). In the human body, THC binds to specific psychoactive and functional effects outside the CNS. It is used for recreational purposes, but because of its wide spectrum of effects and tendency to cause dependency and profound behavioral changes, THC is illegal in the most part of the world.

THC is derived from the leaves, stems, and seeds of the Indian hemp (*Cannabis sativa*). The parts of the plant can be smoked or taken orally, even mixed with food and cooked. When smoked, *Cannabis sativa* leaves, stems, and seeds liberate more than 2000 substances, most of which are produced via pyrolysis (Sharma et al. 2012), but the major psychoactive substance is THC. In the human body, cannabinoids bind to specific cannabinoid receptors (CB1 and CB2) that have physiological ligands (anandamides) that belong to the arachidonate derivatives. The latter act via affecting cAMP intracellular levels and ion transport (calcium and potassium) in different organs.

The multiple physiological effects of cannabinoids are mediated by two types of receptors – CB1 and CB2. CB1 are expressed mainly in the brain areas responsible for the cognitive, memory, pain, reward and anxiety, and endocrine and

motor functions, while CB2 are expressed by the peripheral tissues and organs. The exact mechanisms of action of cannabinoid in the body are not well understood, but it is assumed that the binding to cannabinoid receptors leads to the activation of several signal pathways, including dopamine, serotonin, and norepinephrinergic, GABAergic, opioid, cholinergic, glucocorticosteroid, and prostaglandin systems (Sharma et al. 2012). It is also known that cannabinoids can directly interact with opioid and benzodiazepine receptors and can affect protein, nucleic acid, and prostaglandin synthesis (Sharma et al. 2012), hormone secretion, and DNA repair and replication (Sharma et al. 2012; Reece and Hulse 2016; Li and Lin 1998; Reece 2009). CB2 receptors are expressed in multiple tissues and organs, including the gastrointestinal tract, endocrine glands, and immune systems, and this can at least partially explain the effects of cannabinoids on these structures. Moreover, there is evidence that cannabinoids interact with vanilloid and vanilloid-like receptors (Pertwee 2005) on glutamatergic and alpha-adrenergic receptors and on multiple peripheral tissues.

Cannabinoids, both natural and synthetic, have certain adverse effects on the mental status (including triggering overt psychoses), respiratory tract (including the development of obstructive lung disease and lung cancer), cardiovascular system (changes in blood pressure, ischemic organ damage, inflammatory angiitis, arrhythmias, worsening of the metabolic profile, etc.), bone loss, fetal retardation, etc. (Reece 2009).

Of special interest is the mutagenic, teratogenic, and genotoxic effect of cannabinoids that has become even more visible due to the widespread abuse of cannabis. Of crucial importance are the permanent genetic changes arising during in utero exposure to cannabinoids, leading to the formation of inheritable malignancies, such as childhood neuroblastoma, leukemia, and rhabdomyosarcoma (Reece 2009). These effects are mediated by at least three major mechanisms (Reece 2009):

- Oxidation of DNA plus inhibition of DNA repair (via induction of the formation of nitrogen-centered species and by uncoupling of mitochondrial oxidative phosphorylation), with deoxidation of guanosine to oxo-guanosine being a normal part of the endocannabinoid signaling
- Changes in enzyme activity: stimulation of MAP kinase pathway (an important factor for the induction of non-lymphoblastic leukemia), inhibition of topoisomerase II pathway, and RAD-1 inhibition and damage
- Changes in telomeres due to the inhibition of telomerase (this enzyme is present in stem cells, gonadal/germ cells, and cancer cells but not in the normal somatic cells)

In cell cultures, marijuana smoke condensates have been shown to increase the formation of reactive oxygen species (ROS) and to inhibit the synthesis of the transcription factor p53 that acts as a tumor-suppressor protein (Kim et al. 2012).

The synthetic cannabinoids (fake weed, spice, K2, etc.) are synthetic cannabinoid derivatives with stronger affinity toward the cannabinoid receptors with more pronounced psychomodulating and adverse effects and unknown safety. Their peripheral, long-term, and genetic effects are unknown and hard to predict.

The multiple receptor targets of cannabinoids, their epigenetic and genetic effects, and the unclear mechanism of signaling changes, in combination with their widespread abuse, make the treatment of cannabinoid addiction extremely difficult.

THC and CBD are metabolized mainly in the liver by cytochrome P450 isoenzymes (mainly CYP2Cs and CYP3A4). In vitro studies indicate that THC and CBD both inhibit CYP1A1, CYP1A2, and CYP 1B1 enzymes, and recent studies have indicated that CBD is also a potent inhibitor of CYP2C19 and CYP3A4. Both cannabinoids may interact with other medications metabolized by the same pathway or by inducers/inhibitors of the isoenzymes.

It is important to distinguish different pathways of metabolism, related to different ways of administering both substances. Preparations which have Δ9-THC inhibit CYP2C9 and CYP3A4. CBD inhibits mostly CYP2C19 and

CYP3A4. Marijuana inhalation (pyrolysis) inducts CYP1A1 and CYP1A2. Patients with lower activity of CYP2C9 and/or CYP3A4 (phenotypical or by genotype) could have increased plasma concentrations of other substances they take together with Δ9-THC, while CBD exposure in patients with diminished CYP2C19 and/or CYP3A4 function could lead to unpredictable risk of adverse effects of medication substrates of these enzymes. There are very few documented interactions that are proven in vitro and in vivo, such as TCAs or anticholinergic drugs that can produce significant tachycardia. This may be due to beta-adrenergic effects of cannabis coupled with the anticholinergic effect of tricyclic antidepressants. Most of the other interactions are actually hypothesized, and there is still lack of sufficient controlled trials to state evidence-based approach. Nevertheless, it could be very well expected to have increased plasma concentrations with these substrates, due to the occurrence of possible drug interactions.

	CYP3A4 substrates	CYP2C9 substrates
Δ9-THC preparation	*Antidepressants*: amitriptyline, citalopram, clomipramine, fluoxetine, imipramine, mirtazapine, paroxetine, sertraline, trazodone, venlafaxine *Antipsychotics*: pimozide, quetiapine, risperidone, ziprasidone, aripiprazole, chlorpromazine, clozapine, haloperidol, perphenazine *Benzodiazepines*: clonazepam, diazepam, nitrazepam, alprazolam, midazolam *Sedatives*: zaleplon, zolpidem	*Antidepressants:* Fluoxetine, sertraline, amitriptyline *Selective AT1 angiotensin II receptor antagonists*: losartan, valsartan *Oral hypoglycemic*: sulfonylureas, glimepiride, glipizide, glyburide *NSAIDs* *Others*: phenytoin, S-warfarin, zolpidem
CBD preparations		**CYP2C19 substrates**
		Antidepressants: amitriptyline, citalopram, clomipramine, fluoxetine, imipramine,
	CYP3A4 substrates	CYP2C9 substrates
	Analgesics: buprenorphine, codeine, fentanyl, hydrocodone, tramadol, lidocaine. *Antiarrhythmics*: amiodarone *Ca channel blockers*: amlodipine, diltiazem, nimodipine, verapamil *Beta-blockers*: metoprolol, carvedilol *Protease inhibitors*: ritonavir, lopinavir, nelfinavir, indinavir *NNRTIs*: efavirenz *Antiepileptics*: carbamazepine, ethosuximide, valproic acid, zonisamide *Statins*: atorvastatin, simvastatin *Antibiotics*: azithromycin, clarithromycin, erythromycin *Antifungals*: ketoconazole, fluconazole, miconazol *Others*: dextromethorphan, sildenafil, tamoxifen, ondansetron, PPIs	*sertraline, venlafaxine Barbiturates: hexobarbital, mephobarbital PPI: lansoprazole, omeprazole, pantoprazole, esomeprazole Benzodiazepines: alprazolam, diazepam, flunitrazepam Others: moclobemide, propranolol, nelfinavir*
Marijuana smoking lowers the concentration of CYP1A1, CYP 1A2 substrates, potentially	**1A1 substrates**	**1A2 substrates**
	Compounds of tobacco smoke: heterocyclic amines and polycyclic aromatic hydrocarbons CYP1A1 is a carcinogen-metabolizing enzyme. Its activation or inhibition could	*Antidepressants*: amitriptyline, clomipramine, fluvoxamine, mirtazapine *Antipsychotics*: chlorpromazine, clozapine, fluphenazine, haloperidol, olanzapine, perphenazine,

(continued)

	CYP3A4 substrates	CYP2C9 substrates
	modify the cancer risk factors	thiothixene, trifluoperazine, ziprasidone *Others*: propranolol, caffeine, acetaminophen, riluzole, ropinirole, melatonin, R-warfarin, naproxen, ondansetron

Cannabis produces sedation, impairs psychomotor performance, and increases blood pressure and heart rate. Pharmacodynamic interactions with other sedatives can potentiate the central effects but can be decreased by psychostimulants. This review focuses on the interactions between cannabinoids and alcohol, other drugs of abuse, and prescription medicines.

It is important to note that the ratio between CBD and THC had changed over the years, with Δ9-THC having higher concentration now than in the past. Cannabidiol (CBD) had been used more and more in the treatment of epilepsy, including very recent promising results in the possible improvement of schizophrenia (McGuire et al. 2018).

Case reports suggest that concurrent use of cannabis with other illicit substances could lead to toxic interactions (Lindsey et al. 2012).

Benzodiazepines and Barbiturates

Benzodiazepines (BZDs) are among the most widely prescribed medications all over the world for a broad spectrum of indications, including insomnia, epilepsy, muscle spasms and contractures, alcohol withdrawal, and anxiety (Griffin et al. 2013). They are used in anesthesiology for premedication before general surgery because of their marked anxiolytic effect and the ability to cause anterograde amnesia. The use of some benzodiazepines (i.e., flunitrazepam) has been severely restricted due to their ability to induce abulia in combination with retrograde amnesia and the potential to be used for sexual assault and for "zombification."

Other point of concern is their ability to induce tolerance and dependence/addiction, the changes in their pharmacological profile with age, and the interactions with multiple medications.

The pharmacological effects of benzodiazepines (sedative, anxiolytic, muscle relaxation, antiepileptic) are mediated via modulation of GABAA receptor activity – i.e., in the central nervous system BZD represent GABAA agonists. The gamma-aminobutyric acid (GABA) is a universal inhibitory mediator in the central nervous system that decreases neuronal activity and excitability. The GABA receptors have three major types – A, B, and C – that represent chlorine channels, and BZDs bind selectively to GABAA. The latter is composed of five subunits – two alpha, two beta, and one gamma subunit. BZDs bind to the pocket created by α and γ subunits and change the space structure of the receptor that leads to increased binding of GABA and stimulation of GABA mediation; increased chlorine channel permeability, along with changes in sodium, potassium, and calcium membrane permeability; inhibition of calcium-dependent neurotransmitter release; and inhibition of adenosine neuronal uptake with unknown clinical significance (Quinn et al. 1997; DeVane 2016).

BZDs also have peripheral benzodiazepine receptors (DeVane 2016) that are unrelated to GABAA – in the peripheral nervous system, glia, immune system structures, etc. BZDs also act as mild adenosine reuptake inhibitors thought to explain, at least partially, their anticonvulsant and anxiolytic effects.

Barbiturates bind to a different part of the same receptor as BDZ bind. Barbiturates cause similar effects to the ones that BDZ have (DeVane 2016). While BZDs increase the frequency of the chlorine channel opening, barbiturates increase the duration of the opened state. This leads to increased risk of toxicity of barbiturates. Moreover, barbiturates are known to bind and affect other CNS and peripheral receptors, including inhibition of ionotropic glutamate receptors (kainate and AMPA receptors) and inhibition of

P/Q type of voltage-dependent calcium channels (leading eventually to inhibition of glutamate release). Barbiturates also bind to ligand-gated ionic channels (cationic – nAChR, 5-HT3 receptor, and glycine receptor ionic channels), causing depression of the CNS. The stated extra-GABAA receptor and ionic channel effects of barbiturates leading to more pronounced CNS depression compared to BZDs plus the lack of specific antagonist have led to significant restriction of barbiturate use in the clinical practice (DeVane 2016).

Benzodiazepines

BZDs were first introduced to the clinical practice in the 1960s as tranquilizers and sedatives and subsequently were administered as anticonvulsants and hypnotics.

Depending on the duration of action, BZDs are classified as having short, intermediate, and long duration of action. Short and intermediate acting are used mainly for the treatment of insomnia and long acting for anxiety. There are two main mechanisms of tolerance and addiction in BZDs (Quinn et al. 1997): downregulation of GABAA receptors in the limbic system and increased sensitivity of the benzodiazepine-GABAA receptor complex to inverse agonists. Moreover, similar to alcohol addiction and withdrawal, mechanisms have been described, including, respectively, changes in the expression of corticotropin-releasing hormone (CRF) and CRF receptor sensitivity and neuropeptide Y and increased NMDA and AMPA receptor sensitivity (affecting glutamate neurotransmission). BZD withdrawal syndrome is characterized by sleep disturbance, irritability, anxiety to panic attacks, confusion, nausea, weight loss, changes in blood pressure and heart rate, muscle stiffness, irritability, headache, perceptual changes, psychotic reactions including hallucinations and delusions, and suicidal thoughts and attempts.

In the elderly, BZDs tend to have more unfavorable profile of side effects and are therefore included in the Beers List of inappropriate medications in the elderly.

Specific adverse effects of BZDs on the central nervous system that are even more pronounced in elderly are (Griffin et al. 2013; DeVane 2016):

- Cognitive impairment and other toxic effects on the CNS, including sedation, drowsiness, inattentiveness, motor impairment, anterograde amnesia, and ataxia.
- High risk for development of tolerance and addiction with severe withdrawal symptoms in abrupt cessation.
- Anterograde amnesia, especially concerning the long-term memory and impaired implicit and explicit memory; these effects are particularly dangerous because of the possibility for drug-facilitated sexual abuse, especially in flunitrazepam intake.
- Accumulation and subsequent disinhibition with impaired perception of inherent risk of inappropriate behavior (reckless driving, sexual behavior, etc.).
- BZD-induced delirium states, especially in the elderly and/or hypoxic patients with parenchymal organ failure.

BZDs have significant drug interactions with many prescription, nonprescription, and illicit drugs, including benzodiazepines, opioids, alcohol, and over-the-counter sleep medications. Benzodiazepines are metabolized through the liver, mainly through CYP450 to CYP3A4 isoenzyme.

On the other hand, BZDs have antagonist, flumazenil, used for the acute treatment of overdose, along with the supportive treatment (infusions, antibradycardic, antihypotensive medications, diuretics, etc.).

The treatment of BZD addiction is difficult, as dependence (both psychological and physical) develops relatively quickly and includes flumazenil and cognitive behavioral therapy.

Drug Interactions

Benzodiazepines could have their plasma levels increased through the inhibition of CYP3A4. Potent inhibitors, such as fluoxetine, imipramine, erythromycin, clarithromycin, etc., increased the plasma levels of benzodiazepines through the

inhibition of CYP3A4 by the antidepressants. They could have their plasma concentration lowered (or half-life shortened) when interacting with enzymatic stimulants such as carbamazepine. Benzodiazepines could have synergistic increased depressing effect on CNS and respiratory suppression with mirtazapine, alcohol, barbiturates, and antihistaminic medications.

Barbiturates

Barbiturates were first discovered in 1864, but their widespread use in the clinical practice as hypnotics began in the beginning of the twentieth century. Because of their marked side effects and the risk of dependence and addiction, currently barbiturates are used mainly as hypnotics, anticonvulsants, sedatives, and general anesthetics (sodium thiopental) and for the treatment of severe withdrawal symptoms of alcohol and illicit drug abuse (as sedatives, like BZDs) (DeVane 2016).

Their profile of addiction and withdrawal is similar to benzodiazepines, but they have no known receptor antagonist. Therefore, the intoxication and dependence are more difficult to manage compared to BZDs, having in mind the broader spectrum of receptor and mediator systems affected by barbiturates.

Hallucinogens (LSD, Mescaline, Magic Mushrooms, Ayahuasca, Psilocybin, Dimethyltryptamine, DMT)

The classical hallucinogens are natural substances or their derivatives that cause perceptual alterations of real stimuli, described by the patients as hallucinations, via stimulation of serotonin mediation. These "hallucinations" actually represent distortion of the reality due to changes in perception. Classical hallucinogens have been used for religious and recreational purposes for hundreds of years. Classical hallucinogens are taken orally in the form of "magic drinks or potions," tablets, etc., and are rarely smoked. All act via activation of 5-HT2A receptor pathways. Depending on their chemical structure, classical hallucinogens can be divided into three major groups (Baumeister et al. 2014):

- Tryptamines: psilocin (the psychoactive compound of psilocybin, found in the magic mushrooms) and DMT (*N*,*N*-dimethyltryptamine, the psychoactive compound of ayahuasca) and its derivatives
- Lysergamides: LSD (lysergic acid diethylamide) and LSP (lysergic acid 3-pentylamide) and its derivatives
- Phenylamines: mescaline (the psychoactive compound of the peyote cactus) and DOI (2,5-dimethoxy-4-iodoamphetamine) and their derivatives

Psilocybin, dimethyltryptamine, and mescaline occur in nature, and the rest are synthetic. Ketamine, MDMA, and salvinorin A are able to induce similar changes in the state of consciousness but are not classified as classical hallucinogens because their effects are mediated via other receptor pathways.

All three types of classical hallucinogens act via the stimulation of 5-HT2A receptors (Baumeister et al. 2014) and cause distortion of environmental stimuli (i.e., temporary distortion of reality), perceived as hallucinations, traveling to or contact with other worlds, etc. They also bind to metabotropic serotonin receptors and affect large number of intracellular signaling pathways, the significance of which is not clear (Baumeister et al. 2014). These substances also tend to have antidepressant effects.

The pharmacodynamic studies on these substances have shown that they act via the stimulation of 5-HT2A receptors and can therefore cause cross-tolerance (i.e., between psilocybin and LSD). 5-HT2A antagonists (ketanserin and risperidone) can block their action. Classical hallucinogens cause rapid downregulation of the stated receptors and the subsequent development of tolerance.

The activation of serotonin receptors initiates several signal transduction pathways: G q/11

signaling route activating phospholipase C, decreasing the activity of protein kinase C and increasing the release of calcium ion from the cells. Classical hallucinogens are also known to stimulate phospholipase A2 (with formation of arachidonate) independently of the described pathways. These substances are known to change gene expression in the brain: induction of c-fos (LSD and derivatives), egr-1 and egr-2 (LSD), etc. Hallucinogens affect the Gi/o proteins with subsequent activation of Src. They also affect metabotropic receptors (e.g., glutamate receptor mGluR2) and lead to metabolic and behavioral effect changes. It has been suggested that their hallucinogenic effects are mediated by the co-activation of 5-HT2A and mGluR2, as well as Gi/o proteins and their cascades.

5-HT2A receptors are located in several brain areas: pyramidal neurons of layer V projecting into layer I of the cortex and the thalamus (reticular nucleus, regulating the signal processing from the thalamus to the cortex). The reticular nucleus sends inhibitory GABAergic projections to the thalamus and allows the transfer of more sensory stimuli to the cortex. Psilocybin is known to decrease the metabolic activity in the thalamus, and this could be the mechanism of sensory alterations in its abuse. Probably the changes in information transfer via the stimulation of 5-HT2A receptor pathways represent the mechanism for the development of sensory hallucinations (literally opening doors of perception, described by Huxley in 1954). Classical hallucinogens have different receptor affinities and different half-lives, the longest having LSD up to 12 h).

The physiology of addiction to these substances is associated with persistent activation of 5-HT2A receptors (Baumeister et al. 2014) with epigenetic modifications leading to changes in information transfer and channeling and the need for repetitive stimuli to maintain the same level of information transfer. Medical strategies that counteract these changes are 5-HT2A receptor antagonists, such as ketanserin and risperidone. Still, some of the effects of hallucinogens, including the changes in information channeling, tend to persist in time and cause flashbacks months later, even years after the cessation of drug abuse.

Dissociative Anesthetics (Nitrous Oxide, Ketamine, Dextromethorphan, Phencyclidine, *Salvia divinorum*)

Dissociative drugs are hallucinogenic substances that alter the perceptions and the connection with the environment. They generate the feeling of detachment from the reality and from self.

The classic dissociative drugs have also anesthetic properties, and some of them are currently used in anesthesiology (both in humans and in animals). The following substances are classified as dissociative anesthetics: nitrous oxide, phencyclidine (PCP), ketamine, dextromethorphan (DXM), and *Salvia divinorum*.

Three main mechanisms of action of dissociative anesthetics have been described (Jevtovic-Todorovic et al. 1998; Sleigh et al. 2014; Anis et al. 1983; Capasso et al. 2006):

- Disruption of glutamate-mediated neurotransmission of signals via antagonizing *N*-methyl-D-aspartate (NMDA) receptors – nitrous oxide, phencyclidine, ketamine, dextromethorphan.
- Activation of kappa-opioid receptors: *Salvia divinorum.*
- Interaction with other receptors – e.g., phencyclidine inhibits nicotine acetylcholine (nACh) receptors and directly interacts with endorphin and enkephalin receptors and sigma2 receptors; phencyclidine and ketamine are partial dopamine D2-receptor agonists; and nitrous oxide blocks beta-2-subunit containing nACh channels; inhibits kainite, GABAc, AMPA, and 5HT-3 receptors; potentiates GABAa and glycine receptors; and activates two-pore-domain potassium channels.

The inhibition of NMDA receptors is known to have three effects: dissociative, neuroprotective (via inhibition of glutamatergic stimulation of neurons, which is beneficial in ischemic brain injury), and neurotoxic (inhibition of GABA and

cholinergic stimulation that could cause neuronal damage) (Jevtovic-Todorovic et al. 1998). These effects are age-dependent (Jevtovic-Todorovic et al. 1998).

The first dissociative anesthetic synthetized was the nitrous oxide – discovered in 1772 by Joseph Priestley and later used as "laughing gas." Subsequently, its anesthetic purposes were discovered, and it became clear that besides its anesthetic properties, it can be administered as recreational and neuroprotective agent with risk of neurotoxicity (Jevtovic-Todorovic et al. 1998).

Subsequently, ketamine and phencyclidine were discovered and were introduced as general anesthetics. Dextromethorphan was synthetized as an opioid analogue cough-suppressing agent and is currently part of many over-the-counter syrups against cough, in combination with antihistamines, paracetamol, and decongestants. *Salvia divinorum* is a plant abundant in Mexico and South America, traditionally used for religious purposes (for divination) and gastrointestinal motility problems [Capasso]. Its major psychoactive substance, a structurally unique trans-neoclerodane diterpenoids, is known as salvinorin A. It represents a potent kappa-opioid receptor agonist. Moreover, it inhibits enteral cholinergic transmission (explaining its anti-diarrheic effect) and has some mu-opioid receptor agonist action (Capasso et al. 2006).

Dissociative anesthetics can be administered via inhalation (nitrous oxide) and ingestion, intravenously (ketamine, phencyclidine, dextromethorphan), or chewing of the leaves (*Salvia divinorum*). A major problem of ketamine is that the dry substance has no taste, odor, nor color and if added to a drink may cause dissociative state that can be used for sexual assault and kidnapping.

As it was mentioned above, the classic dissociative anesthetics (PCP, ketamine, and DXM) act mainly via antagonizing NMDA receptors, i.e., inhibition of glutamate-mediated neurotransmission. This mediator transfers the excitatory signal to the adjacent cells, and as it is one of the major players in cognition and nociception, the inhibition of its signal pathways leads to changes in cognitive functions and inhibition of pain sensation. Moreover, the inhibition of glutamate transmission may lead to disruption of vital functions, changes in the mood, sensation of detachment from the environment and from self, depersonalization, derealization, etc. One should not forget that PCP and ketamine affect dopaminergic transmission, i.e., the "reward pathway."

As a parallel with classical hallucinogens, the exact mechanisms of receptor and postreceptor action of dissociative drugs are not well understood, but currently it is assumed that these substances act via temporary blocking of the communications between neurotransmitter systems in the brain and in the spinal cord and causing disorganization of the information channeling that regulates perception (including nociception), vital functions (regulation of sleep, hunger, body temperature, muscle control, sexual behavior), mood, and cognition. In time, due to the phenomenon of neuroplasticity, permanent changes in these information channeling systems may develop, resulting in permanent changes in mood, sleep, hunger, motion, etc.

Moreover, the combined intake of dissociative anesthetics with other addictive and psychoactive substances can be extremely dangerous, due to their strong effects on vital and mental functions. The concomitant intake with antidepressant drugs can cause serotonin syndrome with lethal consequences. In combination with stimulants, dissociative anesthetics can increase the heart rate and the blood pressure to life-threatening levels. In combination with sedatives and alcohol, they can suppress breathing.

In anesthesiology the medication with nitrous oxide and ketamine is accompanied by oxygen supplementation because without oxygen these substances can decrease oxygen saturation, especially the inhalation of pure nitric oxide. The latter also leads to depletion of vitamin B12 stored and subsequent development of megaloblastic anemia and peripheral neuron damage.

DMX is taken largely in the form of cough syrup where it is combined with other substances, such as antihistamines, paracetamol, and decongestants. Extremely large quantities of cough syrup are needed to achieve hallucinations with

DMX, and these contain high quantities of antihistamines, paracetamol, and decongestants that could be toxic for both the body and the brain, causing inhibition of vital functions, hepatotoxicity, rhythm and conduction disturbances of the heart, etc. Approximately 5–10% of the Caucasian population have genetic polymorphism affecting DMX metabolism that leads to increased risk of overdose.

The treatment of addiction and withdrawal to dissociative anesthetics is very difficult, because multiple receptor systems are engaged and frequently changes in vital functions, cognition, and behavior have already developed. Because their mechanism of action involves loss of GABAergic inhibition of the cholinergic excitatory mediation, GABA agonists and cholinolytic medications have beneficial effect in such cases.

Alcohol

The ethylic alcohol (spiritus vini) is the most frequently abused mood-changing substance in the world. It affects the life of millions of people worldwide causing alcohol-related diseases. Every year approximately two million people die of alcohol-related conditions, including cirrhosis, cancer, alcohol dependence syndrome, and traumatism (Quinn et al. 1997). Alcohol has been known to mankind since the dawn of human history. Some cultures and beliefs even have their gods of happiness in alcohol-containing beverages.

From pharmacodynamic point of view, the effects of alcohol are not mediated by any specific receptor systems but rather are triggered by changes in membrane fluidity, disruption of ion channels, and changes in phospholipase and protein kinase C activity (similar to that in LSD use, but much less severe). The alcohol is known to stimulate glutamate-mediated transmission in the central nervous system via *N*-methyl-D-aspartate (NMDA) receptor activation. This effect is thought to be responsible for its sedative and amnesic effects, and the overstimulation of this receptor can cause neuronal death. The latter mechanism is probably the underlying process of development of organic brain syndrome in chronic alcohol abuse. Alcohol is known to stimulate GABAA receptors. Moreover, chronic alcohol consumption leads to alterations in GABAA benzodiazepine receptor, and probably the withdrawal symptoms of alcohol (especially anxiety and seizures) are related to these changes. Alcohol consumption changes norepinephrine levels due to persistent inhibition of alpha2-adrenergic receptors in chronic abuse and leads to dopamine release and subsequent depletion of the nucleus accumbens. Therefore, withdrawal symptoms may be, at least partially, related to norepinephrine over-reactivity and dopamine release and depletion. In favor of this hypothesis is the beneficial effect of alpha2-adrenergic blockers, dopamine and serotonin antagonists, in alcohol withdrawal syndrome (Quinn et al. 1997).

Alcohol is also known to increase the release of 5-HT from central and from peripheral nerve endings and to interact with the opioid receptors in the prefrontal cortex and cause euphoria (Quinn et al. 1997). The administration of opioid receptor antagonists, such as naltrexone, is known to suppress alcohol dependence.

The risk of alcoholism is determined by certain environmental factors, by its interaction with multiple receptor systems (including adrenergic, dopamine, serotonin, NMDA, glutamate, and opioid receptor systems), and by genetic polymorphisms of alcohol and aldehyde dehydrogenase (Quinn et al. 1997; Higuchi et al. 1995).

The alcohol withdrawal syndrome is managed with sedatives, including benzodiazepines, beta-adrenergic blockers, dopamine and 5-HT antagonists, and naltrexone, used to counteract the described receptor and signaling pathways. Alcohol dependence is very difficult to manage because this substance is widely available and easy to access. Several strategies have been applied, including sedative and disulfiram intake and behavioral therapy (Quinn et al. 1997).

To exert its desired effect, a drug generally must travel through the bloodstream to its site of action, where it produces some change in an organ or tissue. The drug's effects then diminish as it is processed (metabolized) by enzymes and

eliminated from the body. Alcohol behaves similarly, traveling through the bloodstream, acting upon the brain to cause intoxication, and finally being metabolized and eliminated, principally by the liver. The extent to which an administered dose of a drug reaches its site of action may be termed its availability. Alcohol can influence the effectiveness of a drug by altering its availability.

Alcohol has three main pathways of metabolism:

1. *Alcohol dehydrogenase (ADH)* metabolizes alcohol to acetaldehyde. It is a toxic compound. It is also a proven carcinogen, with its high activity.
2. *Aldehyde dehydrogenase (ALDH)* turns acetaldehyde into acetate, which is less active and not a carcinogen. Acetate ends up being turned into water and carbon dioxide.

   Genotype of ADH, such as ADH1B*2, is more active. It increases the levels of alcohol-derived acetaldehyde quickly and with high intensity. It is common in people of Chinese, Japanese, and Korean descent but rare in people of European and African descent. This could be a protective factor, considering that intoxication leads to unpleasant experience and possibly would lower the risk for repetitive use (and developing of addiction). ALDH1A1*2 and ALDH1A1*3 on the other hand are the most frequently formed enzymes in patients with alcohol use disorder from African-American descent. There are many environmental factors (food, stress levels, other genetic influences) that shape the alcohol metabolism in different populations. This perhaps could explain the equal distribution of alcoholism and alcohol use disorders among Caucasians, Asians, Native Americans, and African-Americans per recent observations.

   Acetaldehyde alters glial cells' function. It has psychiatric and behavioral repercussions. Normal amounts could have euphoria-inducing, anxiety-reducing, hypnotic, and memory-inhibiting effects. With higher plasma levels, aggression could occur. Acetaldehyde could also lower the preference of alcohol, i.e., aversion to voluntary ethanol consumption. This is the key concept in the use of disulfiram. The disulfiram reaction occurs when alcohol is consumed in the presence of disulfiram, which blocks irreversibly aldehyde dehydrogenase, thus increasing steadily the levels of acetaldehyde. The latter's side effects are flushing, headaches, tachycardia, arrhythmia, nausea, vomiting, and hypotension. Disulfiram also inhibits CYP2E1 too, thus increasing the plasma levels of warfarin, phenytoin, and theophylline and also decreasing the clearance of benzodiazepines such as diazepam, oxazepam, and chlordiazepoxide and also caffeine and tricyclic antidepressants such as desipramine and imipramine. Disulfiram has two toxic metabolites: diethyldithiocarbamate (DDC) and its metabolite carbon disulfide (CS2). DDC blocks the activity of dopamine beta-hydroxylase, through copper chelation. With higher levels of DDD, dopamine can no longer turn into norepinephrine. Presynaptic norepinephrine gets depleted. Dopamine accumulation leads to secondary cardiac abnormalities.
3. *Alcohol could be also metabolized through P450 2E1 (CYP2E1)* and catalase. These enzymes break down alcohol to acetaldehyde. CYP2E1 "switches on" after large amounts of alcohol, that is, after the ADH and ALDH capacities are overwhelmed. Catalase also contributes to alcohol metabolism only at a very small extent (Edenberg 2007). CYP2E1 accounts for roughly 7% of all CYP450 isoenzymes. It is located on chromosome 10q26.3. There are three key polymorphisms in CYP2E1 gene studied as of recently. CYP2E*5 on 5′-regulatory region has two variants, G1293C (PstI) and C1053T (RsaI). CYP2E*6 is the third variant, which is detected by Dra I, one of the restriction enzymes used to digest complete genomes and pulsed field gel electrophoresis. This variant has lower activity and could lead to potentially toxic levels of acetaminophen, ethanol, and styrene (Haufroid V et al. 2002; McGraw 2014). There is different data regarding its ethnic distribution. Per Mittal et al. (2015) it is distributed as follows: Caucasians 9%, African-Americans 9%, and

Japanese 35%; while CYP2E1*6 seems to have been found in 19.6% of Asians, 10.1% of Africans, and 7.7% of Caucasians (Gurusamy and Shewade 2014).

4. *Fatty acid interactions, forming fatty acid ethyl esters (FAEEs),* which damage the liver and pancreas (Vonlaufen et al. 2007).

Drug Interactions

The alcohol-related types of drug interactions have three main dimensions: temporal, enzyme specific, and pharmacodynamic. They could be very complex and difficult to predict at times:

1. Temporal
   - Acute intoxication, especially after prolonged period of sobriety and low concentration and grade for enzyme synthesis – potential for inhibition of drugs' metabolism, through competing interaction with the same enzymes. Exposure to acute alcohol intoxication while on any of the antibiotics such as furazolidone, griseofulvin, and metronidazole can lead to disulfiram-like reaction – headaches, nausea, vomiting, and possibly convulsions. Some tricyclic antidepressants could become toxic after acute ingestion of alcohol. Warfarin could increase its plasma concentration leading to problematic bleeding. Gastrointestinal bleeding could happen with combination of non-opioid pain medications, aspirin, and alcohol, since aspirin increases the availability of alcohol.
   - Chronic use leads to a decrease in the drugs' availability, diminishing their effects, even in the absence of alcohol, for weeks after cessation of drinking. The clinical importance of this is marked as the need for increasing the dose of certain medications, which patient with chronic alcohol abuse had been taking before entering early remission from alcohol. The doses of these medications required by nondrinkers might be way lower. Chronic alcohol use could decrease the availability of rifampin. To the same token, the dose of propofol required to induce anesthesia might be increased in patients with chronic alcohol abuse. There is also increased risk of liver damage by the anesthetic gases enflurane and halothane. Chronic use of alcohol lowers the dose of antiepileptic medications, due to the stimulation of the same enzymes responsible for their metabolism, thus increasing the risk of seizures, even during sobriety periods. Propranolol could have its plasma concentration reduced, thus increasing the risk of hypertensive crisis. Acetaminophen could be transformed into toxic metabolites.
2. Enzyme-specific reactions

   Chronic alcohol use could influence carcinogenesis by several mechanisms. Acetaldehyde is a carcinogen, binding to DNA. It could form active substances, such as malondialdehyde adduct, which mediate lipid peroxidation and nucleic acid oxidation. Inducing CYP2E1 pathway also contributes to forming acetaldehyde and radicals and enhances degradation of retinoic acid affecting signaling pathways, such as estrogen signaling, favoring proliferation and malignant transformation of precancerous cells. Chronic ethanol intake is also associated with the failure of immune surveillance of tumor cells (Ratina and Mandrekar 2017). CYP2E1 catalyzes the metabolism of procarcinogens such as *N*-nitrosamines, aniline, vinyl chloride, benzene, styrene, and urethane. There is consistency in different sets of epidemiological data showing a dose-response correlation between chronic alcohol consumption and increase in the risk for breast cancer (Baan et al. 2007; Schwab 2011). The full impact of chronic alcohol use on cancer is yet to be elucidated fully. CYP2E1 is involved in the metabolism of drugs such as acetaminophen, isoniazid, chlorzoxazone and fluorinated anesthetics, hormones, and xenobiotic toxins (Schmidt and Taylor 1987).
3. Pharmacodynamic interactions

   Alcohol can potentiate the sedative effect of opioids (morphine, codeine, meperidine),

tricyclic antidepressants, antihistamine medications, antipsychotics, benzodiazepines, and hypnotics. Specific alcohol drinks such as beer and wine could lead to hypertensive crises even in moderate amounts, if combined with monoamine oxidase inhibitors, especially if taken also with foods containing tyramine (cheese, some processed meat) or specific alcohol drinks such as beer and wine with monoamine oxidase inhibitors that could lead to hypertensive crises, even in moderate amounts. Dizziness and risk of falls could be exacerbated during acute intoxication with alcohol for someone taking antihypertensive medications such as nitroglycerin, hydrazaline, or with medications for Parkinson's disease, such as methyldopa.

Nicotine

Tobacco has toxic effects on virtually all organs in the human body. These effects are generally caused by substances other than nicotine, but still, this is the major addictive substance in tobacco smoke. Nicotine is a tertiary amine, found in the tobacco plant. Both (S)- and (R)-nicotine bind stereoselectively to nicotinic cholinergic receptors (nAChRs) with the (S)-type being a more potent aAChR agonist. When nicotine enters the body (with the cigarette smoke, when dried tobacco is sniffed or when tobacco leaves are chewed), it quickly enters the bloodstream and reaches two major sites where it exerts its physiological effects – the brain and the adrenal gland.

In the brain, the stimulation of CNS nAChRs leads to activation of dopaminergic transmission (also within the "reward circle" – midbrain – nucleus accumbens and further activation of parts of the limbic system, including cortical areas). The activation of nAChRs leads to activation of other receptor pathways, including acetylcholine, norepinephrine, serotonin, GABA, glutamate, and endorphins (Benowitz 2009). In tobacco smoking, dopamine release in the brain is facilitated by nicotine-mediated augmentation of glutamate release and by inhibition of GABA release (Benowitz 2009). Moreover, in chronic tobacco smoking, inhibition of monoamine oxidase (MAO) A and B is observed, which is associated with further increase of dopamine and norepinephrine in the synaptic cleft. Therefore, the two major pathways of nicotine addiction are (1) the increase in dopamine and norepinephrine in the CNS (the nAChR- mediated stimulation of the brain reward function) and (2) activation of the limbic system.

The second binding site of nicotine are the ganglion-type nAChRs in the chromaffin cells within the adrenal medulla with further epinephrine release leading to increased pulse rate, blood pressure, and contrainsular effects (Benowitz 2009).

The pharmacological interventions in nicotine addiction are directed against stopping the tobacco smoking, generally because of the undesired effects of other smoke ingredients. It consists generally of nicotine replacement – via transdermal patches, etc., and nicotine-blocking treatment (Quinn et al. 1997).

Bath Salts (Synthetic Cathinones)

Synthetic cathinones are phenylalkylamine derivatives chemically similar to the natural monoamine alkaloid cathinone (benzylethanamine, β-keto amphetamine) derived from the plant khat (*Catha edulis*). These substances were first synthesized approximately a century ago but became popular as recreational drugs in the first decade of the twenty-first century. The commonly abused synthetic cathinones (called "bk-amphetamines" for their beta-ketone moiety) resemble amphetamine in their chemical structure and mode of action (including binding to the monoamine transporters and monoamine release, reuptake and signaling within the brain, modulation of serotonin action).

The natural alkaloid cathinone is (*S*)-2-amino-1-phenyl-1-propanone – a beta-ketone amphetamine analogue. It is found in the fresh leaves of the khat plant. Khat leaves are popular for recreational purposes in the Middle East, particularly in

Yemen. The intake of cathinone has sympathomimetic effect, close to that of amphetamine – euphoria, alertness, and increase in pulse rate and blood pressure. Well-known synthetic cathinones are mephedrone, methedrone, methylenedioxypyrovalerone, methylone, butylone, dimethylcathinone, ethcathinone, ethylone, fluoromethcathinone, and pyrovalerone.

The first cathinone derivate – methcathinone – was synthesized in 1928, and 1 year later mephedrone was synthesized. The only synthetic cathinone currently approved for medical purposes is bupropion. Methcathinone was used for the treatment of depression in the 1930s–1940s and was administered for recreational purposes until the late 1990s. The administration of pyrovalerone for chronic fatigue and obesity has been investigated, but due to abuse and dependency, the drug was withdrawn.

In the beginning of the twenty-first century, there was a renaissance of synthetic cathinones as recreational drugs, initially in the UK and subsequently in the USA. These substances have been often referred to as "designer drugs" and have been sold under the name of "bath salts." The commonly sold bath salts in Europe usually contain mephedrone, and in the USA methylenedioxypyrovalerone, along with the different derivatives of pipradrol and pyrovalerone.

The main routes of administration of synthetic cathinones are nasal insufflations (snorting) or oral ingestion, but rectal, gingival, and inhalation delivery and intramuscular and intravenous injection have also been described. Moreover, synthetic cathinones are often administered in combination with other recreational substances, with or without alcohol. Their psychoactive effect appears 10–15 min after the intake and is expected to last for ½–4 h, depending on the route of administration.

As it was mentioned above, all synthetic cathinones are phenylalkylamine derivatives with bk-moiety that resemble amphetamines, so they can modulate the levels and action of biogenic amines (stimulant, sympathomimetic effects) and serotonin (effects on the mood and appetite, psychoactive effects) in the brain. Cathinones have higher polarity compared to amphetamines, and therefore they have lower penetration through the blood-brain barrier. Their pharmacodynamic and pharmacokinetic properties in humans are not well understood, and the majority of pharmacological data are derived from animal models and in vitro studies. Their effect is known to be due to increased synaptic concentration of dopamine, norepinephrine, and serotonin in the synaptic space via two major mechanisms:

- Inhibition of the monoamine uptake transporters with subsequent inhibition of the synaptic clearance of monoamines
- Release of the neurotransmitters from intracellular depots through the alteration of vesicular pH and concomitant inhibition of the vesicular monoamine transport VMAT2 receptor, responsible for the monoamine reuptake in the vesicles

From a neurobiological point of view, the main factor for the self-administration behavior, abuse, and addiction is the mesolimbic dopamine transmission (Baumann et al. 2014). The mechanisms of action of the following synthetic cathinones have been elucidated, at least in animal models (Baumann et al. 2014; Prosser and Nelson 2012):

- Methylone: inhibition of norepinephrine and dopamine via the suppression of monoamine uptake transporters (equally potent to that of methamphetamine and MDMA), inhibition of VMAT2 receptor (less potent than methamphetamine and MDMA), competitive for norepinephrine uptake and non-competitive for serotonin and dopamine), and reverse transport of neurotransmitters from the nerve terminal to the synapse (analogous to that in methamphetamine intake)
- Mephedrone: the same mechanisms of action, but less potent in increasing serotonin brain levels and faster returning of mediator levels to the baseline compared to MDMA and amphetamine
- Pyrovalerone: inhibition of norepinephrine and dopamine and little effect on serotonin reuptake with the S-enantiomer of pyrovalerone possessing higher biological activity

Based on their similarity to amphetamines, the effects of synthetic cathinones resemble much to the effects of amphetamine derivatives. The most important adverse effects of their intake are related to their sympathomimetic and serotoninergic effects: palpitations; increased blood pressure due to vasospasm; epistaxis; abdominal pain; severe rhabdomyolysis due to vasoconstriction with dehydration due to decreased sensation of thirst and increased physical activity (i.e., dancing); mydriasis with vision abnormalities; increased activity of the central nervous system with agitation, aggression, paranoia, and delusions; tremor; seizures; tachypnea; dyspnea; diaphoresis; and fever. Cases of hyponatremia in bath salt intake have been reported that are thought to be related to overhydration (as in amphetamine/MDMA users who voluntarily increase the intake of fluids because of the risk of dehydration) plus changes in antidiuretic hormone secretion [Prosser]. Cases of acute renal failure have been described that are thought to be due to rhabdomyolysis, dehydration, and severe vasoconstriction. Liver failure in cathinone users is thought to be associated with vasoconstriction, thrombosis, and concomitant use of other hepatotoxic substances. It is unknown whether cathinones have direct hepatotoxic effect, like MDMA, mediated by direct mitochondrial toxicity with oxidative modification of mitochondrial proteins (Moon et al. 2008).

Synthetic cathinones are often consumed with alcohol. Studies show that the concomitant intake of mephedrone and alcohol in rodents leads to enhancement of the psychostimulant effect via additional increase in synaptic dopamine levels (A.Cuidad-Roberts et al., 2015). This alcohol-induced potentiation of the effect of cathione can be blocked by haloperidol, but not by ketanserin.

Concerning the addiction and withdrawal, there has been no systematic research on these processes in cathinone-abusing humans. Observational studies have shown that synthetic cathinones are addictive (Prosser and Nelson 2012), with addiction/dependence symptoms and social dysfunction. Abusers report craving to repeat or increase the dose of mephedrone. Although no physical effect has been reported so far. Severe psychological dependence may be present, including depression, anxiety, and craving for continuous use, even without any reported physical effects (Prosser, Nelson JS, Nelson LS 2012). As cathinones are analogues of amphetamines, one could expect the development of marked and long-lasting changes in brain sympathies and serotoninergic receptor systems.

The treatment of acute cathinone intoxication is generally supportive – hydration and correction of electrolyte disturbances; stimulation of diuresis; increase of gastrointestinal clearance; oral administration of absorptive agents, adrenergic antagonists (beta-blockers), sedatives (benzodiazepines as in amphetamine and cocaine intoxication to counteract monoamine release and reuptake inhibition), and anticonvulsants; treatment of hyperpyrexia; treatment of rhabdomyolysis (saline infusions, intravenous loop diuretics, urine alkalization, mannitol infusions, corticosteroids, dialysis); gastroprotection; antithrombotic prophylaxis; etc.

Pharmacodynamic Interactions Between Addictive Substances

As all addictive drugs follow the same neurotransmitter pathways and the neuromediator systems tend to interact closely with each other and with endocrine signaling, the addictive drugs tend to show significant pharmacodynamic interactions that can have detrimental consequences for the body and the brain. These interactions are of importance for the treatment of drug addiction and withdrawal because of the need to decrease anxiety and CNS overexcitement via decrease in receptor sensitivity and/or inhibition of other receptor systems of the same neurons, affected by the addictive substance.

The pharmacodynamic interactions between different psychoactive medications can be explained by four major phenomena: effect on the same receptor systems, presence of both types of receptors on the same neuron, presence of the two types of specific receptors on contacting/adjacent neurons that interact, and, last but not least, interaction on subcellular and intracellular levels (i.e., action on the same second

messenger or enzyme systems within the cell) (Quinn et al. 1997).

Shown below are some examples of pharmacodynamic interactions between psychoactive substances.

Opioids are known to increase the sedative effects of benzodiazepines, and vice versa. Moreover, naloxone decreases this effect. This could be explained by at least one of the following three phenomena (Quinn et al. 1997):

- Presence of both GABAA (for benzodiazepines) and opioid receptors on the same neurons
- Presence of GABAA and opioid receptors on adjacent but related neurons
- Pharmacodynamic synergism due to changes in intracellular mediators and second messengers – in this case changes in cAMP levels and/or GABAA receptor phosphorylation via cAMP-dependent process

The sedative effects of BZDs have been shown to be diminished by naloxone, probably by downstreaming of GABAA receptors (Quinn et al. 1997).

The sedative effects of GABAA pathway-mediated substances (opioids and benzodiazepines) are also increased by alcohol. Acute alcohol ingestion also potentiates GABAA. Antidepressants, antihistamines, and anticonvulsants that interact with GABAA mediation also show additive synergism with opioids, benzodiazepines, and alcohol. Cocaine and opioids, especially heroin, also tend to interact mainly in a pharmacodynamic way (Quinn et al. 1997) – interactions between opioid and dopaminergic pathways and changes in cAMP intracellular levels.

Amphetamine and its derivatives possess MAO inhibitory properties and therefore tend to interact with MAO inhibitors in a potentially fatal way. The co-administration of both types of drugs leads to marked adrenergic activation with extreme elevation of blood pressure and pulse rate. This could result in confusion, coma, and death (Quinn et al. 1997).

The combined intake of dissociative anesthetics with other addictive and psychoactive substances can also be extremely dangerous. The concomitant intake with antidepressants can cause serotonin syndrome, and in combination with stimulants, they can increase heart rate and the blood pressure. In combination with sedatives and alcohol, they can suppress breathing.

Synthetic cathinones are often consumed with alcohol, and it tends to increase their psychostimulant effect (Ciudad-Roberts et al. 2015), probably via additional increase in synaptic dopamine levels. Their effects are probably stimulated by all illicit substances that increase dopaminergic mediation.

Conclusion

The intake of psychoactive substances has followed mankind since the dawn of human history. These substances have been taken for religious and recreational purposes, as sedative or stimulant medications, and for the treatment of somatic conditions (i.e., gastrointestinal mobility disorders, for pain, etc.) and central and/or peripheral nervous system diseases and conditions. Their pharmacodynamic profiles are of crucial importance for the understanding of their pharmacological and toxic effects, addiction, possible drug interaction, and treatment of withdrawal. Pharmacodynamic drug interactions are of particular importance because of the high prevalence of multidrug abuse. The newer “designer drugs” are an emerging and potentially serious problem, because of the affection of many receptor and signaling pathways and the potential to interact with virtually all substances that affect dopaminergic mediation.

Therefore, the good understanding of the pharmacodynamic characteristics of both older and newer addictive substances will aid the diagnostic and therapeutic process in the everyday clinical practice.

References and Further Reading

Anis NA, Berry NSC, Burton NR, Lodge D (1983) The dissociative anaesthetics, ketamine and phencyclidine, selectively reduce excitation of central mammalian neurons by *N*-methyl-aspartate. Br J Pharmacol 79:565–575

Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, Altieri A, Cogliano V, WHO International Agency for Research on Cancer Monograph Working Group (2007) Carcinogenicity of alcoholic beverages. Lancet Oncol 8:292–293

Baumann MH, Soris E, Watterson LR et al (2014) Bath salts, spice, and related designer drugs: the science behind the headlines. J Neurosci 34(46):15150–15158

Baumeister D, Barnes G, Giaroli G, Tracy D (2014) Classical hallucinogens as antidepressants? A review of pharmacodynamics and putative clinical roles. Ther Adv Psychopharmacol 4(4):156–169

Benowitz NL (2009) Pharmacology of nicotine: addiction, smoking-induced disease and therapeutics. Annu Rev Pharmacol Toxicol 49:57–71

Blessing WW, Seaman B, Pedersen NP, Ootsuka Y (2003) Clozapine reverses hyperthermia and sympathetically mediated cutaneous vasoconstriction induced by 3,4 - methylenedioxymethamphetamine (ecstasy) in rabbits and rats. J Neurosci 23(15):6385–91

Brox B, Ellenbroek B (2018) A genetic reduction in the serotonin transporter differentially influences MDMA and heroin induced behaviours. Psychopharmacology (Berl) 235(7):1907–1914

Calipari ES, Ferris MJ (2013) Amphetamine mechanisms and actions at the dopamine terminal revisited. J Neurosci 33(21):8923–8925

Campbell JE, Cohall D (2017) Pharmacodynamics – a pharmacognosy perspective, Chapter 26. In: Pharmacognosy. Fundamentals, applications, strategies. Academic, Boston, pp 513–525

Capasso R, Borrelli F, Capasso F et al (2006) The hallucinogenic herb Salvia divinorum and its active ingredient salvinorin A inhibit enteric cholinergic transmission in the Guinea-pig ileum. Neurogastroenterol Motil 18(1):69–75

Casey ER, Scott MG, Tang S, Mullins ME (2011) Frequency of false positive amphetamine screens due to bupropion using the Syva EMIT II immunoassay. J Med Toxicol 7(2):105–108

Ciudad-Roberts A, Camarasa J, Ciudad CJ et al (2015) Alcohol enhances the psychostimulant and conditioning effects of mephedrone in adolescent mice; postulation of unique roles of D receptors and BDNF in place preference acquisition. Br J Pharmacol 172:4970–4984

de la Torre R, Farré M, Roset PN, Pizarro N, Abanades S, Segura M, Segura J, Camí J (2004) Human pharmacology of MDMA: pharmacokinetics, metabolism, and disposition. Ther Drug Monit 26(2):137–144

Dean A (2006) Illicit drugs and drug interactions. Pharmacist 25(9):684–689. www.erowid.org/psychoactives

DeVane CL (2016) Clinical pharmacokinetics and pharmacodynamics of anxiolytics and sedative/hypnotics. In: Jann M, Penzak S, Cohen L (eds) Applied clinical pharmacokinetics and pharmacodynamics of psychopharmacological agents. Adis, Cham, pp 247–266

Dumas EO, Pollack GM (2008) Opioid tolerance: a pharmacokinetic/pharmacodynamic perspective. AAPS J 10(4):537–551

Edenberg HJ (2007) The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. Alcohol Res Health 30(1):5–13

Feng XQ, Zhu LL, Zhou Q (2017) Opioid analgesics-related pharmacokinetic drug interactions: from the perspectives of evidence based on randomized controlled trials and clinical risk management. J Pain Res 10:1225–1239. https://doi.org/10.2147/JPR.S138698eCollection2017.

Ganetsky M et al (2013) Effect of excipients of acetaminophen metabolism and its' implications for prevention of liver injury. J Clini Pharmacol 53(4):413–20

Galleli L, Gratteti S, Siniscalchi A, Cione E, Sirico S, Seminara P, Caroleo MC, De Sarro G (2017) Curr Drug Abuse Rev 10(1):25–30

Gay GR, Inaba DS, Sheppard CW, Newmeyer JA (1975) Cocaine: history, epidemiology, human pharmacology, and treatment. A perspective on a new debut for an old girl. Clin Toxicol 8(2):149–178

Ghelardini C, Mannelli LDC, Bianchi E (2015) The pharmacological basis of opioids. Clin Cases Miner Bone Metab 12(3):219–221

Gosnell BA, Kotz CM, Billington CJ, Levine AS (2013) Handbook of biologically active peptides, 2nd Ed. In: Abbaj Kastin (Ed), Ingestive Peptides. Academic Press, Elsevier, pp 1149–1154

Griffin AE III, Kaye AM, Bueno FR, Kaye AD (2013) Benzodiazepine pharmacology and central nervous system – mediated effects. Ochsner J 13(2):214–223

Gurusamy U, Shewade DG (2014) Chapter 46: Pharmacogenomics in India, Handbook of Oharmacogenomics nad Stratified Medicine, Elsevier

Harrison AA, Liem YT, Markou A (2001) Fluoxetine combined with a serotonin-1A receptor antagonist reversed reward deficits observed during nicotine and amphetamine withdrawal in rats. Neuropsychopharmacology 25(1):55–71

Härtel-Petri R, Krampe-Scheidler A, Braunwarth WD, Havemann-Reinecke U, Jeschke P, Looser W, Mühlig S, Schäfer I, Scherbaum N, Bothe L, Schaefer C, Hamdorf W (2017a) Evidence-based guidelines for the pharmacologic management of methamphetamine dependence, relapse prevention, chronic methamphetamine-related, and comorbid psychiatric disorders in post-acute settings. Pharmacopsychiatry 50(3):96–104

Härtel-Petri R, Krampe-Scheidler A, Braunwarth WD, Havemann-Reinecke U, Jeschke P, Looser W, Mühlig S, Schäfer I, Scherbaum N, Bothe L, Schaefer C, Hamdorf W (2017b) Evidence-based

guidelines for the pharmacologic management of methamphetamine dependence, relapse prevention, chronic methamphetamine-related, and comorbid psychiatric disorders in post-acute settings. Pharmacopsychiatry 50:96–104

Higuchi S, Matsushita S, Murayama M et al (1995) Alcohol and aldehyde dehydrogenase polymorphisms and the risk of alcoholism. Am J Psychiatry 152:1219–1221

Haufroid V et al (2002) Interest of genotyping and phenotyping of drug-metabolizing enzymes for the interpretation of biological monitoring of exposure to styrene. Pharmacogenetics 12(9):691–702

Jevtovic-Todorovic V, Todorovic SM, Mennerick S et al (1998) Nitrous oxide (laughing gas) is a NMDA antagonist, neuroprotector and neurotoxin. Nat Med 4(4):460–463

Kavannagh D, Goodship THJ, Richards A (2006) Atypical haemolytic uraemic syndrome. Br Med Bull 77–78:5–22

Kim HR, Son BH, Lee SY et al (2012) The role of p53 in marijuana smoke condensates-induced genotoxicity and apoptosis. Environ Health Toxicol 27:c2012017

Kishi T, Matsuda Y, Iwata N, Correll CU (2013) Antipsychotics for cocaine or psychostimulant dependence: systematic review and meta-analysis of randomized, placebo-controlled trials. J Clin Psychiatry 74(12): e1169–e1180

Koob GF, Le Moal M (2006) Neurobiology of Addiction. Academic Press, Imprint of Elsevier 92:101–4495

Lange RA, Cigarroa RG, Flores ED, McBride W, Kim AS, Wells PJ, Bedotto JB, Danziger RS, Hillis LD (1990) Potentiation of cocaine-induced coronary vasoconstriction by beta-adrenergic blockade. Ann Intern Med 112(12):897–903

Li JH, Lin LF (1998) Genetic toxicology of abused drugs: a brief review. Mutagenesis 13(6):557–565

Lindsey WT, Stewart D, Childress D (2012) Drug interactions between common illicit drugs and prescription therapies. Am J Drug Alcohol Abuse 38(4):334–343

McGraw J (2014) Chapter 16: CYP450 and Ethnicity, Handbook of Pharmacogenomics and Stratified Medicine, Elsevier

McCance-Katz EF, Jatlow P, Rainey P, Friedland G (1998) Methadone effects on zidovudine (AZT) disposition (ACTG 262). J Acquir Immune Defic Syn Hum Retrovirol 18:435–443

McGuire P et al (2018) Cannabidiol (CBD) as an adjunctive therapy in schizophrenia: a multicenter randomized controlled trial. Am J Psychiatry 175(3):225–231

Medicines and Healthcare products Regulatory Agency (2016) Citalopram: suspected drug interaction with cocaine. Drug Saf Update 9(12):2

Mittal B, et al, Advances in Clinical Chemistry; Chapter Four: Cytochrome P450 in Cancer Susceptibility and Treatment. Volume 71 Ed by Makowski GS, (2015), Elsevier.

Mollereau C, Roumy M, Zajac JM (2005) Opioid-modulating peptides: mechanisms of action. Curr Top Chem 5(3):341–355

Moon KH, Upreti VV, Yu LR et al (2008) Mechanism of 3,4-methylenedioxymethamphetamine (MDMA, Ecstasy)-mediated mitochondrial dysfunction in rat liver. Proteomics 8(18):3906–3918

Nestler EJ (2005) The neurobiology of cocaine addiction. Sci Pract Perspect 3(1):4–10

Pani PP, Trogu E, Vecchi S, Amato L (2011) Antidepressants for cocaine dependence and problematic cocaine use. Cochrane Database Syst Rev (12):CD002950

Pasternak GW, Pan YX (2013) Mu opioids and their receptors: evolution of a concept. Pharmacol Rev 65(4):1257–1317

Pertwee RG (2005) Pharmacological actions of cannabinoids. Handb Exp Pharmacol 168:1–51

Pomara C, Cassano T, D'Errico S et al (2012) Data available on the extent of cocaine use and dependence: biochemistry, pharmacologic effects and global burden of disease of cocaine abusers. Cur Med Chem 19(33):5647–5657

Powledge TM (1999) Addiction and the brain. Bioscience 49(7):513–519

Prosser JM, Nelson LS (2012) The toxicology of bath salts: a review of synthetic cathinones. J Med Toxicol 8(1):33–42

Quinn DI, Wodak A, Day RO (1997) Pharmacokinetic and pharmacodynamic principles of illicit drug use and treatment of illicit drug users. Clin Pharmacokinet 33(5):344–400

Ratina A, Mandrekar P (2017) Alcohol and cancer: mechanisms and therapies. Biomolecules 7:(3)

Reece AS (2009) Chronic toxicology of cannabis. Clin Toxicol 47:517–524

Reece AS, Hulse GK (2016) Chromothripsis and epigenomics complete causality criteria for cannabis and addiction-connected carcinogenicity, congenital toxicity and heritable genotoxicity. Mutat Res 789:1–11

Richards JR, Garber D, Laurin EG, Albertson TE, Derlet RW, Amsterdam EA, Olson KR, Ramoska EA, Lange RA (2016) Treatment of cocaine cardiovascular toxicity: a systematic review. Clin Toxicol (Phila) 54(5):345–364

Rietjens SJ, Hondebrink L, Westerink RHS, Meulenbelt J (2012) Pharmacokinetics and pharmacodynamics of 3,4-methylenedioxymethamphetamine (MDMA): interindividual differences due to polymorphisms and drug–drug interactions. Crit Rev Toxicol 42 (10): 854–76

Robison AJ, Nestler EJ (2011) Transcriptional and epigenetic mechanisms of addiction. Nat Rev Neurosci 12(11):623–637

Rudnik G, Wall SC (1992) The molecular mechanism of "ecstasy" [3,4-methylenedioxymethamphetamine (MDMA)]: serotonin transporters are targets for MDMA-induced serotonin release. Proc Natl Acad Sci 89:1817–1821

Ruiz-Garcia A, Bermejo M, Moss A, Casabo VG (2008) Pharmacokinetics in drug discovery. J Pharm Sci 97(2):654–690

Schmidt CJ, Taylor VL (1987) Depression of rat brain tryptophan hydroxylase activity following the acute administration of methylenedioxymethamphetamine. Biochem Pharmacol 36:4095–4102

Schwab M (2011) Encyclopedia of cancer, 3rd edn. Springer, Berlin

Sharma P, Murthy P, Bharath S (2012) Chemistry, metabolism, and toxicology of cannabis: clinical implications. Iran J Psychiatry 7(4):149–156

Sleigh J, Harley M, Voss L, Denny B (2014) Ketamine – more mechanisms of action than just NMDA blockade. Trends Anest Crit Care 4(2–3):76–81

Status and Trend Analysis of Illicit Drug Markets (2015) World drug report. http://www.unodc.org/documents/wdr2015/WDR15_Drug_use_health_consequences.pdf

Stein C, Schäfer M, Machelska H (2003) Attacking pain at its source: new perspectives on opioids. Nat Med 9(8):1003–1008, Publisher Iztok-Zapad, 5 Stara Planina Str, 2nd Floor, Sofia, 1000, Bulgaria, EU; ISBN 978-954-9854-19-0, 248 pp

Tenev V (2008) Reference Book on Drug Interactions in Psychiatry and General medical practice. Bulgarian Psychiatric Association, Publisher Iztok-Zapad, 5 Stara Planina Str, 2nd Floor, Sofia, 1000, Bulgaria, EU; ISBN 978-954-9854-19-0, 248 pages

Volkow ND, Morales M (2015) The brain on drugs: from reward to addiction. Cell 162:712–725

Vonlaufen A, Wilson JS, Pirola RC, Apte MV (2007) Role of alcohol metabolism in chronic pancreatitis. Alcohol Res Health 30(1):48–54

Yubero-Lahoz S, Pardo R, Farré M, O'Mahony B, Torrens M, Mustata C et al (2011) Sex differences in 3, 4-methylenedioxymethamphetamine(MDMA; ecstasy)-induced MDMA, methamphetamine, and CYP2D6 cytochrome P450 2D6 inhibition in humans. Clin Pharmacokinet 50:319–329

Pharmacodynamic Evaluation: Ocular Pharmacology

7

Najam A. Sharif

Contents

N. A. Sharif (✉)
Global Alliances and External Research, Global Ophthalmology Research and Development, Santen Incorporated, Emeryville, CA, USA
e-mail: najam.sharif@santen.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_54

Abstract

The eye is a specialized organ that provides a relatively easy access for direct visualization of the different anatomical structures and assessment of the diseases associated with them. Despite this, however, the diagnosis and treatment of eye diseases has proven difficult over the years. Ocular surface diseases include allergic conjunctivitis, infection (viral, bacterial, and fungal), inflammation, and dry eye. Major anterior chamber and lens-associated disorders include cataracts, presbyopia, iritis/ureitis, elevated intraocular pressure-associated glaucoma and pseudo-exfoliation glaucoma. Diseases that primarily effect the retina in the posterior segment of the eye include wet and dry age-related macular degeneration, diabetic macular edema and retinopathy, and glaucomatous optic neuropathy that involves the retinal sensory neurons (retinal ganglion cells) and their axons that form the optic nerve that connects the retina to the brain. Many decades of basic and applied research have resulted in the discovery and development of different types of pharmacological agents (small molecules), peptides, and antibodies that help clinically manage the various ocular disorders mentioned above. Recent advances in gene- and cellular-therapeutics, and production of suitable miniature devices, have also revolutionized ocular disease management. The pharmacotherapeutic and pharmacodynamic aspects of these modalities will be discussed here. This will include target protein localization, assessment of drug engagement with the target, and mechanism of action of the drug entities in cellular and whole-eye efficacy systems using normal and disease-based assays and animal models. Such in vitro screening and in vivo evaluation and the types of results obtained from such studies will be also described and discussed.

General Introduction

Due to the fact that this chapter aims to cover numerous types of eye diseases that have a diverse set of etiologies and disease pathways, it is the author's opinion that a standardized format is unsuitable. As such, the chapter will cover the key elements of the requisite headings and subheadings but without strict adherence. It is hoped that the format followed is acceptable.

The World Health Organization (WHO 2018), National Eye Institute (NEI 2014), and American Academy of Ophthalmology have estimated that ~250 million individuals, including 36 million blind people, have some form of eyesight impairment. Furthermore, it is believed that the number of blind people will increase to 38.5 million by 2020, and to 115 million by 2050. In the USA alone, the total economic burden related to vision loss is expected to reach ~USD 715 billion by 2050. In fact, chronic eye diseases are one of the main causes of vision loss globally, and an ~90% cases of visual impairment are due to such conditions, and indeed a large portion of these chronic ophthalmological disorders affect back of the eye (WHO 2018; NEI 2014; American Academy of Ophthalmology).

Eyesight is an extremely vital sense in animals and humans and is critical for survival in most species that sense and interact with the environment through ocular cues. In fact, most humans place the highest value, among the five senses, on eyesight. While not life-threatening, visual

impairment due to eye disorders has serious implications for quality of life for the patient, caregivers, and the society at large. As some of the eye diseases are age-related in terms of onset and severity, the incidence of ocular diseases continues to increase as the aging population on our planet increases. For example, cataracts were reported to impact >20 million people in 2010, and glaucoma afflicts >65 million people worldwide (Tham et al. 2014; Weinreb et al. 2014; Jonas et al. 2017) with similar numbers for age-related macular degeneration and related visual acuity disorders. Similarly, dry eye (15 million patients in the USA alone), diabetic retinopathy (>5 million patients), refractive errors (>4.1 million patients), and ocular allergies (>3 million patients/year in the USA) also have a high prevalence. These eye diseases, vascular eye disorders, and myopia, continue to cause a sustained undue suffering, medical burden and expense to the patients and the society (WHO 2018; NEI 2014; American Academy of Ophthalmology). Accordingly, discovery and development of therapeutic agents and devices to treat ocular disorders has gained prominence and importance, eliciting an appropriate heightened sense of urgency to find better treatment modalities and ultimately cures for these maladies.

Introduction to Eye Anatomy and Basic Physiology

Being a uniquely specialized sensory organ, it is important to briefly outline the key anatomical elements of the eye and how the eye encodes the light it receives into well-defined color-coded images for us to see (Fig. 1). Most of the eyeball is encased in a white relatively thick, tough, yet flexible fibrous tissue called the sclera, which provides protection from sharp objects and noxious chemicals and shapes the eyeball. The sclera at the

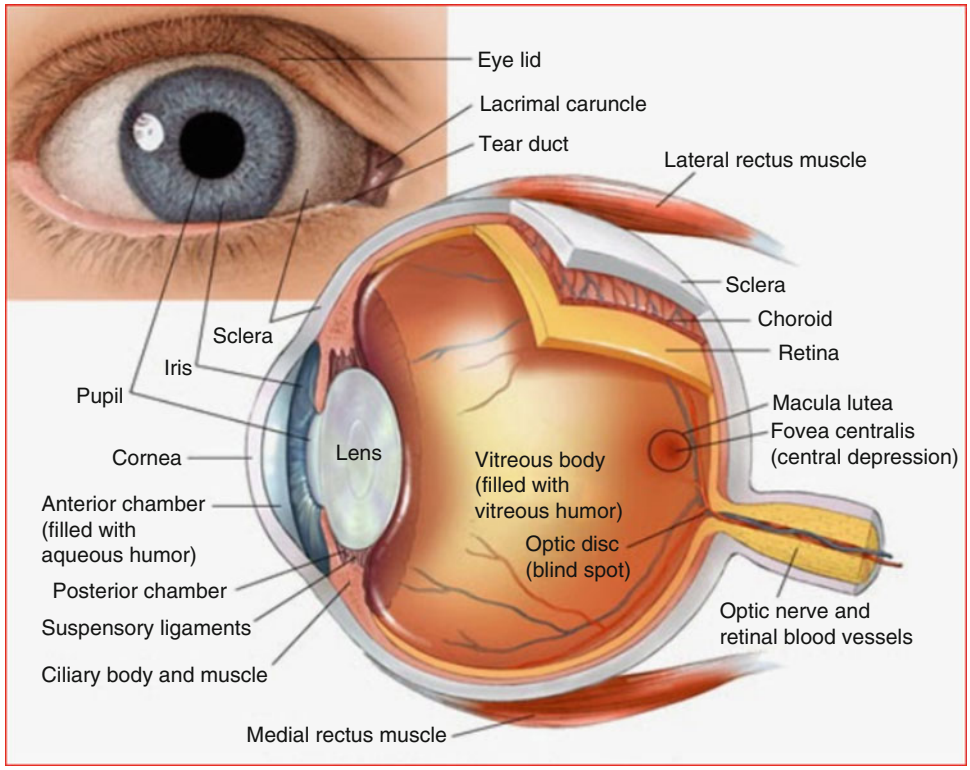


Fig. 1 Depiction of the anatomy and key structures of the human eye pertinent for the discussion of various eye diseases

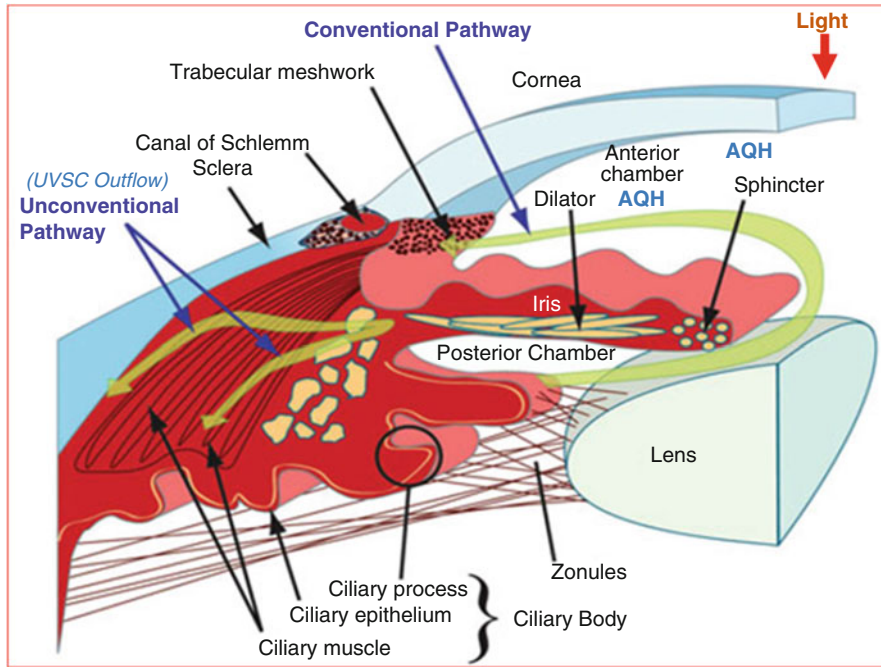


Fig. 2 Key elements of the anterior segment of the human eye are shown. Formation of the AQH by the ciliary processes, its flow from the latter in front of the lens, followed by its drainage from the anterior chamber via TM/SC and via the UVSC pathways is also depicted

front of the eyeball is further specialized and forms the transparent cornea that permits light to enter the eye and that also provides a barrier to airborne chemicals, noxious agents, and pathogens.

The eye is divided into anterior and posterior segments that are separated by the lens. The anterior chamber (ANC) contains a clear fluid (aqueous humor [AQH]), that contains various nutrients and oxygen, is generated by the ciliary epithelium within the ciliary body. As the AQH flows through the ANC, it nourishes the cells of the lens epithelium, corneal endothelium, and trabecular meshwork (TM) (Fig. 2). AQH also removes metabolites, dead cells, and other toxic waste as it drains through the TM/Schlemm's canal (SC) and into the venous circulation.

The posterior segment is filled with vitreous humor (VH), a gelatinous material, that does not turnover as much as the AQH. The AQH and VH, coupled with the sclera, provide the eyeball its unique shape and overall rigidity. In direct contact with the VH is a thin inner limiting membrane that isolates the retina from the VH but one that is fairly permeable to chemicals and gases. The retina is a highly specialized tissue being composed of multiple layers of cells that have unique functions in light perception and neuronal communication to the brain (Fig. 3). For the sake of brevity and focus, the most important cell types related to pathological aspects and treatment of retinal diseases include the retinal ganglion cells (RGCs) and their axons (connected with various forms of glaucoma; glaucomatous optic neuropathy (GON)), retinal pigment epithelial cells (RPE cells; connected with wet age-related macular degeneration (AMD) and diabetic retinopathy (DR), and photoreceptor cells (connected with retinitis pigmentosa (RP) and dry AMD) (Fig. 3). However, the interplay between many of these and other retinal cell types, for example, Muller glia and retinal interneurons, is also very important for visual perception. Thus, in humans a million or so RGC axons are bundled together to form the optic nerve that projects to the brain centers that are involved in visual perception.

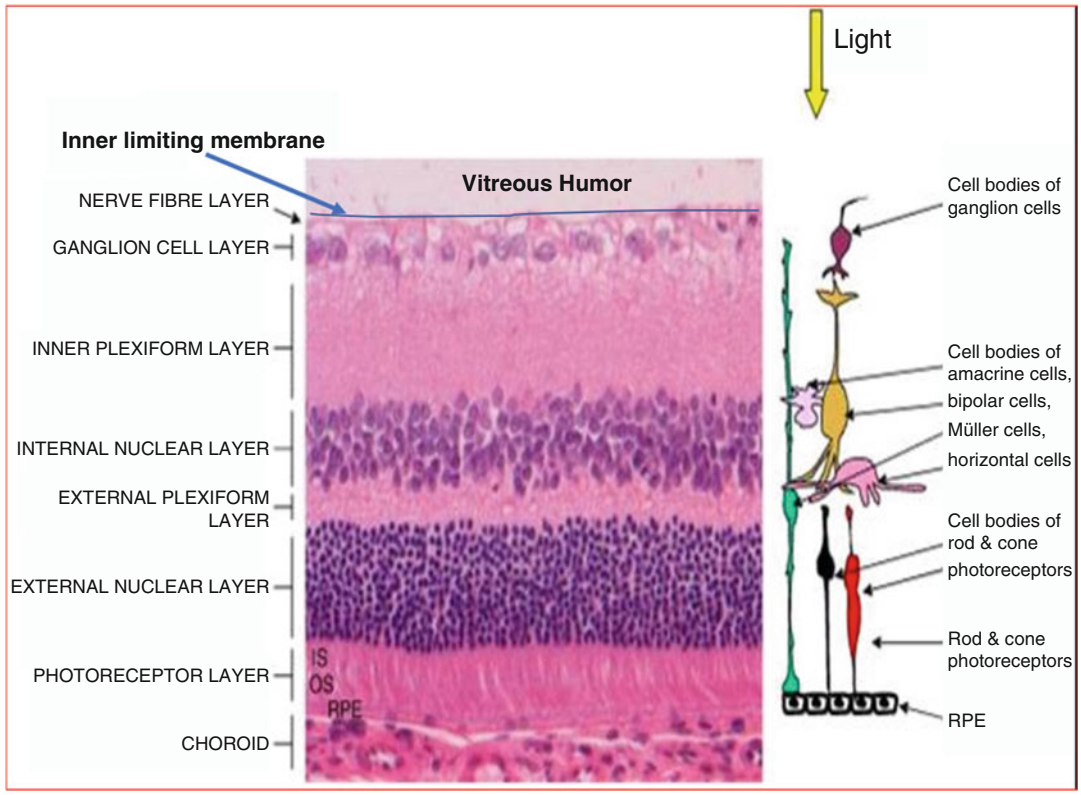


Fig. 3 Various layers of human retinal cells are shown indicating their relative positions and interconnections between them

Wrapped within the optic nerve are the central retinal artery and the central retinal vein that provide and remove, respectively, some of the blood to the retina. Another major artery supplying oxygenated blood to the retina is the posterior ciliary artery.

In terms of visual perception, light from the environment is first focused by the cornea and then passes through the pupil and is focused further by the lens onto the retina. Here, photoreceptors (rods and cones) then convert the light photons into membrane potential changes by closing of Na^+-channels and with resultant cellular hyperpolarization. This electrical change within the rods and cones decreases the release of the excitatory neurotransmitter glutamate (GLU) from their synaptic terminals onto bipolar cells. This changes the level of tonic electrical activity of bipolar cells that then release less GLU onto RGCs. In turn, since GLU activates many types of ionotropic and metabotropic receptor subtypes that modulate various Ca^{2+}-channels and a whole host of transient receptor potential channels, the electrical activity of the RGCs is well regulated. Also, since neighboring horizontal and amacrine cells modulate the activity of bipolar cells through release of inhibitory neurotransmitters such as gamma-butyric acid and glycine, the bipolar cells' ability to encode and transmit signals to the RGCs is extremely well coordinated and finely controlled. Finally, the RGCs integrate all the information received from bipolar cells and relay this as electrical impulses down their axons, within the optic nerve, to the lateral geniculate nucleus/superior colliculus within the thalamus of the mid-brain. From there, visual information is relayed to the visual cerebral cortex for final processing and generation of the exquisite images that we see. In this manner, the eyes represent windows for the brain.

Receptors, Ion-Channels Transporters, and Pharmacodynamics

In short, pharmacology is the study of drug action at a tissue and cellular level, while pharmacodynamics pertains to the study of mechanism of action of drugs at molecular, cellular, and tissue levels that reflects drug affinity and efficacy. In principle, ocular pharmacology is no different to pharmacology of other organs such as the heart, lungs, and brain. However, since the repertoire of receptors/ion-channels/transporters (target proteins; Alexander et al. 2017a, b, c, d) do not necessarily exist or function in exactly the same manner in each tissue/cell of these organs, it is worthwhile considering some of the basic aspects of pharmacology and their application to the eye. This necessitates the understanding of how neurotransmitters, hormones, cytokines, and other mediators and drugs exert their actions by interacting with receptors, enzymes, transporters, ion-channels, and nucleic acids that the cells possess (Alexander et al. 2017a, b, c, d). An important characteristic of most biologically and pharmacologically relevant molecules (whether they be small molecules, peptides, hormones, cytokines, or antibodies) is their affinity ("relative attraction") for the target protein. This affinity parameter is the dissociation constant (K_d or K_i) as determined by the ratio of (rate of dissociation/rate of association) of the ligand from the target protein. The K_d or K_i values are inversely proportional to the affinity of the ligand, thus a low K_d or K_i reflects a high affinity, usually represented as a concentration needed to occupy [or dissociate from half of the ligand binding sites of the target protein. Implied within the K_d and K_i values is a certain amount of specificity of the ligand for the receptor binding site such that mere adsorption onto the target protein can be discounted. Some examples of these pharmacological/pharmacodynamic parameters for ligands binding to various prostaglandin receptor subtypes can be seen in Tables 1 and 2.

Of the overall many druggable target proteins expressed by cells, receptors (39–41% of total), transporters (4–5%), and ion-channels (8–11%) (Alexander et al. 2017a, b, c, d) are primarily embedded in cellular membranes, while most kinase enzymes (22–24%) and proteases (8–11%) are located within the cytoplasm, and of course nucleic acids and nuclear receptors (2%) are located within the cell nucleus. The interplay among these targets and their signal transduction mechanisms, including positive or negative cooperativity, feed-forward or feedback pathways, allows the cellular machinery to dampen, amplify, or subtly modulate the response mechanisms of the cells and tissues to the pharmacological agent, thereby providing extra fine control and specificity. Added to this complexity is the ability of endogenous ligands and exogenous agents to behave either as full-agonists (producing a maximal biological response), or partial-agonists (only producing a submaximal response; e.g., Fig. 4, left panel), or inverse-agonists (inducing a response opposite to an agonist), or antagonists (blocking the actions of an agonist) (e.g., Fig. 4, right panel). Furthermore, there are instances when partial agonists at certain concentrations and under certain conditions can behave as antagonists (Griffin et al. 1999; Sharif and Klimko 2019), and this property can be exploited in the realm of disease management when truly bona fide high potency and high affinity antagonists are not available or are unsuitable as therapeutic drugs.

In general terms, receptors are functional transmembrane proteins that are coupled to signal transduction components such as G-proteins (or kinase enzymes) coupled to certain catabolizing enzymes on the inner leaflet of the cell membrane. Activation of the receptor by an agonist ligand (such as norepinephrine or prostaglandin $F_{2\alpha}$) changes the conformation of the receptor to an active state that triggers specific G-protein(s) and the associated enzyme(s) to generate intracellular second messengers such as cAMP (by activation of G_s-linked to adenylyl cyclase [AC] to increase cAMP and via G_i to inhibit AC and Ca^{2+-}channels, and open K^+-channels), cGMP, inositol phosphates, and diacylglycerol (by activation of G_q-linked to phospholipase C) (Alexander et al. 2017a, b, c, d) (Fig. 5). These second messengers amplify the signal transduction by activating protein kinases, opening/closing

Table 1 Receptor binding inhibition constants and relative receptor selectivities for some key natural PGs

Natural PG	PG binding inhibition constants (K_i, nM) and receptor selectivity (x)							
	DP	EP_1	EP_2	EP_3	EP_4	FP	IP	TP
PGD_2	**81 ± 5**	>19,000 (x 234)	2,973 ± 100	1,115 ± 118 (x 14)	2,139 ± 180 (x 26)	2,500 ± 760 (x 31)	>140,000 (x 1,728)	>35,000 (x 432)
PGE_2	>10,000 (x 667 *vs* EP_1)	**26 ± 10**	**4.9 ± 0.5**	**3 ± 0.2**	**0.9 ± 0.03**	3,400 ± 710 (x 227 *vs* EP_1)	53,708 ± 2,136 (x 3,581 *vs* EP_1)	>10,000 (x 667 *vs* EP_1)
$PGF_{2\alpha}$	18,000 ± 6,460 (x 138)	594 ± 12 (x 5)	964 ± 64	24 ± 8 (x 0.2)	433 ± 25 (x 3)	**130 ± 6**	≥ 50,000 (x 385)	≥ 190,000 (x 1,462)
PGI_2	3,537 (x 3)	>15,000 (x 11)	Nd	5,375 ± 1,394 (x 4)	8,074 ± 254 (x 6)	>86,000 (x 62)	**1,398 ± 724**	>65,000 (x 46)

Data are means ± SEM from 3–141 experiments. The bolded values represent the drug inhibition constant (K_i) of the PG for its primary preferred receptor. In the case of PGE_2 it is clear that it lacks selectivity between the subtypes of the EP receptors. The K_i value is inversely related to affinity of the PG for the receptor, i.e., the smaller the K_i value the greater the affinity. The values in parentheses denote the relative selectivity of the PG for its preferred (cognate) receptor compared to its affinity for other PG receptors
Nd not determined (Sharif et al. 2003)

Table 2 PG receptor inhibition constants and relative selectivities of synthetic PG receptor analogs of the FP-class

PG Analog	PG receptor binding inhibition constants (K_i, nM) and FP receptor selectivity (x)							
	DP Receptors	EP_1 Receptors	EP_2 Receptors	EP_3 Receptors	EP_4 Receptors	FP Receptors	IP Receptors	TP Receptors
Travoprost acid ((+)-Fluprostenol)	52,000 ± 7,200 (x 1,486)	9,540 ± 1,240 (x 273)	Nd	3,501 ± 461 (x 100)	41,000 ± 2,590 (x 1,171)	**35 ± 5**	≥ 90,000 (x 2,571)	≥ 121,000 (x 3,457)
(±)-Fluprostenol	>50,000 (x 510)	12,300 ± 1,240 (x 126)	>100,000#	4,533 ± 597 (x 46)	14,400 ± 1,550 (x 147)	**98 ± 9**	>60,500 (x 617)	121,063 ± 20,714 (x 1,235)
Bimatoprost acid (17-phenyl-$PGF_{2\alpha}$)	>90,000 (x 1,084)	95 ± 27 (x 1)	Nd	387 ± 126 (x 5)	25,700 ± 2,060 (x 310)	**83 ± 2**	>100,000 (x 1,205)	>77,000 (x 928)
Latanoprost acid (PHXA85)	≥ 20,000 (x 204)	2,060 ± 688 (x 21)	39,667 ± 5,589#	7,519 ± 879 (x 77)	75,000 ± 2,830 (x 765)	**98 ± 11**	≥ 90,000 (x 918)	≥ 60,000 (x 612)
Bimatoprost (Amide)	>90,000 (x 14)	19,100 ± 1,450 (x 3)	Nd	>100,000 (x 16)	>100,000 (x 16)	**6,310 ± 1,650**	>100,000 (x 16)	>100,000 (x 16)
Unoprostone (UF-021; Acid)	>43,000 (x 7)	11,700 ± 2,710 (x 2)	Nd	≥ 22,000 (x 4)	15,200 ± 3,500 (x 3)	**5,900 ± 710**	>30,000 (x 5)	>30,000 (x 5)
S-1033 (Na^+-salt)	90,000 (x 4)	13,500 ± 1,670 (x −2)	Nd	≥ 77,000 (x 4)	6,650 ± 610 (x −3)	22,000 ± 2,600	>30,000 (x 1)	>30,000 (x 1)

Data are means ± SEM from 3–8 experiments. The bolded values represent the drug inhibition constant (K_i) of the PG for its primary preferred receptor, except for S-1033 which is left unbolded since it appears to possess a higher affinity for the EP_4 receptor than the FP receptor. The K_i value is inversely related to affinity of the PG for the receptor, i.e., the smaller the K_i value, the greater the affinity. The values in parentheses denotes the relative FP-receptor-selectivity of the PG analog compared to its affinity for the other PG receptors

Nd not determined (Sharif et al. 2003)

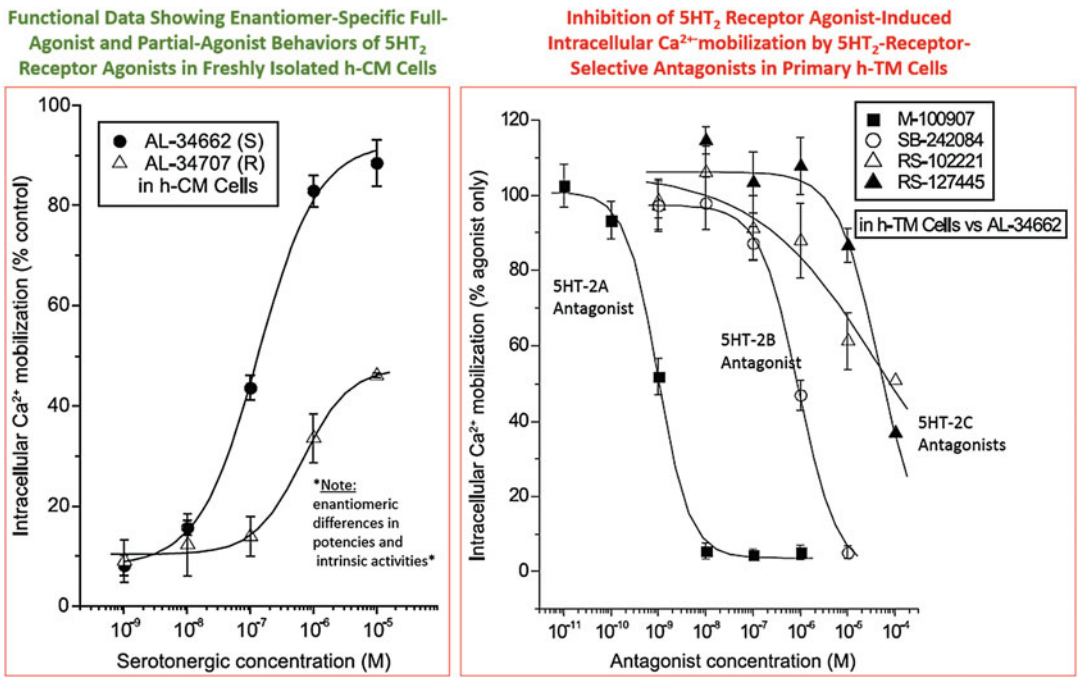


Fig. 4 Left-side panel of this figure shows the increase in intracellular Ca^{2+}-induced by AL-34662 and AL-34707, two enantiomers of a $5HT_2$-receptor agonist. Note that while AL-34662 is a full-agonist, AL-34707 is a much weaker agonist and is a partial-agonist based on its relative potency and intrinsic efficacy compared with AL-34662.The right-side panel depicts the ability of various $5HT_2$ receptor subtype-selective antagonists concentration dependently blocking the actions of AL-34662. (Modified from Sharif et al. (2007))

ion-channels or releasing intracellular Ca^{2+}, which ultimately results in a biological response such as hormone/cytokine release, muscle contraction, induction of inflammation and/or pain. On the other hand, kinase-linked receptors represent another large family of proteins that respond to certain growth factors and cytokines, and they trigger the phosphorylation of intracellular proteins that are involved in other types of signal transduction linked to cell growth, differentiation, and gene activation or inhibition. Furthermore, certain lipophilic ligands such as steroids enter the cell and have to move to the nucleus to activate gene transcription and thus influence protein synthesis (Fig. 5). Thus, rapid transmission of information (milliseconds to minutes) to modulate cellular activity and achieve communication among cells is undertaken by ion-channels/second messengers and intermediary mediators, while slow transfer of information (occurring over hours to days) is accomplished via gene expression changes (Alexander et al. 2017a, b, c, d). When the actions of the endogenous ligand or exogenously delivered drug are completed, these molecules are either degraded by specific enzymes or are taken up by cells via transporter systems and then metabolized and ultimately recycled. Transporters, as the name implies, shuttle chemicals, drugs, metabolites, and/or nutrients from one side of the cell membrane to the cytoplasm or vice versa, and this process occurs over a seconds-minutes time-scale. Thus, by effective and timely use of different receptors, ion-channels, transporters, and diverse signaling mechanisms, the body conserves energy, enhances efficiency and tries to maintain homeostasis and thus good health.

The degree to which the agonist ligand fits the receptor active-site and converts the inactive conformation of the receptor protein to a fully active or partially active state governs the relative

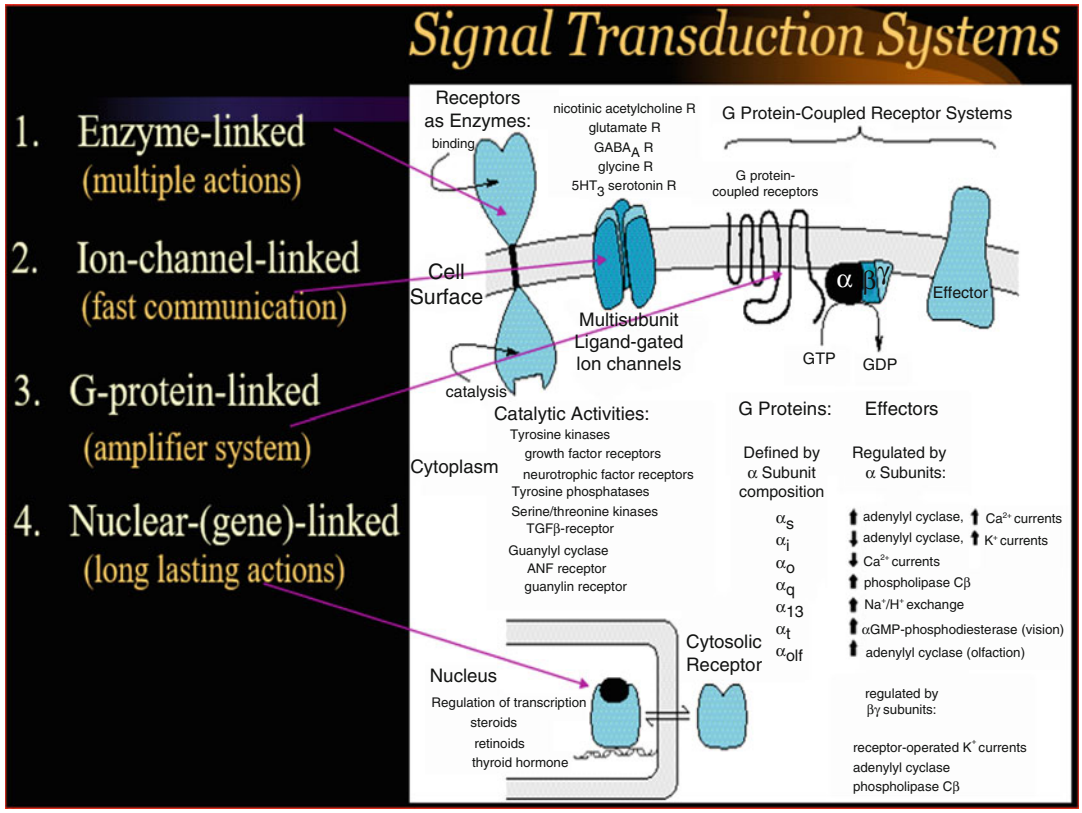


Fig. 5 The signal transduction systems associated with different kinds of receptor proteins present in most mammalian cells are shown

efficacy of the agonist (full- or partial agonist) to elicit the final cell or tissue response. The relative attraction of the ligand for the receptor to begin with is determined by the relative binding affinity of the ligand to fit the receptor-protein-pocket ("key and lock engagement") and the ability of the ligand-receptor complex to be formed and thus to cause a biological effect ("opening the lock phenomenon") (Fig. 5). For an agent to be classified as a pharmacologically relevant entity, its action at the receptor/ion-channel/transporter must follow concentration-response or dose-response relationship characteristics, and other compounds in the same class must demonstrate various degrees of affinity/potency/efficacy (intrinsic activity) with parallel concentration-response curves. This means that an agonist must elicit an increasingly larger response as its concentration or dose is increased until all the receptors/ion-channels/transporter ligand-binding sites are fully occupied and the induced response plateaus. In most cases, overstimulation of the receptor leads to dissociation of the receptor-G-protein complex, and a diminished response is observed. This desensitization phenomenon is well known and is responsible for the development of tolerance to a drug.

The potency of the agonist ligand is defined as the relative concentration needed to induce a given response via receptor activation or by engagement with its recipient protein. The relative potencies of agonist compounds (the concentration needed to induce 50% of the maximal response; EC_{50} or ED_{50}) is used to rank order such compounds (e.g., Tables 3 and 4.). Likewise, the relative potencies of antagonist or inhibitor

Table 3 Functional agonist potencies of selected PGs at various PG receptors

Test PG Compound	Agonist potency (EC_{50}; nM) at various prostaglandin receptors						
	DP-receptor (↑ cAMP)	EP_1-receptor (PI turnover; or other response)	EP_2-receptor (↑ cAMP; or other response)	EP_3-receptor (various functional responses)	EP_4-receptor (↑ cAMP)	IP-receptor (↑ cAMP or other response)	TP-receptor (PI turnover; or other response)
PGD_2	**74**	3,190	58,000	Nd	>10,000	>10,000	>10,000
PGI_2	>10,000	319	>10,000	3,019	>10,000	**7**	>10,000
PGE_2	>1,000	**2.9**	**67**	**19.9; 45; 4.5**	**40**	3,310	>10,000
$PGF_{2\alpha}$	>10,000	**29**	>10,000	691; >10,000; 2,000	>10,000	3,000	>10,000
Bimatoprost acid	>10,000	2.6	>10,000	Nd	>10,000	>10,000	>10,000
Travoprost acid	>10,000	Nd	>10,000	>10,000	>10,000	>10,000	>10,000
Latanoprost acid (PHXA85)	>10,000	119	20,000	12,000	>10,000	>10,000	>10,000
Cloprostenol	>10,000	93	>10,000	228	>10,000	>10,000	>10,000
S-1033	>10,000	>30,000	>10,000	>10,000	>10,000	>10,000	>10,000
Unoprostone (UF-021)	>10,000	>30,000	>10,000	>10,000	>10,000	>10,000	>10,000

Data are average values from up to three experiments.
Nd not determined (Sharif et al. 2003)

compounds (the concentration needed to block/inhibit 50% of the maximal response or event; IC_{50} or K_i) can be used to rank order such compounds in order to choose which compound(s) to pursue in animal studies, for instance. Furthermore, antagonists can be classified as competitive or noncompetitive. Competitive antagonists shift the concentration response of an agonist to the right (dextral shift) without diminishing the maximal effect of the agonist thereby reducing the agonist affinity for the receptor (Griffin et al. 1999). Noncompetitive antagonists invariably produce rightward shifts of the agonist concentration-response curves but prevent the agonist compound achieving its maximum effect (Sharif and Klimko 2019). Just as agonists cause receptor desensitization when the latter are exposed to excessively high a concentration (and/or exposed too often) of the agonist ligand, cells/tissue/animals challenged with very high levels of competitive antagonists (and/or on a high frequency) actually induce generation of more receptors (receptor upregulation) as a compensatory mechanism. This is often associated with a so-called "rebound effect" when the agonist actually produces a greater response than it induced before exposure to the high levels of the antagonist.

Ligand-gated ion-channels are important targets for neurotransmitter and drug interaction and thus for drug discovery/development. There are several different types of ion-channels that are present primarily on cell membranes of neurons and excitable cells where fast communication is needed. The most well-known ion-channels, that are made up of 3–5 protein subunits, are those permeable to Na^+, K^+, Cl^-, and Ca^{2+} ions and are responsible for depolarization or hyperpolarization of cells. Binding of specific ligands (e.g., serotonin, glutamate, zinc, acetylcholine, ATP) to certain subtypes of receptors activates these types of ion-channels (Alexander et al. 2017a, b, c, d). Certain cation-channels of the transient receptor potential (TRP) protein superfamily

Table 4 Functional agonist potencies of various FP-class PG analogs at native or cloned FP receptors in five different cell types of different species

Compound	Agonist potency for stimulating inositol phosphate production in different cell types (EC_{50}; nM)				
	Human ciliary muscle cells (h-CM cells)	Human trabecular meshwork cells (h-TM cells)	Human cells (HEK-293) expressing cloned human ocular FP receptor	Mouse Swiss 3T3 fibroblasts	Rat A7r5 vascular smooth muscle cells
Travoprost acid ((+)-fluprostenol	1.4 ± 0.2	3.6 ± 1.3	2.4 ± 0.3	2.6 ± 0.2	2.6 ± 0.5
(±)-fluprostenol	4.3 ± 1.3	11 ± 2	4.6 ± 0.4	3.7 ± 0.4	4.4 ± 0.2
Bimatoprost acid (17-phenyl-$PGF_{2\alpha}$)	3.8 ± 0.9	28 ± 18	3.3 ± 0.7	2.8 ± 0.2	2.8 ± 0.6
Latanoprost acid (PHXA85)	124 ± 47	35 ± 2	45.7 ± 8.4	32 ± 4	35 ± 8
Travoprost (Isopropyl ester)	123 ± 65	103 ± 27	40.2 ± 8.3	81 ± 18	46 ± 6
Latanoprost (Isopropyl ester)	313 ± 90	564 ± 168	173 ± 58	142 ± 24	110 ± 19
Bimatoprost (amide)	9,600 ± 1,100	3,245 ± 980	681 ± 165	12,100 ± 1,200	6,850 ± 1,590
Unoprostone (UF-021)	3,503 ± 1,107	3,306 ± 1,700	3,220 ± 358	617 ± 99	878 ± 473
Unoprostone isopropyl ester	8,420 ± 912	2,310 ± 1,240	9,100 ± 2,870	560 ± 200	458 ± 85
S-1033	4,701 ± 2,031	7,000 ± 2,600	2,610 ± 463	670 ± 320	767 ± 93
$PGF_{2\alpha}$	104 ± 19	62 ± 16	29 ± 2	26 ± 3	31 ± 3

Data are mean ± SEM from 3–23 experiments. *nd* not determined. $PGF_{2\alpha}$ (K_i = 122 ± 40 nM) and latanoprost acid (K_i = 149 ± 9 nM) exhibited relatively high affinity for the prostaglandin transporter transfected in host cells. The functional PI turnover activities of various PGs were blocked by the FP-receptor selective antagonist, AL-8810. The pooled antagonist potencies of AL-8810 were: cloned human FP receptor $K_i = 1.9 \pm 0.3$ μM; h-TM cell $K_i = 2.6 \pm 0.5$ μM; h-CM cell $K_i = 5.7$ μM; rat A7r5 cell $K_i = 0.4 \pm 0.1$ μM; and mouse 3 T3 cell $K_i = 0.2 \pm 0.06$ μM using a variety of FP agonists including fluprostenol, travoprost acid, unoprostone, bimatoprost, and bimatoprost acid (Sharif et al. 2003)

exist on nonexcitable and excitable cells and act as sensors of heat/cold, changes in osmolarity, odorants, and mechanical stimuli. The transient receptor potential vanilloid-1 (TRPV1) channel, for example, responds to capsaicin and detects "hot taste" associated with chili-peppers but is also activated by noxious heat (>43 °C), low pH, voltage, and various lipids.

All the pharmacodynamic aspects of drug-receptor (or drug-ion-channel) interactions mentioned above apply to in vitro and in vivo situations. Obviously, all drugs have side effects, some toxicological in nature, and thus the risk-benefit ratio must be determined for each in order to ensure that the therapeutic index is high and that the side effects are minimized for the subject. The drug safety elements also have to account for the dose of the drug administered, routes of administration, the speed with which the active drug reaches its intended site of action, the duration of action of the drug, and its rate of and safe elimination from the body. Such data are obtained

from pharmacokinetic (PK) and absorption/distribution/metabolism/elimination (ADME) studies conducted in suitable normal healthy animals or human subjects during the drug development processes. Direct efficacy of the drug can then be assessed at appropriate dosage(s) in animal models of disease and of course in humans in clinical trials in due course.

Due to the diversity of cell types involved in the anatomy and physiology of the eye and the mediation of pathogenesis of various ocular diseases, the eye offers a great opportunity to apply the principles of pharmacology. There is thus a rich history of delineation of physiological and pharmacological actions of endogenous ligands and exogenous experimental chemical agents. A few examples include the role of endogenous epinephrine and norepinephrine (NE) in promoting the synthesis of AQH by the nonpigmented ciliary epithelial cells of the ciliary body in the anterior chamber (ANC) of the eye, and the role of vascular endothelial growth factor (VEGF) to stimulate the growth of new blood vessels into the vitreous humor in wet AMD and diabetic retinopathy. The NE-induced AQH production is mediated through action at alpha- and beta-adrenoceptors and by raising the intracellular levels of cAMP, while the VEGF-induced neovascularization of the choroidal blood vessels is mediated through receptor-tyrosine kinase (RTK)-coupling. Hence, β-blockers and anti-VEGF antibodies reduce AQH production and neovascularization, respectively. Similarly, allergen-induced conjunctival mast cell degranulation results in the release of histamine and other mediators into the tear film with subsequent activation of histamine receptor-1 in conjunctival and corneal epithelial cells (Sharif et al. 1994, 1996). This raises intracellular Ca^{2+} and then results in secretion of various proinflammatory cytokines, that together with histamine, cause increased vascular permeability of both tissues causing ocular itching and redness. And thus, treatment of the ocular surface with H_1-receptor antagonists such as emedastine and olopatadine blocks the effects of histamine and curtails the itching and redness (Sharif et al. 1994, 1996; Yanni et al. 1999). Other examples of ocularly relevant receptors/ion-channels/transporters, and drugs that interact with them, will be discussed below in more detail.

Application of Pharmacodynamic Principles in Ocular Drug Discovery and Development

The first step in any drug discovery program is to identify the target protein whose activity needs to be modulated to achieve the therapeutic benefit. Next, it is important to verify the localization and relative distribution of the target protein in the specific tissue/cells connected with the disease process and in other areas where potential side effects may occur. Such target localization and visualization can be accomplished using techniques such as autoradiography (Fig. 6, left panel), immunohistochemistry (Fig. 6, right panel), and *in situ* hybridization. From an ocular perspective, the relative density and distribution and cellular localization of the target protein in animal and human eye sections is paramount. Examples of use of such techniques and results obtained are shown in Fig. 6 and Table 5.

Testing funnel schemes are drawn up next that place the target protein into ligand binding assays (and/or directly into functional assays) using cell/tissue homogenates enriched in the target protein (Fig. 7). If the target receptor/transporter/ion-channel can be found at high levels in animal/human cells/tissues, then the naturally occurring protein can be used for screening purposes, and this is actually preferred (Sharif 2018a). The alternative strategy is to avail recombinant molecular biological techniques to express the desired protein at high levels in host cells to make the assay sensitive and reproducible. Using known high-affinity (usually nanomolar K_d) tritiated or iodinated radiolabeled ligands to tag the target protein, the relative affinity of unlabeled test compounds to compete for, and thus displace, the radioligand from its binding sites is determined using rapid vacuum filtration techniques to separate the free from bound radioligand followed by liquid

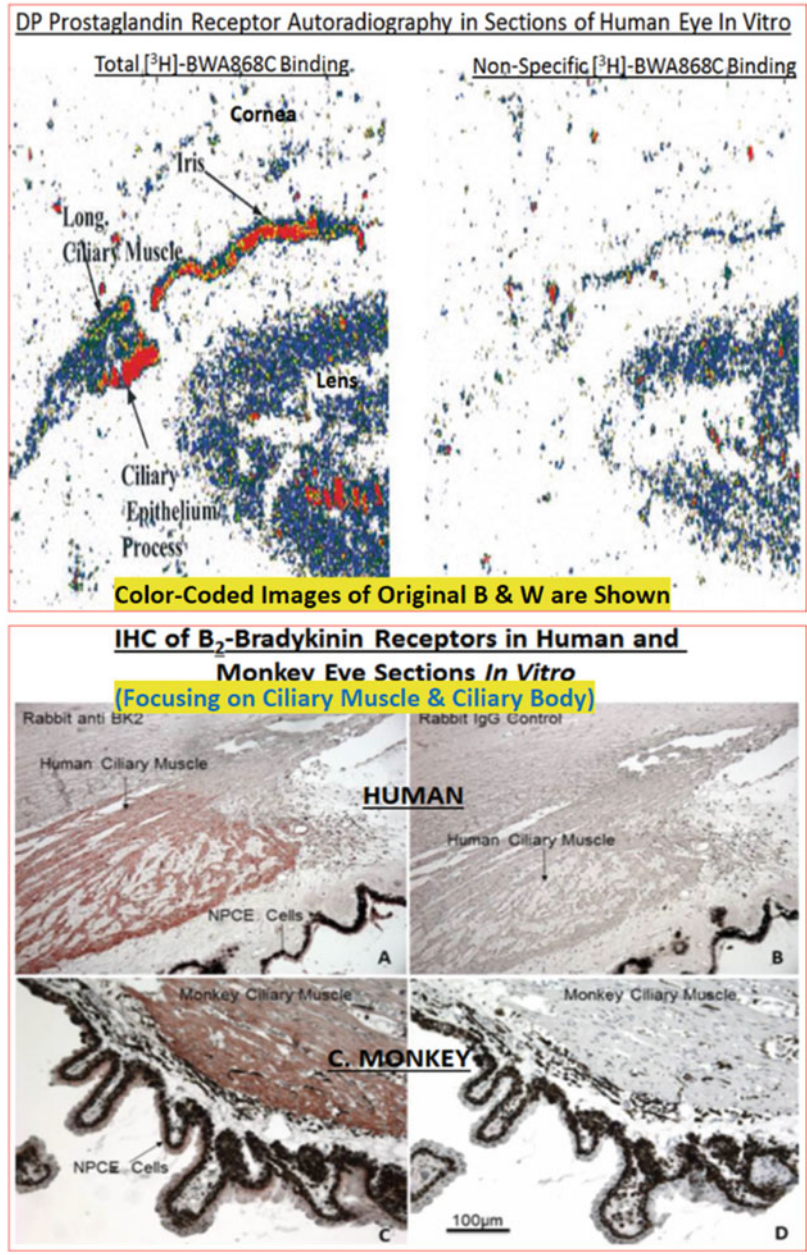


Fig. 6 Two different techniques of localizing and visualizing membrane-bound receptors in human and monkey eye sections are illustrated. The top panel shows the autoradiographic distribution of DP PG receptor binding sites in the anterior segment of the human eye using radio-labeled DP-receptor antagonist, [^{3}H]-BWA868C. Top-left panel shows the total binding, while the top-right panel shows the nonspecific binding of the radioligand. (Modified from Sharif et al. 2000). The bottom panel depicts localization of bradykinin B_2-receptors in human and monkey eye sections using the IHC technique– specific binding of the B_2-receptor anti-body is shown in the left-side panels (russet color), while the control (nonspecific binding) is shown in the right-side panels for each species.(Modified from Sharif et al. (2014).

scintillation spectrometry (Fig. 8, left panel; Sharif 2018a). Alternatively, homogeneous proximity-scintillation ligand binding assays can be utilized. The ligand binding affinity data obtained from such experiments are used to rank-order compounds and only those meeting specified pass criteria (e.g., $K_i \leq 50$ nM) are then tested in functional assays. Here, isolated primary cells from target tissue are used to determine if the compounds are agonists or antagonists (Sharif et al. 2007) (e.g., Fig. 8, right panel) or activators or inhibitors of enzymes (Chen et al. 2014) or channels (Patil et al. 2016). Compounds can then be ranked according to their relative functional potencies for instance for their ability to stimulate the production of intracellular cAMP (Crider and Sharif 2001) or cGMP (Katoli

Table 5 Quantitative autoradiographic distribution of various drug receptors in postmortem human eye sections

	All 5-HT Receptors	5-HT_2 Receptors	β-Adrenoceptors	DP PG Receptors	FP PG Receptors
	[^{3}H]-5-HT binding (5-HT receptors)	*[^{3}H]-ketanserin binding (5-HT_2 receptors)*	*[^{3}H]-levobetaxolol binding (β-receptors)*[a]	*[^{3}H]-BWA868C binding (DP prostaglandin)*[b]	*[^{3}H]-$PGF_{2\alpha}$ binding (FP prostaglandin)*[c]
Tissue	*Specific binding; DLU/mm^2 and (% specific binding)*	*Specific binding; DLU/mm^2 and (% specific binding)*	*Specific binding; DLU/mm^2 and (% specific binding)*	*Specific binding; DLU/mm^2 and (% specific binding)*	*Specific binding; DLU/mm^2 and (% specific binding)*
Ciliary epithelium (process)	71,780 ± 2,725 (70%)	13,683 ± 5,870 (40%)	51,459 (76%)	67,000 (82%)	2,554 (35%)
Longitudinal ciliary muscle	14,232 ± 7,937 (48%)	14,459 ± 3,683 (47%)	27,543 (83%)	37,900 (79%)	12,741 (68%)
Iris	174,943 ± 20,092 (74%)	20,026 ± 11,276 (20%)	78,140 (62%)	55,500 (54%)	3,776 (40%)
Lens	2,225 ± 582 (14%)	4,151 ± 2,762 (20%)	1,507 (18%)	23,000 (5%)	1,886 (12%)
Choroid	33,238 ± 5,950 (60%)	16,304 ± 3,792 (40%)	19,244 (72%)	39,700 (54%)	1,671 (28%)

DLU digital light units. These units represent an index of the relative density of the receptor population found in the tissues studied

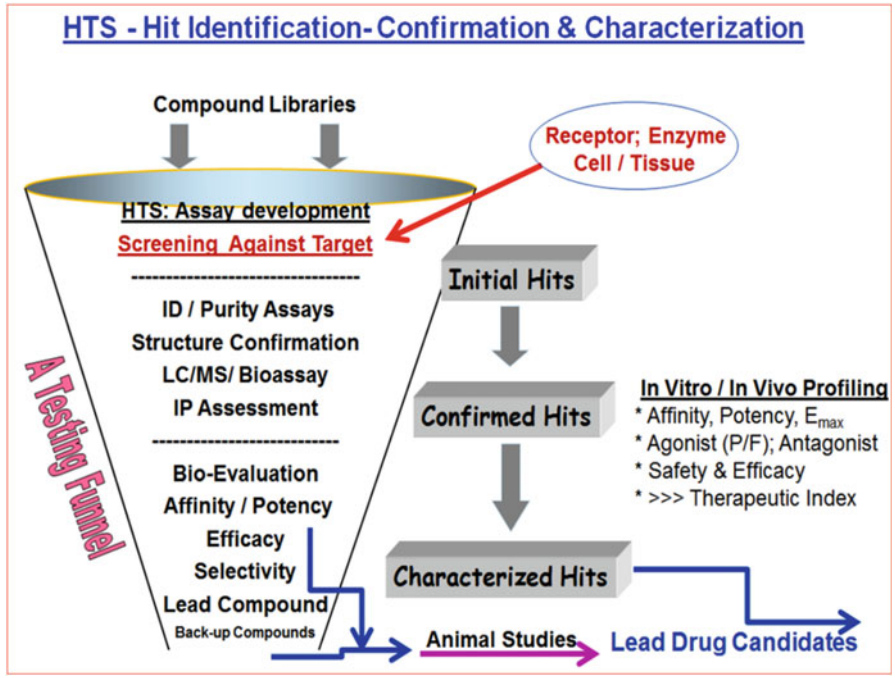


Fig. 7 A schematic illustration of testing funnels for profiling and selecting compounds as "Hits" and fully characterized compounds is shown. Note that the target can be a receptor, enzyme, transporter, or ion-channel expressed in isolated normal primary cells or genetically engineered cells or in isolated animal/human tissue samples

et al. 2010; Cavet and DeCory 2018; Fig. 9, left panel) or to mobilize intracellular Ca^{2+} (Kelly et al. 2003; Sharif et al. 2007; Fig. 8, right panel) or to induce cell/tissue contraction (Ohia et al. 2018).

Sometimes the radioligand binding assays are omitted from the testing funnels and compounds are directly screened in functional activity assay systems using multiple compound concentrations to construct full concentration-response curves (e.g., Fig. 4, left panel). Once again relative potency data are used to triage and select most potent and efficacious agonists (e.g., $EC_{50} \leq 10$ nM), or most potent antagonists ($K_i \leq 10$ nM), to advance into *in vivo* testing in animal models of ocular safety (e.g., guinea pigs or rabbits for ocular irritation, redness [hyperemia], or inflammation [mucus-containing discharge]), followed by testing in animal models of eye disease (e.g., ability of compounds to lower intraocular pressure (IOP) in rodent, rabbit, and monkey eyes, either normotensive or ocular hypertensive) (e.g., Fig. 9, right panel, Cavet and DeCory 2018; Fig. 10, Sharif et al. 2014).

In other cases where compound supply is limited, only one or two concentrations are tested to generate receptor binding and/or functional-response data to help make decisions. If none of the tested compounds meet the prespecified criteria of affinity/potency/in vitro efficacy, or in order to improve the overall profile of the compounds, the medicinal chemists would need to modify the chemical structure of the compounds and the biologists would need to retest the new compounds. This iterative process can be short or long depending on the complexity of the synthetic process for the compounds/peptides/antibodies and the degree of difficulty and complexity of the assay systems. Based on the physiochemical properties and structure-activity relationship (SAR) profile of the compounds, a suitable lead compound and possible backup compounds are selected for advancement to full development that requires gathering much more detailed information on the ocular/systemic safety, PK, ADME, and toxicology (genotoxicity; central and systemic toxicology in multiple species after repeated

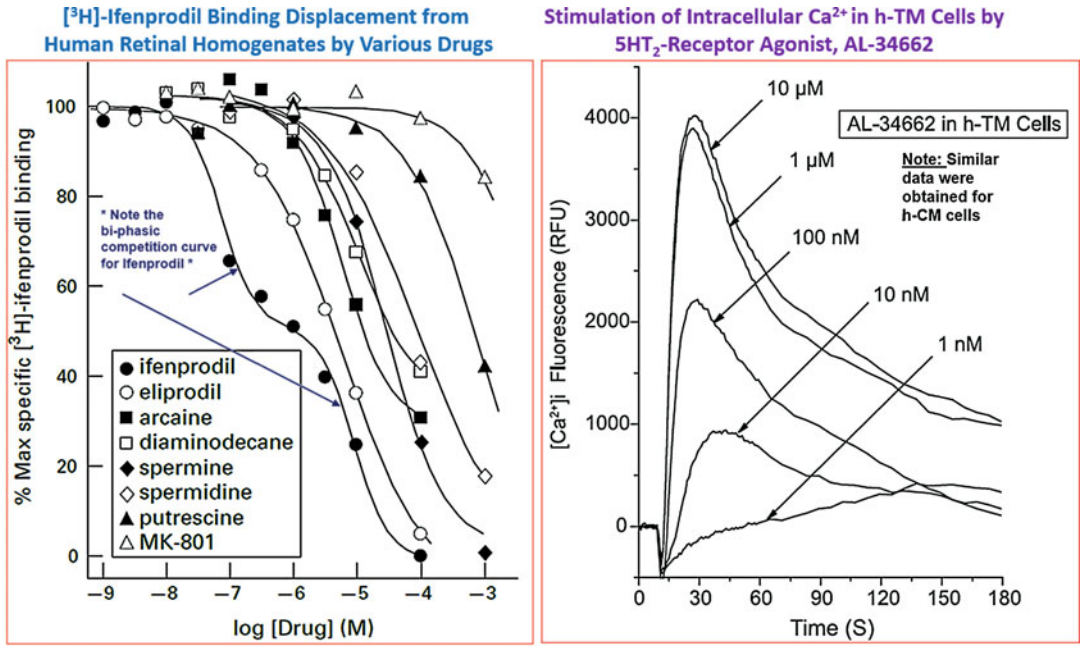


Fig. 8 Two different types of assays are shown. The left-side panel depicts a typical receptor-binding experiment where various compounds of interest are screened for their ability to displace a receptor-specific radioligand, in case from human retinal homogenates. Note how a compound can be shown to interact with the bind site in a high- and a low-affinity state(ifenprodil displacing [^{3}H]-ifenprodil), or to only bind to a low-affinity state of the receptor. The relative affinities of the different compounds are easily observed and determined in this manner. (Modified from Sharif and Xu, 1999). The right-side panel shows a typical profile of intracellular Ca^{2+} increase induced by different concentrations of the $5HT_2$-receptor agonist, AL-34662. The peak induction of responses from such traces are then used to construct concentration–response curves depicted previously in Fig. 4. (Modified from Sharif et al. (2007))

dosing at multiple doses) information in suitable animal species and test systems. Additional work requiring building a data package covering optimized formulation, route(s) of administration, duration, and mechanism of action of the lead compound (and a backup compound) would be undertaken next. These are lengthy, laborious, expensive, but necessary experiments to discover suitable drugs.

Ultimately, the lead compound data package would have met all the necessary requirements to be considered for clinical trials in humans. At this stage, a formal request to conduct Phase-1 (primarily for ocular safety at multiple doses) clinical trials is made to the regulatory agency of the country whence the complete data package on the Investigational New Drug is submitted for approval. If the lead compound is pronounced safe in a limited number of healthy humans (e.g., 20–30) at multiple doses administered once daily or multiple times/day via the best route of administration (e.g., topical ocular; intracameral; intravitreal; Hartman and Kompella 2018), Phase-II studies can be conducted. About 70% of drugs that enter Phase-I will be successful enough to proceed to Phase-II. Here, clinical studies are conducted in a small number of age-matched control patients lacking the ocular disease, and in those patients having the ocular disease. A control vehicle (placebo), multiple doses of the test drug candidate and a single dose of a suitable comparator drug (previously approved by the regulatory authority) are tested in patients who have the ocular disease (e.g., 30–50 per each treatment arm). Such Phase-II studies are conducted over several months and allow the selection of the most optimum dose of the lead compound. In Phase-III studies, the optimum dose of the lead compound is compared against specified marketed comparator drug (in

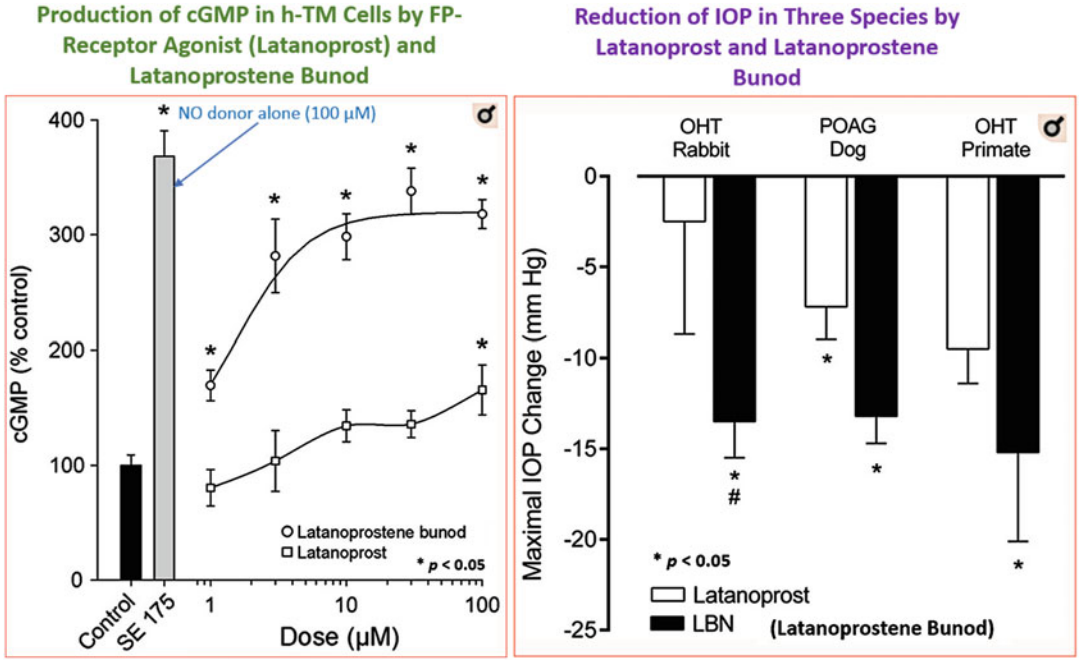


Fig. 9 The left-side panel illustrates the production of a second messenger, cGMP, in h-TM cells by various ligands, while the right-side panel shows how the different ligands lower IOP in rabbits, dogs, and monkeys. The cell-based and animal model-based data correlate well with each other. (Modified from Cavet and DeCory (2018).

the same drug class) in a much larger human population (e.g., 100–800 patients/treatment arm) of patients with the ocular disease and the trials conducted for several months. It is important to "statistically power" the clinical trials (with sufficient number of patients per treatment arm) in order to show statistically and clinically relevant efficacy and benefit to the patient suffering from the ocular disease.

Having established the necessary safety and efficacy of the drug candidate molecule, the regulatory agency can be approached for approval to market the drug by submitting a New Drug Application. This dossier contains all the necessary guidance on drug manufacturing procedures, stability and formulation data, PK, ADME, and other necessary data, and of course all the preclinical and clinical safety and efficacy data, as required by the specific health authority. After meeting all the regulatory authority's criteria and requirements, the agency may grant an approval to market the drug and thus make it available for use by clinicians to treat the patients with the ocular disease with appropriate guidance on dose/frequency of dosing/route of administration and of course side effects, etc. Based on the above, it is not surprising that drug discovery, drug development and approval by a health authority is a very long (taking 10–12 years from discovery to marketing and before being introduced into clinical medicine) and very expensive process (typically costing $50–100 million depending on the type of drug/its cost of goods/frequency of dosing, etc.). Nevertheless, such due diligence, time, and cost is worthwhile in order to reduce and/or prevent visual impairment and preserve eyesight for a global population of humans afflicted with the eye disorders for which the treatment was sought.

Eye Diseases and Their Pharmacological Treatments

As the earlier discourse above illustrates, the eye is a very complex organ being composed of a heterogeneous population of specialized tissues/

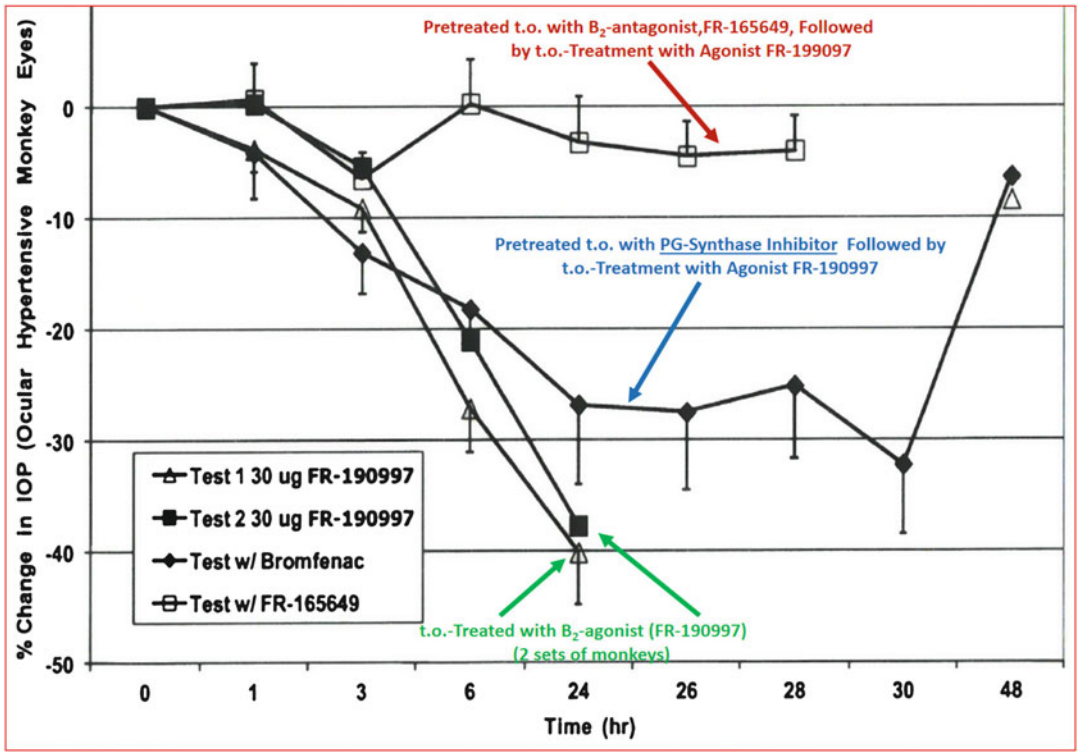


Fig. 10 IOP reduction in conscious cynomolgus monkeys. This figure shows how the nonpeptide B_2-bradykinin receptor agonist FR-190997 lowered IOP of ocular hypertensive eyes of the monkeys when delivered topical ocularly (t.o.). The effect of FR-190997 was reproduced in two different colonies of cynomolgus monkeys. Additionally, it was important to demonstrate that predosing t.o. with a B_2-receptor antagonist FR-165649 could block the IOP-lowering effects of the agonist. Likewise, in order to show that endogenous PGs were mediating some (or all) effects of activating the B_2-receptor, a PG synthase inhibitor partially reduced the agonistic effects of FR-190997. Such studies illustrate the role of different pharmacological treatments to better understand the mechanism of action of drugs in vivo. (Modified from Sharif et al. (2014))

cells which have diverse functions ranging from providing structural support, limiting pathogenic infiltration, keeping corneal transparency, focusing light, transducing light into electrical impulses, absorbing excess light, phagocytosing toxins and cellular debris, production and drainage of AQH, etc. It is therefore not surprising that dysfunction of these cell types results in various ocular disorders such as cataracts (51% of total eye diseases), glaucoma (8%), AMD (5%), corneal opacity/scarring (4%), pediatric eye diseases (4%), trachoma (3%), uncorrected refractive errors (e.g., myopia; presbyopia; 3%), diabetic retinopathy (1%), and infections, ocular allergies, inflammation, retinoblastoma, etc. (21%) (WHO 2018; NEI 2014).

The other aspects worthy of note include the fact that while somewhat isolated from the rest of the body, medications administered to the eye topically still drain into the naso-pharynx area and eventually into the stomach and then into the blood stream. As the blood is supplied to each organ, the latter are then exposed to the drug that was administered to the eye, albeit at lower concentrations. Depending on the concentration/affinity/potency of the drug or its metabolite(s), off-target side effects are always possible and thus the physician and the patient need to be aware of such issues. Likewise, drugs that are delivered to the anterior or posterior chambers of the eye via different routes of administration also are routed via the venous circulation and eventually

metabolized by the liver and then eliminated via the kidneys and/or the alimentary canal. Due to these reasons, local ocular and systemic and central side effects of drug treatments to the eye have serious consequences and need to be considered in the treatment regimens.

In order to emphasize the pharmacodynamic and translational aspects of ocular pathologies and their treatments, eye disorders that are more prevalent and which are debilitating from a visual impairment perspective will be covered first followed by others that are lesser common and where pharmacological elements are less well characterized from the structure activity and pharmacological viewpoint. Due to space limitations, only certain eye diseases will be discussed in detail where the pharmacology is quite well defined and robust. It is hoped that readers will follow-up on the cited references, especially review articles, to help advance their understanding of the eye disease pathogenesis and the treatment modalities available.

Primary Open-Angle Glaucoma (POAG) and Ocular Hypertension (OHT)

Glaucoma is a heterogeneous group of ocular diseases that have their origin due to pathogenic events occurring in at least three distinct parts of the eye, namely, ANC, retina, and optic nerve. Over >65 million people around the world suffer from the most common form of glaucoma (primary open-angle glaucoma (POAG)). It is a slowly developing ocular disease that the patient only notices once some of the peripheral vision has already been lost and when only 60% of the original RGCs remain functional and capable of transmitting retinal signals to the brain. Eyesight is lost due to POAG and other types of glaucoma when the RGCs die, their axons atrophy and the RGCs are disconnected from the brain where neurons also die (Yucel et al. 2000). Other characteristic features of POAG are thinning of the retinal nerve fiber layer (RNFL) due to RGC axon atrophy, excavation of the optic nerve head (OHN), and subsequent cupping of the optic nerve disc (Tham et al. 2014; Weinreb et al. 2014; Jonas et al. 2017).

Behind cataracts, POAG is the next most prevalent preventable blinding eye disease. Despite linkage of numerous risk factors to POAG, abnormally high IOP is the most high IOP accepted modifiable element in the disease process. This ocular hypertension (OHT) is caused by an imbalance between the amount of AQH being produced and the amount draining from the ANC. The blockage of the TM and SC with aberrant or age-related accumulation of extracellular matrix (ECM) and cell debris is primarily the culprit responsible for OHT (Xu et al. 2014). The OHT has been shown to be directly responsible for demise of RGCs and their axons in animal models of POAG and in humans such that it is calculated that every 1 mmHg reduction of IOP results in up to 13% lowering of the progression of POAG (Weinreb et al. 2014; Jonas et al. 2017). Therefore, reducing IOP has been an effective way to treat OHT/POAG for many decades. This has resulted from the fundamental understanding that either the production of AQH can be reduced and/or the efflux of AQH can be accelerated in order to lower IOP. Even though inhibiting AQH generation by blocking the Na^+-K^+-ATPase in the CEP cells of the ciliary processes is not usually recommended, since the AQH provides nutrients to the ANC cells and removes their waste products, early treatment options were quite limited and clinicians had no choice but to inhibit the inflow process. The relatively recent discovery and development of drugs, surgical procedures and devices that can effectively reduce IOP by stimulating the drainage of AQH via the trabecular meshwork and/or uveoscleral pathways and newly created pathways have revolutionized clinical management of OHT/POAG.

Several decades ago, pharmacological management of IOP was achieved using pilocarpine, a plant-derived alkaloid muscarinic receptor agonist. However, while it reduced IOP it also caused miosis and brow ache by contracting the iris sphincter and ciliary muscle (CM). Additionally, its IOP-lowering effect was rather short-lived (4–6 h depending on the concentration of the drug and dosing frequency). Since that time, an enormous amount of progress has been made in discovering, developing, and marketing a variety of pharmacological agents that have

diverse mechanisms of action to lower IOP. Today, prostaglandin FP-receptor agonist analogs (PGAs; e.g., latanoprost; travoprost; tafluprost; bimatoprost) (Hellberg et al. 2002) represent the drugs of choice for treating ocular hypertension and POAG since they exhibit excellent efficacy, long duration of action (at least 24 h) with relatively low incidence and severity of side effects after a single topical ocularly administered drop of the drug. These PGAs promote egress of AQH from the ANC of the eye via the uveoscleral pathway and to some extent via the conventional TM-SC outflow pathway by releasing matrix metalloproteinases that digest ECM and other cellular debris and thus enlarging existing spaces between CM bundles and the sclera and/or creating new drainage channels in the latter tissues. The notable side effects of PGAs, however, are hyperemia (eye redness), darkening of the iris and orbital skin, lengthening and thickening of eyelashes, deepening of the orbital sulcus, and to a lesser extent, cystoid macular edema.

While alpha-2 adrenergic agonists like brimonidine and apraclonidine lower IOP by inhibiting inflow (generating cAMP that suppress Na^+-K^+-ATPase in ciliary epithelium) and by stimulating some outflow, their ocular (e.g., ocular allergy) and systemic and central (lowering CNS activity and causing lethargy) side effects limit their utility. Other agents that inhibit production of AQH, inflow inhibitors, also include carbonic anhydrase inhibitors (e.g., dorzolamide and brinzolamide), and beta-blockers (e.g., timolol and betaxolol) that do lower IOP but exhibit a number of ocular side effects (burning, stinging, foreign-body sensation) and systemic side effects (drop in blood pressure, bradycardia, palpitations, arrythmias, and bronchospasms). The latter inflow and outflow drugs have been combined in suitable formulations to generate so-called "combination products" with certain degree of enhanced efficacy (Holló et al. 2014), that have expanded the treatment options for treating elevated IOP and glaucoma. The exceptional value of the inflow suppressors and outflow stimulators is highlighted by the latter combination products but also their utility in glaucoma and OHT patients who become refractory to or are nonresponders to PGAs, especially to latanoprost.

In order to overcome some of the deficiencies and side effects of the aforementioned drugs for the treatment of OHT and glallcoma, two recent FDA-approved medications have been marketed: netarsudil (Rhopressa; Lin et al. 2018) and latanoprostene bunod (Vyzulta; Cavet and DeCory 2018). While netarsudil is a rho-kinase inhibitor that lowers IOP by relaxing the ciliary muscle and TM (Lin et al. 2018) thereby stimulating AQH to flow out of the ANC, latanoprostene bunod is a conjugate drug made up of latanoprost and a nitric oxide donating agent (Cavet and DeCory 2018), that reduces IOP by relaxing CM and TM tissues (engaging the outflow pathway) and by activating the UVS pathway. How these new drugs will fare after being introduced into clinical management of OHT/glaucoma remains to be seen, but in the meantime, the search for even better pharmacological agents with unique characteristics and better side effect profiles continues. It is encouraging that continued research in this area is poised to deliver additional drugs as judged by the multitude of reports published in recent years. The most succinct way to show this progress is via a tabular listing of such pharmacological agents (Table 6). Pharmacodynamic aspects, including mechanism of action, of each class of these agents is shown in this table and also described in detail in the relevant citations. Another exciting recent development is the ability of implanted microdevices, after Minimally Invasive Glaucoma Surgeries (MIGS), to literally drain the excess AQH from the ANC of OHT/POAG patients (e.g., Batlle et al. 2016; Fig. 11) without causing collapse of the ANC.

The MIGS-related devices have revolutionized AQH drainage from the ANC and have added another means to lower IOP, which was previously dominated by tubes and trabeculectomies (Batlle et al. 2016).

Receptor Binding and Functional Assays to Discover New IOP-Lowering Agents

Testing of potential ocular hypotensive agents is simplified since in most cases the target protein and its signal transduction mechanism is known.

Table 6 Pharmacological agents that lower IOP in various mammals and the mode of action of the compounds

Compound classes	Pharmacological agent	Reported or potential mode(s) of action
Conventional outflow (TM) stimulators		
Muscarinic receptor agonists (mostly M_1 receptor agonists)	Pilocarpine; Acecledine; Carbachol	Contract ciliary muscle/TM to promote outflow of AQH via the TM-SC pathway
Inhibitors of chloride transport	Ethacrynic acid; Ticrynafen; Indacrinone	Inhibition of Na^+-K^+-Cl^--transporter activity in the TM changes cell shape and volume and thus AQH efflux is increased
Kinase inhibitors	*Rho kinase (ROCK) inhibitors*: AR-12286 (Netarsudil); Ripasudil (K115); Y-27632; Y-39983; AMA-0076; H-7; ML-9; Chelerythrine; Staurosporin	Modification of actomyosin contractility that leads to changes in actin cytoskeleton of TM (relaxation) and this leads to AQH efflux
	LIM-K inhibitor	
	Myosin-II ATPase inhibitor: Blebbistatin	
	Src kinase inhibitor	
Marine macrolids	Latrunculins A and B; Bumetanide; Swinholide	Promote sequestration of actin monomers and dimers in TM; cause cell TM shape change and thus AH efflux
Guanylate cyclase activators	*Natriuretic peptides*: ANP; CNP; SHP-639	Type-A and type-B receptor activation leads to cGMP production, TM relaxation and AQH efflux via TM
NO donors	Sodium nitroprusside; Hydralazine; 3-morpholinosyndnonimine; (S)-nitroso-acetyl-penicillamine; NCX-125	
Soluble guanylate cyclase activators	YC-1; BAY-58-2667; IWP-953	
κ-opioid receptor agonists	Bremazocine; dynorphin	Release natriuretic peptides and thus raise cGMP in TM leading to its relaxation and thus AQH efflux
Cannabinoid receptor agonists	WIN55212-2; CP55940; SR141716A	Receptor stimulation opens BKC-channels and relaxes TM which then causes AQH efflux via TM and SC
FP-class PG-receptor agonists	Latanoprost; Travoprost; Tafluprost; Bimatoprost; Unoprostone isopropyl ester	Some clinical evidence of promoting conventional outflow in addition to UVS outflow
Serotonin-2 receptor antagonists	BVT-28949; ketanserin and its analogs	Unknown and unverifiable mechanism(s) of action (may block beta-adrenergic receptors indirectly?)
Releasers of MMP and AP-1	FP-class PGs (see above); and *t*-butylhydroquinone (t-BHQ); β-naphthoflavone;	Local production of MMPs; ECM degradation; stimulation of AQH efflux via TM
Uveoslceral (UVSC) outflow stimulators (via gaps in CM fiber bundles and scleral tissue)		
FP-class PG-receptor agonists	Latanoprost; Travoprost; Tafluprost; Bimatoprost; Unoprostone isopropyl ester	FP receptor activation in CM causes release of MMPs that breakdown ECM ("clog") around CM bundles and within sclera thus causing UVS outflow of AQH
EP_2- and EP_4-PG-receptor agonists	Omidenepag Isopropyl (DE-117); Butaprost; AL-6598; ONO-AE1-259-01; PF-04217329; PF-04475270	Receptor activation increases cAMP that relaxes CM and TM; EP_2 agonists also cause release of MMPs that breakdown ECM ("clog") around CM bundles and within sclera thus causing UVSC outflow of AQH
Serotonin-2 ($5HT_2$) receptor agonists	(R)-DOI; α-methyl-5HT; AL-34662	Contraction/relaxation of CM and TM by activation of $5HT_2$ receptors. May also release MMPs and/or PGs or other local

(continued)

Table 6 (continued)

Compound classes	Pharmacological agent	Reported or potential mode(s) of action
		mediators that promote CM remodeling and thus promote UVS outflow
Bradykinin B_2-receptor agonists	Bradykinin; FR-190997; BKA278	B_2-receptor activation causes PI hydrolysis production of IPs and DAG; cause PG release and release of MMPs that digest ECM and this promote UVS outflow in cynomolgus monkey; conventional outflow also stimulated in isolated bovine/porcine anterior eye segments
Dual activity PGs, and conjugated compounds	FP/EP_3 receptor agonist (ONO-954)	Promotes UVSC outflow
	AL-6598 (DP/EP_2 receptor agonist)	Inhibits inflow and stimulates outflow (TM and UVSC)
	Latanoprostene bunod (latanoprost-NO donor conjugate)	Promotes UVSC and TM outflow
Inflow inhibitors (reduce AQH production)		
β-adrenergic antagonists	Timolol; Betaxolol; Levobetaxolol; Levobunolol; Metipranolol	Block β-adrenergic receptors in the ciliary process, decrease cAMP generation and thus decrease AQH formation
β-adrenergic receptor silencer	SYL-040012; siRNA (Bamosiran)	Downregulates endogenous β-adrenergic receptors and their signaling
α2-adrenergic agonists	Brimonidine; Apraclonidine; Clonidine	Intracellular cAMP reduced in CP that decreases AQH generation; may also prevent NE release Brimonidine also promotes TM outflow
Carbonic anhydrase inhibitors (CAIs)	Dorzolamide; Brinzolamide	Inhibit ciliary process CA-II and CA-IV and thus reduce bicarbonate production that in turn reduces AQH generation
Chloride channels inhibitors	5-nitro-2-(3-phenylpropylamino)-benzoate (NPPB)	Ion flux of CP NPE cells causes reduction of AQH formation
Na^+-K^+-ATPase inhibitors	Ouabain; Digoxin analogs	Ciliary process Na^+-K^+-ATPase inhibited leading to inhibition of AQH production
Dopamine receptor agonists	PD-128907; CHF-1035; CHF-1024; SDZ GLC-756; (S)-(-)-3-hydroxyphenyl)-N-n-propylpiperidine (3-PPP)	Inhibit release of NE and prevent AQH production; may also release natriuretic peptides
Na^+-K^+-ATPase inhibitors	Ouabain; Digoxin analogs	Ciliary process Na^+-K^+-ATPase inhibited leading to inhibition of AQH production
Aquaporin inhibitors	Various aromatic sulfonamides and dihydrobenzofurans	Inhibit release of NE and prevent AQH production
Additional IOP-lowering agents		
Mas receptor stimulator	DIZE via ACE-2 activation	Prevent ECM (including TGFβ) accumulation (outflow stimulation?)
Angiotensin-II receptor antagonists	CS-088	Various mechanisms of action; not robust IOP-lowering
Ca^{2+}-channel inhibitors	Lomerazine; Nivaldipine; Nifedipine; Nimodipine; Verapamil; Brovincamine; Iganidipine	Enhance retinal blood flow; some may lower IOP; work well in normal tension glaucoma patients
Alpha-adrenergic receptor antagonists	Oxymetazoline; 5-methylurapidil; Ketanserin	Work mostly via outflow mechanism but this needs to be defined
Other prostaglandin receptor agonists	AL-6598 (DP/EP_2 receptor agonist); AGN192093 (TP receptor agonist); BW245C (DP receptor agonist); Sulprostone (EP_3 receptor agonist)	These work through multiple mechanisms of action involving cAMP production, Ca^{2+} mobilization leading to relaxation/ contraction of ciliary muscles/TM

(continued)

Table 6 (continued)

Compound classes	Pharmacological agent	Reported or potential mode(s) of action
PG-conjugates	Latanoprostene Bunod (NO donor coupled to latanoprost)	Combination of NO-cGMP production and FP-receptor activation
Combination products	Brinzolamide-brimonidine; Brinzolamide-brimonidine; Acetozolamide-Timolol-Brimonidine; Travoprost-brimonidine; Bimatoprost-brimonidine; Tafluprost-Timolol	Complementary mechanisms of action encompassing inflow-outflow inhibition, and inflow-uveoscleral outflow inhibition

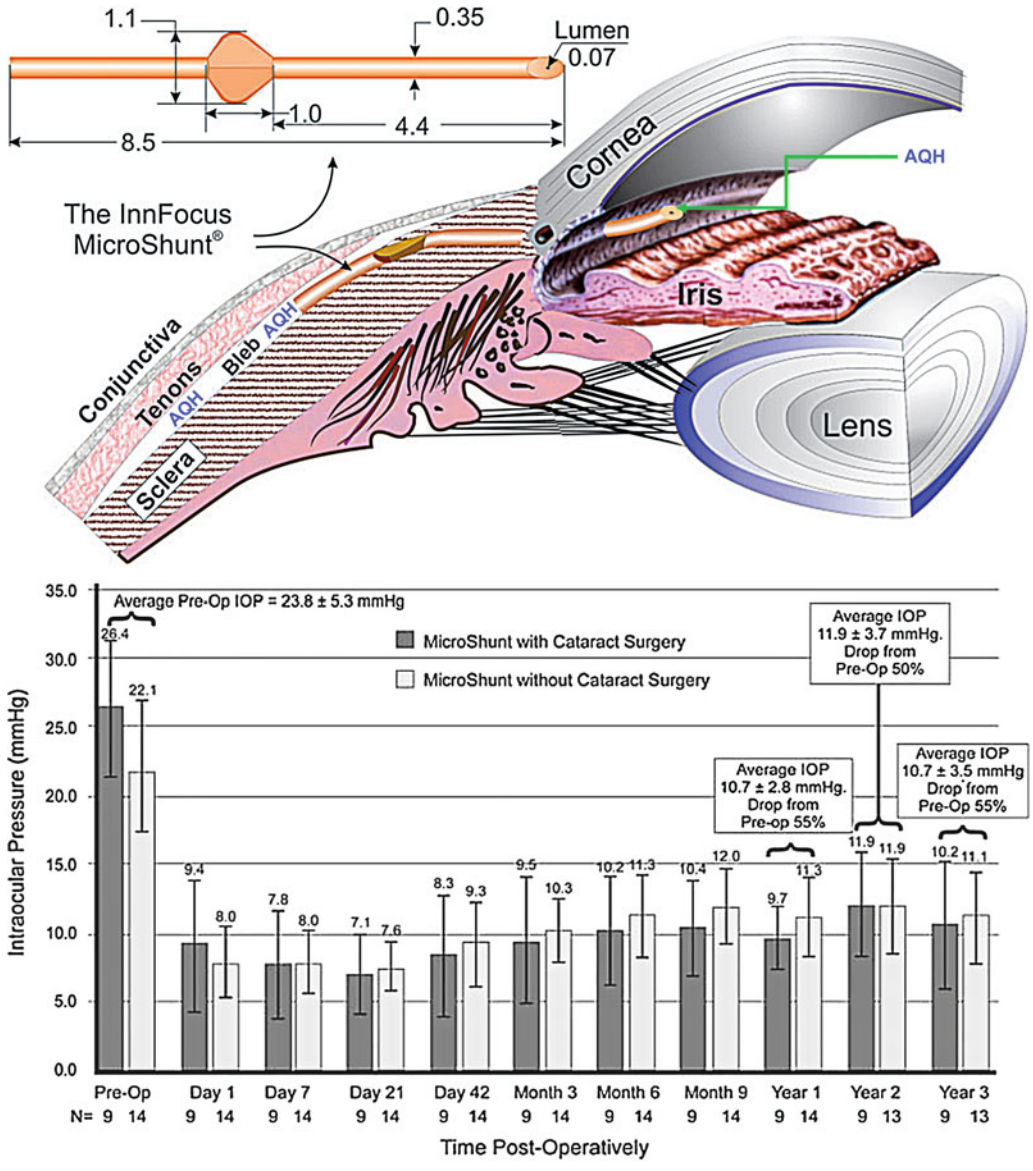


Fig. 11 Top panel shows the dimensions of and the placement of the InnFocus Microshunt in the anterior chamber of the eye to drain the AQH in order to lower IOP. The efficacy of this MIGS device is clearly demonstrated by the data in the lower panel in longitudinal studies performed in human OHT/POAG patients. (Modified from Batlle et al. (2016))

Thus, for instance, if a new small molecule needs to be discovered for an EP_2 or FP prostaglandin receptor, then ligand binding assays for all the known PG receptors and their subtypes, preferably human cloned or using fresh target tissues/cells, and using selective radiolabeled ligand, are established and validated. Then test compounds are used to compete for the receptor-radioligand complex in cell/tissue homogenates suspended in suitable buffers at equilibrium using rapid filtration techniques and thus evaluated for their relative affinities, relative selectivities, and then ranked ordered accordingly. Examples of such receptor-ligand binding assays and the data obtained are shown in Tables 1 and 2 (Sharif et al. 2003; Kirihara et al. 2018).

Since agonist compounds are being sought as novel ocular hypotensives, the compounds selected from the above assays are then tested for their ability to stimulate the desired receptor in freshly isolated target cells (preferably human CM or TM cells known to be involved in AQH dynamics and IOP regulation) (e.g., 5HT2-receptor-mediated intracellular Ca^{2+}-mobilization, Fig. 4, left panel), or in host cells transfected with the human cloned receptors of interest, and appropriate second messengers detected and quantitated. The data are graphed to determine the relative potencies of the test compounds relative to a positive control drug. Thus, for instance, cAMP (Crider and Sharif 2001; Kirihara et al. 2018) or cGMP (Katoli et al. 2010) or inositol phosphates (Sharif et al. 1994, 1996) or intracellular Ca^{2+} (Sharif et al. 2007) are used as receptor-induced signal readouts (Fig. 4, left panel; Fig. 8, right panel) and the receptor-specific blockade of these responses (e.g., Fig. 4, right panel). Other assays to examine functional activity in vitro of potential IOP-reducing agents involve use of cell- (Ramachandran et al. 2011) and/or tissue-based relaxation/contraction assays (Ohia et al. 2018) using cells/tissues known to be mediators of AQH dynamics and thus ocular hypotensive activity in vivo (e.g., CM and TM [Sharif et al. 2007, 2014] or SC [Dismuke et al. 2010] cells or CM/TM tissue strips [Wiederholt et al. 2000; Ohia et al. 2018]), coupled with cellular impedance changes that reflect cell relaxation/contraction (Wang et al. 2013) or cell-volume changes (Dismuke et al. 2009, 2010). Examples of some of these data for different classes of ocular hypotensive drugs are shown in Figs. 4, 8, and 9; Tables 3 and 4.

Testing of Compounds for AQH/Fluid Extrusion in Ex-Vivo Systems

The results from cell/tissue-based experiments are useful to define the agonist/antagonist nature of test compounds but they only provide a glimpse into the pharmacological activity of the latter. Therefore, researchers have utilized partially intact ANC segments of the eye from a number of species (human, bovine, and porcine) (e.g., Sharif et al. 2014) in culture to assess the ability of drugs to stimulate outflow of the fluid from such models. The perfusion buffer can also contain suitable inhibitors of receptors or enzymes in order to dissect potential molecular mechanisms of action of the perfused test agents.

Animal Models Used to Discover Novel Ocular Hypotensive Drugs

Testing of potential IOP-lowering agents is obviously best performed in animal models where the initial baseline IOP is naturally high so that the change in IOP induced by the test compound is more easily identified. To this end, there are only a few animal species that express a naturally high IOP and that are readily accessible. These include DBA/2J strain of mice, Dutch-Belt rabbits, and Beagle dogs (see McNally and O'Brien 2014; Sharif 2018a, b, for reviews). In the absence of the latter, researchers have had to artificially elevate IOPs in rodents and monkeys by a number of different techniques. For rodents, the TM can be partially or completely destroyed using injection of hypertonic saline via the episcleral vein, by injecting latex or magnetic microbeads into the ANC of the eye, or by lasering the TM or episcleral veins (reviewed in McNally and O'Brien 2014; Sharif 2018a, b). In the monkeys, the preferred reproducible method of raising IOP is by lasering the TM where the IOP can be elevated and maintained around 30–40 mmHg

for many months and in some cases for years. To obtain optimal results from testing new compounds for ocular hypotensive activity, it has been reported that the animals need to be unrestrained, well trained, comfortable, and conscious to have their IOPs measured on a frequent basis accompanied by suitable rewards. Some investigators anesthetize animals before they measure IOPs, but the data from such methods are somewhat suspect since unconscious animals have a different set of baseline IOPs and also respond differentially from fully awake animals. The age of the test animals also needs careful attention and requires standardization to older animals to ensure that OHT/POAG condition in humans is being well represented.

Ocular normotensive and ocular hypertensive (naturally or induced) animals are used to directly assess IOP-lowering potential of new test compounds. Due to cost and genetic considerations, rodent models of OHT have been preferred for primary screening using t.o. dosing of known and new compounds dissolved or suspended in suitable ocularly compatible and previously approved vehicle solvents/buffers. First, test compounds are freshly prepared in the vehicle of choice and evaluated using either single installation (20–30 μl drop) or after multiple administrations for ocular safety using rabbits and/or guinea pigs. Albino New Zealand White rabbits are preferred since guinea pigs are overtly too sensitive to most t.o. drugs. When rodents are used for IOP-lowering efficacy studies, they can also be used for ocular safety assessments as well. Compounds are considered safe when they do not induce excessive irritation/redness, excessive blinking, and discharge, and when no vocalization is recorded.

For the actual ocular hypotensive efficacy experiments, animals are grouped according to the treatment they will receive. The eye is first mildly numbed with proparacaine and IOP measured 3–4 times in a quiet and dimly lit environment using various pneumotonometers (e.g., Tono-Pen, Tono-Lab, Goldman appellation pneumotonometer) in order to establish the baseline IOP. Future IOPs are determined at the same time every day in the same room. Usually, a 5–10 μl drop of the vehicle or positive control drug (for rodents; or 20–30 μl drop for larger species) or the new test agent is instilled in one eye of the mice/rats and the IOPs recorded over time in each eye of each group of animals. A number of IOP readings are taken at each time point to ensure accuracy and reproducibility of the data. Those compounds that meet acceptable criteria of ocular safety (see above) and sufficient IOP reduction (e.g., 20% from baseline) are then tested in secondary models (e.g., ocular normotensive New Zealand White rabbits or naturally ocular hypertensive Dutch-Belted rabbits), and then ultimately in tertiary screening models (Beagle dogs; nonhuman primates such as Cynomolgus or Rhesus monkeys, either ocular normotensive and/or ocular hypertensive) (Fig. 12; Fig. 13, top panel) (Sharif et al. 2014; Kirihara et al. 2018; Fuwa et al. 2018). Typically, compounds that reduce IOP by >20% from baseline are considered worthy of further pursuit. Now, more detailed dose-response and mechanistic studies, involving testing of selective enzyme inhibitors and/or receptor antagonists can be undertaken (e.g., Sharif et al. 2014; Fig. 10). Additionally, AQH dynamic studies (Sharif et al. 2014) can be conducted in order to define whether the test compound(s) promote outflow of AQH from TM and/or UVSC pathways or inhibit the production of AQH (e.g., Fig. 13, bottom panel; Fuwa et al. 2018). In order for these results to be biologically and perhaps clinically relevant, the latter AQH modulation studies are usually performed in anesthetized ocular hypertensive nonhuman primates before being tested in human subjects in clinical trials.

Neuroprotective Therapeutics for Treating Glaucomatous Optic Neuropathy

It is unfortunate that despite use of many different classes of ocular hypotensive agents for multiple decades to lower and control IOP, patients with OHT/POAG (and even those with normal IOPs),

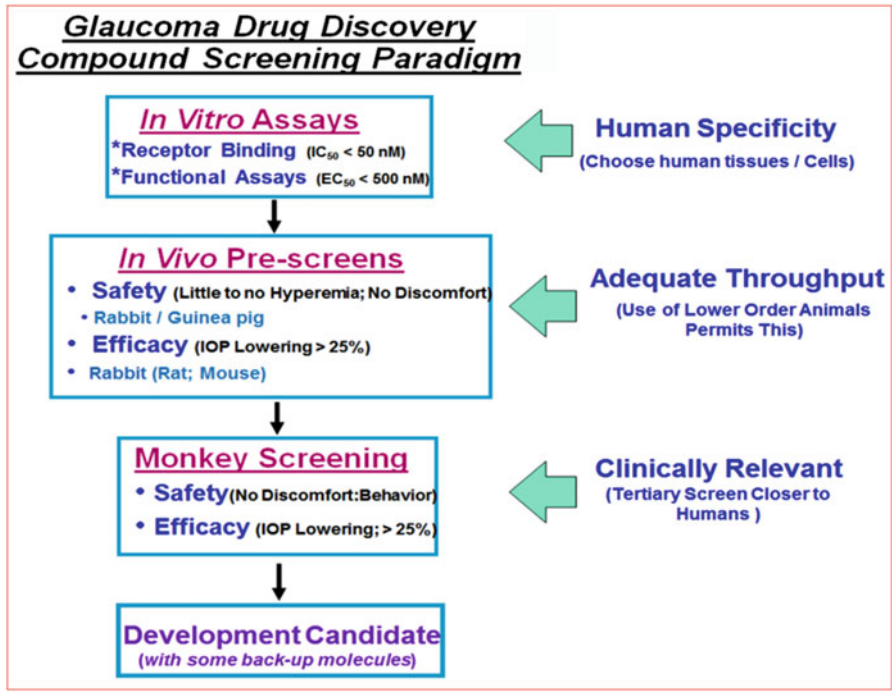


Fig. 12 A typical testing funnel for screening compounds directed at ocular hypertension and POAG is shown. Once a target protein is identified and incorporated into the in vitro screening paradigm, radioligand binding assays and/or functional cell-based assays are conducted in order to rank-order the test compounds in terms of their affinity and functional potency/intrinsic activity (agonist or antagonist). This information is used to select compounds to be tested for ocular/systemic toxicity/irritability and efficacy in rodents and/or rabbits. Those compounds meeting selection criteria are then tested in ocular normotensive and OHT monkey eyes for safety and efficacy. All this information is ultimately used to improve the structure–activity relationship of compounds by medicinal chemists. This iterative process is expected to yield clinical candidate drug molecules

continue to lose vision, and some may ultimately succumb to blindness. Thus, it is now accepted that direct protection of the retinal neurons (especially RGCs), their axons and terminals within the brain need to be strongly considered in addition to IOP reduction. Due to the multiplicity of damaging factors and insults impacting the RGC-brain axis that results in glaucomatous optic neuropathy (GON) (Fig. 14), including constriction of optic nerve axons at ONH (Hollander et al. 1995), complement activation (Tezel et al. 2010), locally released inflammatory/toxic substances (excess neurotoxic amino acids, cytokines, endothelins, NO, etc.) and tissue/cell remodeling enzymes (MMPs; calpains; caspases), blockage of neurotrophin (Quigley et al. 2000) and mitochondrial transport up and down the RGC axons (hence loss of energy [Thomas et al. 2000], etc., a multipronged neuroprotective strategy is necessary to preserve RGCs and their axons. As far as optic neuropathy (ON; encompassing Leber's hereditary ON, nonarteritic ischemic ON) is concerned, there has already been a number of a drug launched/approved such as the powerful antioxidant and Ca^{2+}-channel blocker idebenone, with others in late-stage clinical trials (e.g., lenadogene/nolparvovec (mitochondrially encoded NADH dehydrogenease-4 expression enhancer), QPI-1007 (caspase-2 expression inhibitor), and elamipretide (apoptosis inhibitor). Other novel strategies for protecting RGCs and optic nerve components are discussed in more detail in recent publications (Smedowski et al. 2016; Hines-Beard et al. 2016; Venugopalan et al. 2016; Williams et al. 2017; Mead et al. 2018; Sharif 2018b).

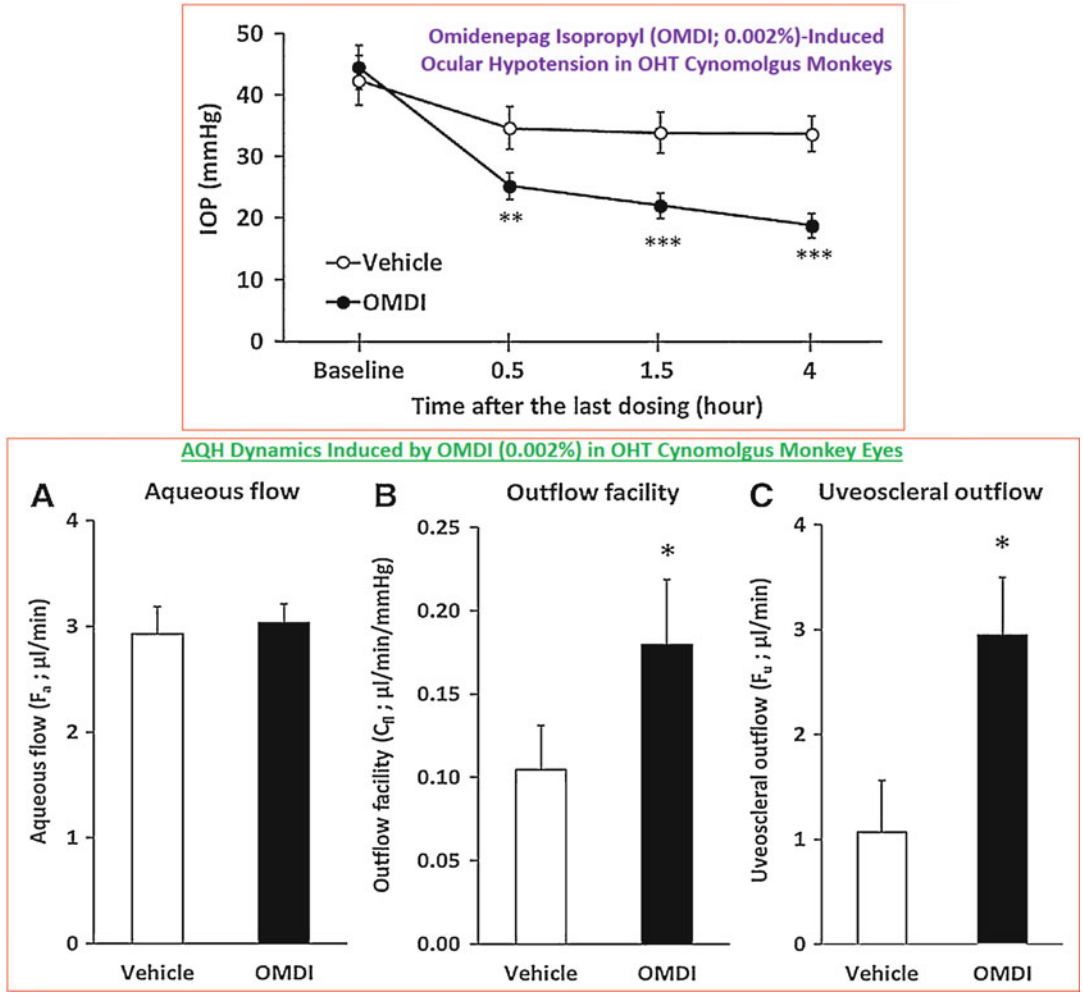


Fig. 13 A new-generation non-PG EP_2-receptor agonist omidenepag isopropyl (OMDI) and its IOP-lowering effects in monkeys (top panel), and its stimulation of both conventional outflow and uveoscleral outflow in OHT eyes of monkeys (lower panel) is shown.(Modified from Fuwa et al. (2018))

Cell-Based Assays and Animal Models for Discovering Neuroprotective Drugs

A number of cell/tissue-based assay systems have been developed to assess the potential protective activity of selected compounds of interest. These have included isolated primary rat RGCs, co-cultures of RGCs and other retinal cells, ONH or retinal astrocytes, surrogate CNS neurons, retinal explants, whole retinas, and whole retina-optic nerve explants (see He et al. 2018; Sharif 2018b for recent reviews). Despite the fact that GON is caused by a plethora of chemical, biological, mechanical, and local environmental factors simultaneously or in a close time frame at the level of the RGCs, their axons and terminals (Fig. 14), the majority of the reported test systems investigating neuroprotection in vitro have only introduced a single insult to the retinal/axonal cells/tissues. The types of challenges used have included hypoxia (to mimic retinal ischemia), elevated hydrostatic pressure (to mimic high IOP), glutamate- or NMDA- or TNFα- or IL-1-β-induced cell death (to recapitulate neuronal toxicity), glucose or neurotrophic factor withdrawal (to mimic hypoglycemia as a result of ischemia

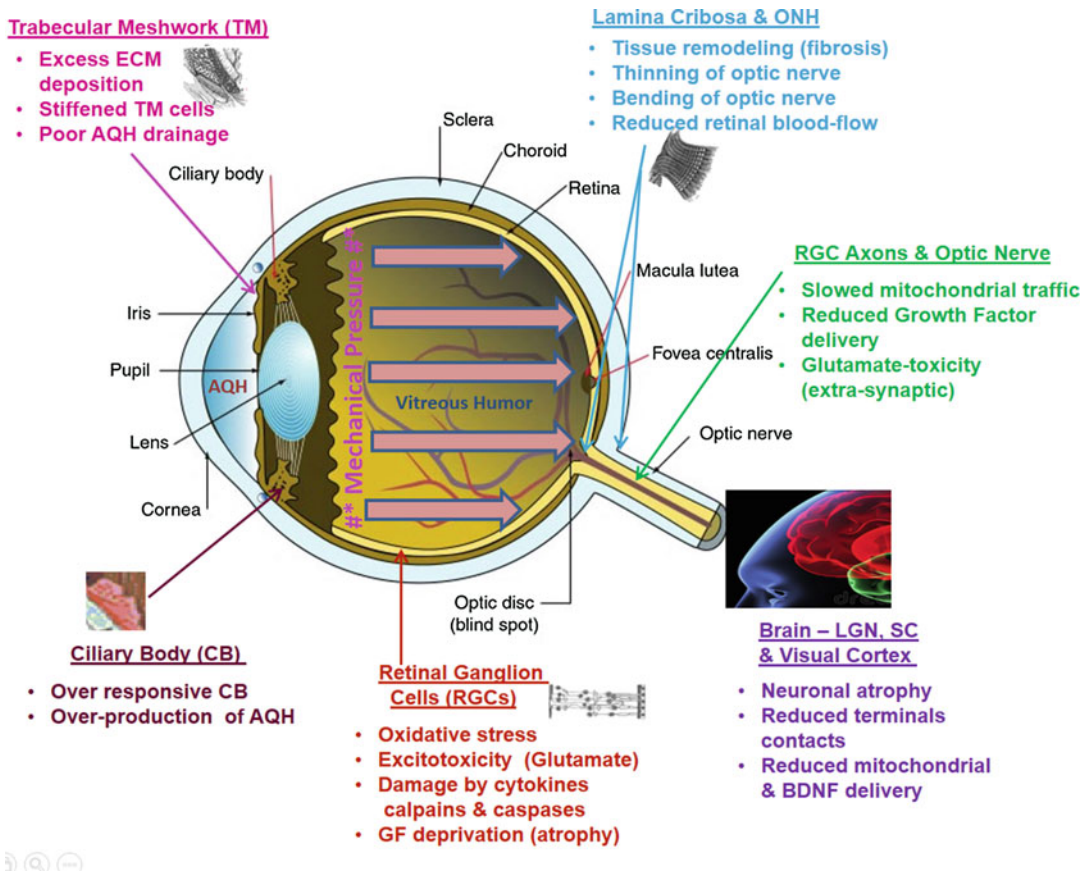


Fig. 14 Illustrationof the numerous sites within the eye-brain axis where various pathological events (e.g., mechanical, chemical, and environmental) occur to damage the anterior chamber cells (e.g., TM cells), retinal cells (e.g., RGCs), RGC-axons, and the optic nerve connections to the brain. The net result of such damage is progressive visual impairment that can ultimately lead to blindness as in various forms of glaucoma. (Modified from Sharif (2018b))

(Osborne et al. 2014), and axonal/optic nerve constriction/blockage, respectively). Techniques and types of signal readouts that have provided means to gauge neuroprotective activity of test compounds in these in vitro systems included the following: cell viability (cellular toxicity) assays using measurement of extracellular lactate dehydrogenase and/or cytochrome c oxidase (released when cell membranes become compromised due to cell ill health), quantifying membrane potential using JC-1 dye, detecting mitochondrial destruction by fission, quantitating mitochondrial viability by cyan fluorescent protein labeling, monitoring cell apoptosis using the TdT-mediated dUTP nick-end labeling (TUNEL assay), Brn3A-staining, propidium iodide labeling to assess cell death, measuring caspase-1/3 activity, multiple electrode array recording of cellular activity, measuring cellular levels of ATP using nuclear magnetic resonance, calcein AM staining, gelatin zymography for MMPs and other proteases, inhibition of [^{3}H]-D-aspartate release from retinas as an index of neurotoxicity prevention, neurite elongation using high content screening, etc. (see references in: He et al. 2018; Sharif 2018b). If compounds show cytoprotective activity in a number of these assay systems, then chances are high that they will exhibit some level of neuroprotective efficacy in vivo.

Animal models of GON are somewhat limited and generally tend to be labor-intensive, yielding data of variable value. The major reason for this is again the singularity approach of chemical, mechanical, and other insults/challenges used to evaluate the in vivo neuroprotective efficacy of compounds. This necessitates the utility of multiple animal models to robustly assess the neurotherapeutic activity of any compound that may ultimately be used in a clinical setting.

Despite the above issues, much progress has been made in finding novel compounds and new pathways that represent useful intervention points with therapeutic end points for testing in animal models of retinal/optic nerve damage as encountered in GON/POAG. Consequently, numerous classes of compounds have been qualified as neuroprotective based on their ability to reduce the loss of RGCs and/or their axons. The following represent some of the most commonly used animal models to study GON and to screen compounds for their neuroprotective efficacy: rodent models of acute/chronically elevated IOP that results in retinal ischemia; rodent models of partial transection or crush of the optic nerve at the level of the ONH; ivt injections of neurotoxins such as NMDA, endothelin, amyloid-beta peptides, or staurosporin, or phorbol ester to capitulate endogenous inflammatory reactions common in GON; uveitic glaucoma model and inflammatory demyelination models (see references in Sharif 2018b). In the majority of the cases, the number of RGCs and their axons are quantified postmortem to assess the degree of damage/protection in control *vs.* the treated animals. This is achieved by retrograde-labeling of RGCs using fluorogold- or Brn3A-labeling after injection of the latter markers into the superior colliculus of the animals, and by axonal counts in transverse sections of the optic nerve, respectively.

In some cases, intravenously injected fluorescently labeled annexin-5 has been used to monitor and quantify retinal cells undergoing apoptosis in living animals subjected to various experimental challenges pertinent to GON as described above. Interestingly, this technique has now been used in patients with POAG/OHT and certain neurodegenerative diseases and useful baseline data gathered (Cordeiro et al. 2017). Perhaps such diagnostic/prognostic biomarkers can be utilized to evaluate neuroprotective drugs in nonhuman primates and human subjects in the near future. Likewise, the recent use of flavoprotein fluorescence to monitor mitochondrial health of retinal cells in vivo in POAG and control patients appears a promising tool for assessing cytoprotective actions of compounds believed to possess neuroprotective efficacy (see references in Sharif 2018b). Furthermore, the use of novel technologies such as high-resolution adaptive optics and visible-light-OCT coupled with standard functional readouts like visual-evoked potential measurements will enhance our understanding of the GON and lead to better therapies in the future.

Age-Related Macular Degeneration (AMD)

Another ocular disease that has a relatively rich history of drug discovery and pharmacodynamics associated with it is age-related macular degeneration (AMD) (Lambert et al. 2016). Unlike POAG where peripheral vision is lost, in AMD the loss of macular photoreceptors impacts central vision. AMD is believed to be responsible for nearly half of all severe vision loss in the US adults over the age of 40. AMD has been divided into two forms: nonexudative or "dry" AMD (dAMD; 90% of total) and exudative or "wet" AMD (wAMD; 10% of the total). While dAMD is characterized by the loss of photoreceptor cells in the macula following the death of supporting RPE cells, wAMD's hallmarks are retinal edema and rampant neovascularization of choroidal capillaries (choroidal neovascularization or CNV) (Al-Zamil and Yassin 2017; Hernandez-Zimbron et al. 2018). Such aberrant angiogenesis causes retinal fibrosis, scar formation (and perhaps retinal detachment), culminating in loss of central visual acuity.

While dAMD develops over months and years, wAMD is highly progressive and rapidly develops to rob vision of the patient unless treatment is sought. Advancing age is the strongest demographic risk factor associated with AMD

although Caucasian heritage predisposes patients to this ocular disease. However, chronic excessive oxidative stress from cigarette smoke and other sources, autoimmune disease involving complement activation, and chronic local inflammation are also causative factors in the development of dAMD (Rickman et al. 2013). An increased blood plasma concentration of the proinflammatory proteins C-reactive protein, IL-6, cholesterol/triglycerides, and a family history of AMD also been positively correlated to some degree with AMD development and progression. Recent angiography-coupled-OCT evidence suggests that poor choriocapillary blood flow leads to poor clearance of cellular and other debris and may be the source of deposited drusen and related materials that ultimately lead to dAMD/GA (Qin et al. 2018).

RPE cell dysfunction likely begins with intralysosomal accumulation of a fluorescent material called lipofuscin, a complex mixture rich in polyunsaturated lipids and probably is derived from phagocytosed photoreceptor outer segments that cannot be broken down. Over time, this material renders lysosomal enzymes inactive and raises the pH level causing lysosomal membrane dysfunction and RPE cell death. A likely functionally important component of lipofuscin is the amphiphilic pyridinium ion N-retinyldene-N-retinylethanolamine (A2E) which is generated from the condensation of phosphatidylethanolamine with 11-*trans* retinaldehyde, followed by phospholipase D-catalyzed dephosphorylation. It has been postulated that A2E might be the major toxic chemical in lipofuscin producing reactive oxygen species and oxiranes in the presence of light and O_2 (Fig. 15). Additionally, A2E likely acts as a detergent causing leakage of toxic reagents into the RPE cell, thereby killing the latter cells.

Once RPE cells are injured, their ability to phagocytose photoreceptor cell outer segments is hindered and this causes the incomplete recycling of the components of the latter. This is responsible for accumulation of cellular debris, and the production and release of additional inflammatory

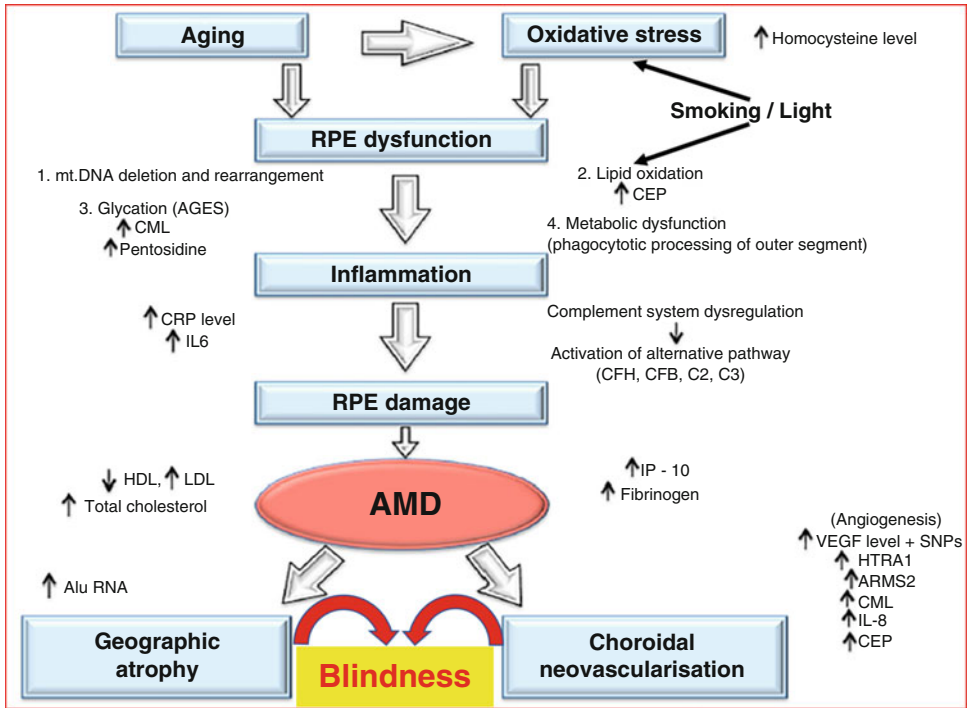


Fig. 15 Pathogenesis of dAMD/GA. (Modified from Ambati et al. (2003)).

agents from the dying RPEs such as oxidized lipids, proinflammatory cytokines (e.g., IL-1β and IL-17A [Zhang et al. 2016]), and acute-phase inflammatory proteins, that accumulate between the RPE cell layer and the Bruch's membrane that separates the RPEs from the choroidal capillaries. All this deposited toxic material (drusen) then kills more RPE cells and the homeostasis is disrupted further (Rickman et al. 2013) (Fig. 15). The drusen and other cellular debris also act as a barrier and less O_2 and nutrients are made available to the rest of the retina, and this results in the development of a hypoxic environment at the back of the eye. In order to overcome this situation, hypoxia-inducible transcription factors like HIF-1α are upregulated and vascular endothelial growth actor (VEGF) and other angiogenic factors are locally released that cause the generation of new blood vessels from the existing choroidal capillaries. This aberrant neovascularization temporarily helps remove some of the metabolic and cellular waste and provides nutrients and O_2 to the remaining RPE and photoreceptor cells. However, the new blood vessels breach the Bruch's membrane and start to grow into the retina and eventually into the vitreous and start to interfere with light transmission to the retina and the communication among the inner retinal cells. Additionally, since the new blood vessels are leaky, blood starts to accumulate in the vitreous and local hemorrhages develop at the rear of the globe that begins to detach the retina. The patient loses more and more visual acuity and requires urgent care to curb the loss of all vision in that eye. This is how dAMD can cause wAMD. However, there are patients who do not develop wAMD, and their dAMD keeps progressing till their retina has widespread drusen deposited that is characterized as geographic atrophy (GA) (advanced dAMD) that robs the patient's central and eventually peripheral vision and thus causes irreversible blindness (Fig. 16).

As to the possible treatment options for combating dAMD/GA, there is currently no accepted pharmacological treatment in use in the clinical setting. While dietary supplementation with a combination of high-dose vitamins E and C, beta-carotene, and zinc afforded a 25% risk reduction for progression of high risk dAMD patients to wAMD form, the interpretation of the results of this study has been controversial. It is thought, however, that based on the involvement of oxidative stress and inflammation in the etiology of dAMD that drugs that can reduce the severity of the latter conditions may have beneficial effects in dAMD patients. Therefore, additional research and clinical trials are in progress using a number of agents and treatment options for dAMD/GA which are at various stages of development (Ishikawa et al. 2015; Waugh et al. 2018) (e.g., Table 7).

Several wAMD treatments centered around reducing levels of VEGF have been successfully developed and introduced into the clinical management of the disease including use of bevacizumab and ranibizumab (both truncated antibodies bind VEGF and remove it), and aflibercept (VEGF-receptor as a VEGF trap) (Ishikawa et al. 2015; Al-Zamil and Yassin 2017; Hernandez-Zimbron et al. 2018). Despite the undeniable success of the anti-VEGF treatment modalities, there is increasing concern about the reported development of tolerance/resistance to these medications after the 2nd-year of treatment (Maguire et al. 2016; Yang et al. 2016). Thus, it is imperative that additional therapeutic agents be discovered and developed to mitigate such issues. Consequently, several therapeutic strategies have been proposed to reduce and/or prevent the development and progression of wAMD that appears to be driven by the angiogenic factors such as HIF-1α and VEGF (Table 8), along with other emerging growth factor culprits such as angiopoietins that trigger angiogenesis via the Tie-1/2 receptors. It may be necessary to also begin combinatorial therapy for wAMD (e.g., endogolin + anti-VEGF combination [Shen et al. 2018]), as was the case for OHT/POAG when it was realized that conventional single-agent therapies were unsuitable for recalcitrant patients and also those who were not maximally controlled by monotherapeutic agents (see above). Likewise, the novel cell-replacement therapies involving growth of polarized monolayer of human embryonic stem cell-derived RPE (hESC-RPE) cells on an ultrathin parlene substrate and placing them into human subjects' retinas with severe vision loss due to dAMD/GA as a cell-patch (Kashani et al. 2018), or placing similarly engineered hESC-RPE patches

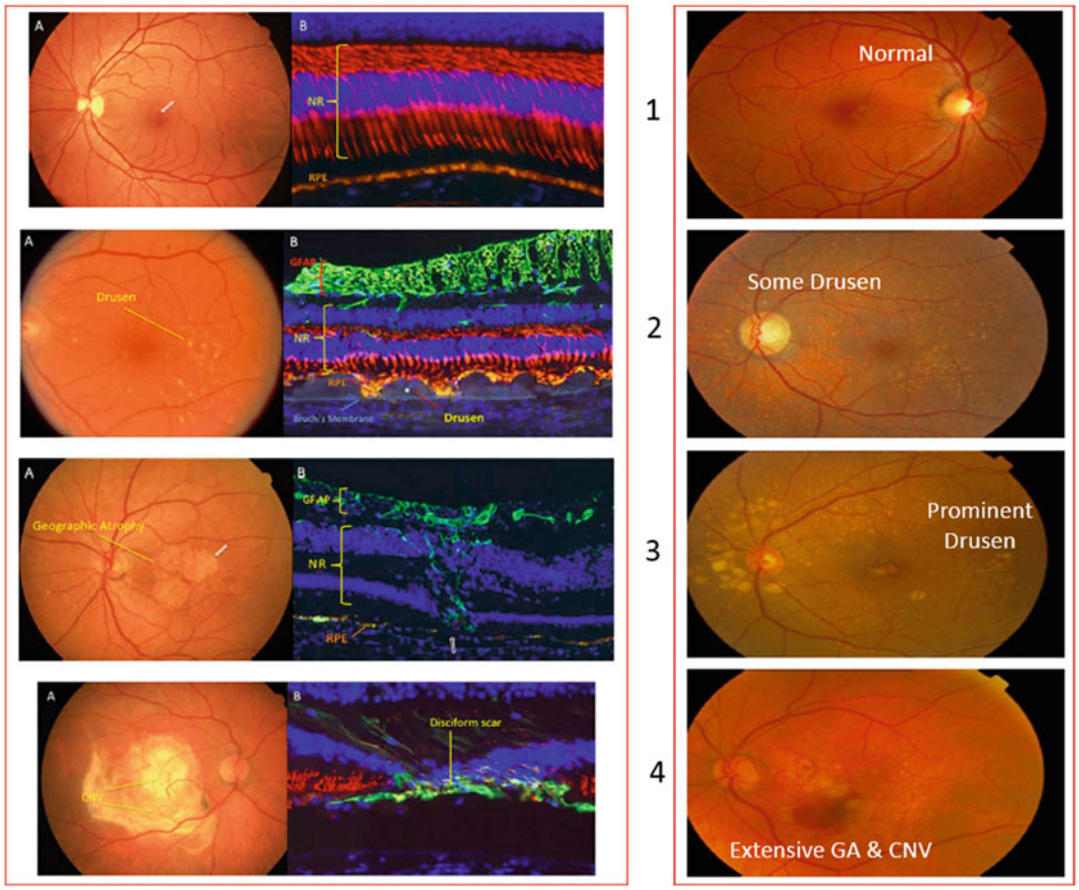


Fig. 16 Pictorial depiction, via fundus photos, of human dAMD progressing to widespread GA due to increasing drusen deposition in the retina.(Modified from Ambati et al. (2003))

into subretinal space of patients with wAMD (da Cruz et al. 2018) to improve visual acuity holds tremendous promise.

Assay Systems Deployed for Anti-dAMD/Anti-GA Drug Discovery

As with OHT/POAG and GON, cell-based assays utilize anatomically relevant retinal cells known to be involved in the etiology of or as comprised cell type(s) in the disease process of dAMD/GA for screening purposes. Since the cascade of events linked to the disease involve inflammation, deposition of ECM in Bruch's membrane, and drusen accumulation in the RPE cells, the latter cells have been used as target cells to study ways to prevent the former and latter phenomena. Anderson et al. (2013) used ARP-19 cells and tested various components of drusen (e.g., carboxyethypyrrole (CEP)-modified proteins, amyloid-β (1-42), Nε-(carboxymethyl) lysine (CML)-modified proteins, and aggregated vitronectin) and the key component of lipofucin (A2E) and monitored production of major inflammatory cytokines, chemokines, and VEGF-A as biomarkers. They concluded that A2E was the most active proinflammatory substance studied, and that it promoted the release the aforementioned biomarkers from RPE-19 cells by activating the inflammasome (NLRP3/caspase-1) pathway activation. Such a screening tools can thus be used to find suitable blockers of the NLRP3/caspase-1 system, and these may prove useful in ameliorating the dAMD/GA.

Table 7 Therapeutic agents, devices and other treatments under development for dAMD/GA

Mode of action	Therapeutic agent/ product	Sponsoring institution	Disease indication	Development phase	Route of delivery
Small molecules					
Visual cycle modulator	Emixustat HCl	Acucela	GA/dry AMD	Phase-2b/3	Oral daily tablet
Modulating choroidal blood flow	MC-1101	Macuclear	Dry AMD	Phase-2/3	Topical, twice daily
Tetracycline antibiotic 40 mg doxycycline	Oracea	University of Virginia	GA/dry AMD	Phase-2/3	Daily oral capsules
Antioxidant; slows DNA damage, reduces ROS levels	Metformin	University of California San Francisco	Nondiabetic GA/dry AMD	Phase-2	Daily oral tablets
Aptamers/peptides/antibodies					
mAb fragment for complement factor D	Lampalizumab	Genentech	GA/ advanced AMD	Phase-3	Intravitreal injection
Inhibition of complement C3	APL-2	Apellis Pharma	GA/AMD	Phase-2	Intravitreal injection
Antibody for C5 complement	LFG 316	Novartis	GA/AMD	Phase-2	Intravitreal injection
Aptamer; inhibits complement Factor C5	Zimura	Ophthotech	GA/dry AMD	Phase-2/3	Monthly intravitreal injection
Hu-mAb for Aβ peptide	GSK933776	GlaxoSmith-Kline	Retinal amyloidosis/ GA/dry AMD	Phase-2a	Intravenous infusion
Cellular therapeutics					
Stem cell transplantation	Hu-CNS SC	StemCells Inc.	GA/ Advanced dry AMD	Phase-1/2	Subretinal transplantation
hESC-derived RPE cells	MA09-hRPE	Ocata Therapeutics – Astellas Pharma	Advanced dry AMD	Phase-1/2	Subretinal transplantation
Human umbilical tissue-derived cells	CNTO 2476	Janssen Research and Development	GA/AMD	Phase-1/2a	Subretinal administration
hESC-derived RPE cells seeded on polymeric substrate	CPCB-RPE1	Regenerative Patch Technologies (RPT)	Advanced dry AMD	Phase-1/2	Subretinal implantation
hESC-derived RPE cells	OpRegen	Cell Cure Neurosciences	Advanced AMD	Phase-1/2	Subretinal transplantation
Patient-derived iPSC transplantation	iPSC-derived RPE cells	Moorefields Eye Hospital NHS Foundation Trust	Wet and dry AMD	Phase-1/2	Transplantation
Autologous bone marrow–derived stem cells	BMSC- SCOTS study	Retina Associates of South Florida and MD Stem Cells	AMD	Early stage interventional study	Sub-Tenon injection

(continued)

Table 7 (continued)

Mode of action	Therapeutic agent/ product	Sponsoring institution	Disease indication	Development phase	Route of delivery
Devices and other treatments					
Photobiomodula-tion or low light level therapy	LumiThera LT 300 light delivery system	LumiThera	Vision loss associated with AMD	Early stage interventional study	Light exposure
Electrical stimulation	Argus-II System	Second Sight Medical Products	Late-stage AMD	Phase-1	Implantable visual retinal prosthesis
Transpalpebral microcurrent electrical stimulation	Nova Oculus	The Eye Machine Canada	Vision loss associated with dry AMD	Early stage interventional study	Externally applied microcurrent electrical stimulation
Electrophysiologic methods (double plasma filtration or cascade filtration)	Rheohemapheresis	University Hospital Hradec Kralove	High-risk dry AMD	Phase-4	Cascade filtration

Abbreviations: *AMD* age-related macular degeneration, *GA* geographic atrophy, *hESC* human embryonic stem cell, *iPSC* induced pluripotent stem cell, *mAb* monoclonal antibody, *MOA* mode of action, *ROS* reactive oxygen species, *RPE* retinal pigment epithelium

Primary human RPE (ph-RPE) cells represent a better cellular system than the cell line ARPE-19. Zhang et al. (2016) demonstrated that ph-RPE cells expressed all three IL-17 receptors and that addition of IL-7A to these cells upregulated the production of IL-1β secretion via the NLRP3 inflammasome activation mechanism. Importantly, these authors found that inhibiting caspase-1 activity and silencing NLRP3 significantly reduced IL-1 β release from RPE cells. Thus, this assay system can be used to find new potent and efficacious blockers of caspase-1 and NLRP3 to help prevent dAMD/GA. Additional work in this area using human patient-derived iPSC-RPE cells is also very encouraging (Galloway et al. 2017, 2018).

Animal Models to Find Anti-dAMD/ Anti-GA Drugs

While not truly reflecting the human dAMD/GA disease, a number of animal models have been developed to study the condition and use for potential drug discovery efforts. The light damage models utilize rodents and expose them to very bright light (white or blue) for a number of days/ weeks and the retinal damaged assessed by electroretinograms and histology (Chader 2002). The genetic rodent models include the Royal College of Surgeons (RCS) rats (a recessive genetic defect that prevents phagocytosis of rod-outer-segments by RPE cells), P23H rhodopsin defect rat, and an ABCR-1- rat (which has the transporter of 11-trans-retinaldehyde (ABCR-protein) knocked-out). The most appealing of the animal models created thus far is the chemokine receptor-2 knockout mouse which exhibits many of the hallmark features of human dAMD/GA (drusen accumulation under RPE, photoreceptor loss followed by CNV; Ambati et al. 2003). How these animal models are exploited to discovery novel therapeutics for treating dAMD/GA remains to be seen.

Cell-Based Assays for Finding New Anti-wAMD/Anti-CNV Drugs

As described above, wAMD involves poor retinal circulation leading to retinal hypoxia, release of angiogenic factors (e.g., HIF-1/2α; VEGF; ANG-2), and abnormal growth and development of leaky new blood vessels from the choroidal

Table 8 Recently approved and emerging treatment options being pursued for GA and neovascular AMD

Drug agent	Drug class	Developer/ sponsor	Targeted pathology	Mechanism of action	Route of administration
Ranibizumab (Lucentis)	Anti-VEGF antibody	Novartis	wAMD/ Neovascular AMD	Removes VEGF	Intravitreal
Aflibercept (Eylea)	VEGF-Receptor (VEGF-trap)	Regeneron	wAMD/ Neovascular AMD	Removes VEGF	Intravitreal
Conbercept	VEGF-Receptor (KH-902; Biosimilar)	Chengdu Kanghong	wAMD/ Neovascular AMD	Removes VEGF	Intravitreal
Brolucizumab	Anti-VEGF antibody (RTH-258)	Novartis	wAMD/ Neovascular AMD	Removes VEGF	Intravitreal
Lampalizumab	Anti-factor D Fab	Genentech	GA	Anti-factor D Fab	Intravitreal
MA09-hRPE	Cell therapy	Astellis Pharma	GA	Human umbilical tissue-derived cells	Subretinal injection
Brimonidine tartrate implant	α2-adrenoceptor agonist	Allergan	GA	Alpha-2-agonist	Intravitreal implant
Eculizumab	Anti-C5 mAb	Alexion	GA	mAb against complement factor C5	Intravitreal
Fovista (E10030)	Anti-PDGF aptamer	Ophthotech	Neovascular AMD	Anti-PDGF PEGylated aptamer	Intravitreal
Abicipar pegol	Anti-VEGF Aptamer	Allergan/ Molecular Partners	Neovascular AMD	Anti-VEGF	Intravitreal injection
RBM-007	Anti-FGF aptamer	Ribomic	Neovascular AMD		
PAN-90806	RTK Inhibitor	PanOptica	Neovascular AMD	Kinase inhibitor	Topical ocular
RXI-109	rX-RNA	RXi Pharma	Neovascular AMD	CTGF expression inhibitor	Intravitreal
Sunitinib	RTK Inhibitor	Graybug	Neovascular AMD	Multikinase inhibitor	Topical
APL-2	C3 Inhibitor	Apellis	Neovascular AMD	Complement-C3 Inhibitor	Intravitreal
Faricimab	Bispecific Ab	Chugai	Neovascular AMD	Anti-VEGF-A/ Ang-2 Ab	Intravitreal
RGX-314	Gene therapy vector (VEGF neutralizer)	RegenxBio	Neovascular AMD	Anti-VEGF AAV	Intravitreal
Retinostat	Gene therapy vector (Angiostatic stimulator)	Oxford Biomedica	Neovascular AMD	Lentivirus vector	Intravitreal

Abbreviations: *Ab* antibody, *Ang-2* angiopoeitin-2, *AMD* age-related macular degeneration, *CTGF* connective tissue growth factor, *GA* geographic atrophy, *mAb* monoclonal antibody, *PDGF* platelet-derived growth factor, *VEGF* vascular endothelial growth factor

system. Such CNV compromises Bruch's membrane (due to release of MMPs from RPE and Muller cell that digest the latter), retinal interneurons and eventually RGCs as the new blood vessels branch out into the vitreous and cause local hemorrhages and loss of vision, especially at the macula. Therefore, pathologically relevant cell types (e.g., human micro- and macrovascular

retinal endothelial cells [REC]; RPE [including primary cells and ARPE-19 cells, and Muller glial cells), surrogate cells (e.g., human umbilical vein endothelial cells (HUVEC)), and co-cultures (Chen et al. 2017), have been utilized to evaluate new drug modalities directed at the neovascular component of wAMD. Functional readouts relevant to the CNV using in vitro assays have involved RPE-stretch-induced release of VEGF (Farjood and Vargis 2018), growth factor-induced REC proliferation/migration/tube formation using a matrigel assay (Chen et al. 2016), cellular permeability, and expression and secretion of various MMPs (Di and Chen 2018).

Animal Models to Find Anti-wAMD/Anti-CNV Drugs

Aberrant ocular angiogenesis is the hallmark of CNV/wAMD. Animal models that have been used to study the pathological and drug discovery aspects of these diseases have involved laser-induced and surgically induced CNV in rodents, rabbits, and monkeys (reviewed by Liu et al. 2017). Retinopathy of prematurity (ROI) as induced by high levels of oxygen soon after birth of rodents (OIR) is another useful model of CNV and wAMD (Liu et al. 2017). These models are self-explanatory and have been deployed to study effects of disease-prevention/reduction using vascular permeability, measurement of VEGF release, and angiogenesis as biomarkers and readouts. In view of the importance of these retinal disorders, a number of transgenic mouse models with spontaneous sub- and intraretinal angiogenesis have also been established for possible screening purposes where other angiogenic signaling molecules may be involved other than or in addition to VEGF (Liu et al. 2017). Due to the aberrant involvement of the complement system in the etiology of wAMD/CNV, an important interaction between the oxidative stress and the latter systems is also being studied (Du et al. 2016), as is the measurement of numerous cytokines and chemokines using multiplex and microarray technologies to find additional targets for wAMD/CNV conditions that may lead to discovery and development of new medicines to treat these blinding disorders (Liu et al. 2016; Lambert et al. 2016).

Diabetic Macular Edema and Diabetic Retinopathy

Diabetes-induced retinopathy is also a major ocular disease that causes preventable blindness around the world, currently estimated at 93 million worldwide in total. Two forms of diabetic retinopathy (DR) have been described: non-proliferative (NPDR) and proliferative (PDR). Abnormal neovascularization induced by prolonged hyperglycemia and retinal hypoxia causes PDR which results in vision loss (Wang and Lo 2018). Diabetic macular edema (DME) is a hallmark of NPDR and corticosteroids and biologics are used in its treatment (Wang and Lo 2018). Since many of the signs and symptoms of PDR resemble those of wAMD, anti-VEGFs biologies (antibodies (ABs) and small AB fragments [Fabs]) have paved the therapeutic pathways for treatment of PDR in addition to traditional laser photocoagulation of leaky retinal blood vessels. Similarly, the side effects associated with the latter, and the tachyphylactic responses to anti-VEGFs observed in wAMD patients will also necessitate requirement of alternative treatment options for DR and DME. Some novel approaches encompass the explorative use of bispecific antibodies (RO-6867461 [Anti-Ang-2 + Anti-VEGF]), AKB-9778 (Tie-2 activator), EBI-031 and Tocilizumab (IL-6 inhibitors), Luminate (Integrin inhibitor), MTP-131 (cardiolipin inhibitor), Lutein, and ALA (mitochondrion-specific antioxidant) (Wang and Lo 2018). Recently, it has been proposed that not only are the components of aberrant neovascularization and inflammation important in the etiology and ultimate resolution of PDR and NPDR but that degeneration of the neural retinal cells is also a key element leading to blindness resulting from diabetes (Simo et al. 2018). Therefore, interventional and perhaps even prophylactic neuroprotective therapies should be considered for both wAMD and DR.

Assay Systems and Animal Models for Discovering New Treatments for DR and DME

In an effort to explore possible means to ameliorate the effects of DR and DME, researchers have focused attention on identifying possible interventions point of these retinal disorders. Since hyperglycemia is the core of the problem, the effects of high glucose on RPE, Muller glia, vascular endothelial cells, and pericytes have been studied and used for screening compounds for reducing various biomarkers of disease (e.g., secretion of cytokines and angiogenic factors, cell permeability, mitochondrial dysfunction; microglial activation, RPE, and REC tube formation) (e.g., Tien et al. 2017; also see above in section "Cell-Based Assays for Finding New Anti-wAMD/Anti-CNV Drugs"). While none of these in vitro systems truly capitulates the DR/DME, direct effect of potential drugs of benefit in treating the latter have been examined using a variety of animal models (chemically [alloxan/streptozotocin/diet]/surgically [pancreatomy]-induced and genetic) ranging from zebrafish to rodents, to cats, dogs, pigs, and monkey; Olivares et al. 2017). It is worth mentioning that alloxan/streptozotocin/diet-induced retinal disease outcomes resemble human DR/DME, where rodents generally exhibit the major defects in terms of hyperglycemia, damaged pancreatic beta-cells, damaged/reduced pericytes, increased acellular capillaries, microglial changes, basement membrane thickening, microaneurysms, RGC, and inner retinal cell loss (Olivares et al. 2017). Retinal neovascularization was observed in rodents after systemic hyperglycemia and was most profound under hypoxic conditions in zebrafish, rodents, and monkey (Olivares et al. 2017). Of the genetically induced DR/DME, rodents were the most susceptible showing many of the aforementioned phenotypic changes in their retinal anatomy and pathology. These results emphasize the need to use multiple assays and animal models to assess the therapeutic efficacy of any treatment modalities for DR/DME.

Ocular Surface Diseases

Despite the protection afforded by the blink response, the placement of the eyeball within the orbital socket, and protection provided by the sclera, the eye still remains a target for airborne allergens, pollutants, bacteria, and viruses that fall onto the ocular surface. The tear film covering the cornea and conjunctiva also acts as a barrier but can harbor some of the agents mentioned above.

Allergic Conjunctivitis

Seasonal allergic conjunctivitis (SAC) and perennial allergic conjunctivitis (PAC) are allergic reaction of the cornea and conjunctiva to airborne allergens such as pollen, mold, pet dander, and air pollution (Yanni et al. 1999; O'Brien 2013). SAC afflicts millions of patients of all ages every few months and causes debilitating and extremely bothersome excessive tearing, intense itching, grittiness, burning, photophobia, redness and swelling of the eyelids (O'Brien 2013; Gomes 2014). These symptoms are caused by release of histamine, prostaglandins, cytokines, and chemokines from resident mast cells in conjunctiva of the eyelids (Sharif et al. 1996; Yanni et al. 1999). SAC leads to decreased work productivity, increased absenteeism from work and school, limitation of everyday activities, significantly reduced quality of life, including decreased sleep quality. These SAC symptoms combined with seasonal rhinitis cause further ill health and detrimental psychological ill effects leading to impaired social interaction on top of the physical morbidity. Overall, SAC and rhinitis due to their perennial occurrence requires potent and efficacious treatment options. Accordingly, there are now several approved/launched histamine-1 (H_1) antagonists that are used in clinical management of SAC. These include emedastine (Sharif et al. 1994), olopatadine (Sharif et al. 1996), epinastine, alcaftadine, and cetirizine. However, by far the most effective agent, with a dual mechanism of action, that provides 24-h relief from SAC upon a single t.o. dose is the H_1-antagonist/mast cell

stabilizer, olopatadine (Patanol (o.1% olopatadine); Pataday (0.2% olopatadine); Pazeo (0.7% olopatadine). Elesion/Elestat [0.05% epinastine) is also quite and effective drug for the treatment of signs and symptoms of SAC. A number of other treatment options for SAC at various stages of development are allergy vaccines (MK-3641; MK-8237; SQ tree SLIT-tablet), an aldehyde scavenger (ADX-102), Syk tyrosine kinase inhibitor (PRT-2761), and an anti-IgE monoclonal antibody (xmab7195) (Gomes 2014).

Assays and Animals Models for Discovering Drugs to Treat Allergic Conjunctivitis

Once again, exploiting the knowledge of the possible disease-causing elements, relevant cell types, and animal models have been deployed in the screening for new drugs to treat seasonal and perennial allergic conjunctivitis. Thus, isolated human primary conjunctival epithelial and mast cell (Sharif et al. 1996; Yanni et al. 1997), and corneal epithelial cells (Offord et al. 1999) has been at the forefront of the cell-based assays systems. These cells have been challenged with various allergens and the release of cytokines and other mediators quantified in the presence or absence of test drugs of interest. These assays proved effective in yielding compelling data for advancing H_1-antagonists like emedastine (Sharif et al. 1994) and dual pharmacophoric drugs like olopatadine (H_1-antagonist and mast cell stabilizer) (Sharif et al. 1996; Yanni et al. 1997) into animal models of allergic conjunctivitis (Yanni et al. 1997).

Several different mammalian species have been deployed to test drugs for inflammatory and allergic conjunctivitis including guinea pigs and rodents (Groneberg et al. 2003). The most frequently and preferred models use active immunization sensitization protocols. In the guinea pig model, for instance (Yanni et al. 1997), animals are sensitized with anti-ovalbumin (OA) serum injected subconjunctivally in one eye. Twenty-four hours after passive sensitization, ovalbumin (OA) was administered either intravenously (i.v.) or topically onto the eye. The anti-allergic effect of compounds following i.v. antigen administration was determined as follows: 30 min prior to i.v. antigen challenge, the animals received 20 μl of the drug solution or saline applied topically to the eye. The animals are then challenged i.v. via the marginal ear vein or lateral tail vein with 1.0 ml of an OA:Evans Blue solution (100 pg: l mg, guinea pigs; 1 mg:2.5 mg, rats). For assessment of the allergic response following topical ocular antigen challenge, 20 μl of ovalbumin (1.0%, w/v) was administered to the sensitized eye 5 min after topical ocular application of the test drug (20 μl). During dose-response studies, the order of compound administration is randomized. Thirty minutes later, the reaction is quantitated, using a scoring of ocular allergic reactions accounting for swelling, discharge, and congestion of the conjunctiva/eye lids.

In the histamine-induced vascular permeability model using guinea pigs, animals are injected i.v. via the marginal ear vein with 1.0 ml of Evans Blue dye (1.0 mg/ml). Forty-five minutes post-dye injection, 20 μl of test compound or saline vehicle is applied topically onto one eye of each experimental animal. Thirty minutes following topical drug application, the guinea pigs are anesthetized and challenged subconjunctivally with histamine (300 ng/10 μl). The appearance of blue color on the ocular surface is then quantitated (Yanni et al., 1996).

Dry Eye Disease (DED)

Dry eye disease (DED), keratoconjunctivitis sicca, is a multifactorial disease of the tears and ocular surface that results in symptoms of discomfort, visual disturbance, and tear film instability with potential damage to the cornea and conjunctiva (Messmer 2015; Marshall and Roach 2016; Baudouin et al. 2018; Dogru et al. 2018). It is accompanied by increased osmolarity of the tear film and inflammation of the ocular surface. Sjogren's syndrome is present in ~10% of the total DED patients. Other patients tend to be women who are postmenopausal, pregnant, or who are on hormone replacement therapy or are

taking oral contraceptives. Defects of the meibomian gland, that normally produces the oily/lipid components of the tear film which limit evaporation of the aqueous fluid of the tears, are also now recognized as key contributors to the development of DED (Baudouin et al. 2018). The signs and symptoms of DED have been difficult to address over the years apart from the use of artificial tears to provide brief and temporary relief. However, there are now four drugs available to the physician to treat DED including various formulations of cyclosporine (Restasis; Ikervis; Cyclokat; all are anti-inflammatories that are calcineurin inhibitors), diquafosol tetrasodium (a P_2Y_2 receptor agonist that stimulates tear production), rebamipide (Mucosta; mucin liberator that works on conjunctival goblet cells), and lifitigrast (Xiidra; an LFA-1 antagonist) (Messmer 2015; Marshall and Roach 2016; Baudouin et al. 2018; Dogru et al. 2018).

Several other classes of drugs are currently in early-late-stage research and development that cover diverse mechanisms of action such as thymosin-beta-4 ligand (RGN-259), TRPV1 expression inhibitor (siRNA; SYL-1001), NGF receptor agonist (tavilermide; MIM-D3), ICAM1 expression inhibitor/antioxidant (visomitin; SkQ1), nicotinic receptor agonist (cytisine; OC-02), alpha-4 integrin antagonist (AXR-159), JAK3/Syk Kinase inhibitor (R932348), aldehyde scavenger (reproxalap; ADX-102), multikinase inhibitor (TOP-1630), and RAR-gamma receptor agonist (palovarotene; RG-667) (Messmer 2015; Marshall and Roach 2016). It is hoped that some of these agents will prove to be safe and effective alternative treatment options for the DED patients in the near future.

In Vitro Assays and Animal Models of Dry Eye Disease

There are a number of therapeutic intervention points in DED including the ocular surface itself (corneal and conjunctival epithelia and goblet cells), and tear and meibomian glands. Accordingly, in vitro assays have been established that incorporate cells derived from these tissues. Cultured epithelial cells and cell lines of human cornea, conjunctiva (epithelial and goblet cells), and meibomian gland have been exposed to various treatments to simulate dry eye conditions including hyperosmolarity (400–500 mOsM) (Clouzeau et al. 2012), desiccation Hovakimyan et al. 2012), and inflammation (e.g., exposure to formaldehyde; Vitoux et al. 2018) and the effects on cell viability (lysosomal integrity), cell apoptosis/death (cell membrane permeability and chromatin condensation), secreted cytokines (e.g., IL-1, IL-8), oxidative stress (reactive oxygen species and superoxide anion), and cellular hyperpolarization measured. These indices were then used to determine the potential therapeutic effects of test compounds (e.g., Hagan et al. 2018).

Animal models for assessing the impact of test compounds in dry eye conditions have been difficult to establish and correlate with the human disease. Nevertheless, there has been progress made in understanding the DED processes and some level of screening performed using mice subjected to a dry, drafty environment (20% humidity) for 5–10 days after they receive daily subcutaneous injections of scopolamine that inhibits tear secretion. The animals exhibited a desiccating-stress-activated innate immune response resulting in release of cytokines, chemokines and MMPs on the ocular surface, increased intercellular adhesion molecule-1, increased loss of conjunctival epithelial and goblet cells, and CD4 T-cell infiltration (Stern and Pflugfelder 2017). Rabbits have also been employed for dry eye models using desiccation (Gamache et al. 2002) and lacrimal gland inflammation (Negelhout et al. 2005) with limited success.

Bacterial Infection/Ocular Inflammation

Even though bacterial infection of the cornea is fairly rare due to increased public awareness and enhanced ocular hygiene, bacterial keratitis is often caused by *Staphylococcus aureus* and *Pseudomonas aeruginosa* (Willcox 2011). Infections caused by these bacteria are best treated with the fluroquinolones levofloxacin, moxifloxacin and gatifloxicin, and/or with the aminoglycoside tobramycin. Inflammation of the eyelid hair follicles often caused by bacteria (blepharitis) is a

bothersome disorder that causes itching, redness, and irritation. Common treatments include t.o. antibiotic alone, or antibiotic + a corticosteroid. However, due to development of bacterial resistance, other drugs like ciprofloxacin, ofloxacin, and levofloxacin are less prescribed but may still be useful and the only drugs available in less developed world. New broad spectrum antibiotics are eagerly awaited in order that the bacterial resistance can be overcome.

Postoperative inflammation and pain are common when patients undergo elective Lasik, cataract or photorefractive keratectomy (PRK) eye surgery. PRK is used to correct mild to moderate nearsightedness, farsightedness, and/or astigmatism and/or mild myopia. Patients often receive nonsteroidal anti-inflammatory drugs such as the prostaglandin synthase inhibitors bromfenac, nepafenac, ketorolac, and/or low-dose steroids such as loteprednol to reduce the pain and inflammation associated with the eye surgeries (Waterbury et al. 2011).

Uveitis is caused by inflammatory responses inside the eye as a result of tissue damage, bacterial/viral infection, or due to toxins (Tsirouki et al. 2018). The disease will cause various symptoms, such as decreased vision, pain, light sensitivity, and increased "floaters" in the vitreous, and it causes 10–15% of blindness in the USA. In many cases, the cause is unknown and thus is idiopathic. Anterior uveitis occurs in the front of the eye and is the most common form of uveitis, predominantly occurring in young and middle-aged people. Many cases occur in healthy people and may only affect one eye but some are associated with rheumatologic, skin, gastrointestinal, lung, and infectious diseases. Intermediate uveitis is commonly seen in young adults and is observed in the vitreous. It has been linked to several disorders including, sarcoidosis and multiple sclerosis. Posterior uveitis is the least common form of uveitis, and it occurs in the back of the eye, often involving both the retina and the choroid. It is often called choroditis or chorioretinitis (Tsirouki et al. 2018). There are many infectious and noninfectious causes to posterior uveitis. Lastly, panuveitis is a term used when all three major parts of the eye are affected by inflammation. Behcet's disease is one of the most well-known forms of panuveitis, and it greatly damages the retina. Intermediate, posterior, and panuveitis are the most severe and highly recurrent forms of uveitis. They often cause blindness if left untreated. Treatment modalities for uveitis primarily try to eliminate inflammation, alleviate pain, prevent further tissue damage, and restore any loss of vision. Treatments depend on the type of uveitis a patient displays. Some, such as using corticosteroid eye drops and injections around the eye or inside the eye, may exclusively target the eye whereas other treatments, such immunosuppressive agents taken by mouth, may be used when the disease is occurring in both eyes, particularly in the back of both eyes (Tsirouki et al. 2018). However, these steroid-based treatments adversely affect the body and can cause glaucoma. A recent development for treating posterior uveitis that overcomes the latter side effect issues associated with corticosteroids centers around ivt injection of the immunosuppressant inhibitor of mammalian target of rapamycin, sirolimus (Nguyen et al. 2018). This drug has shown significant efficacy in a number of patients suffering from noninfectious posterior uveitis (Nguyen et al. 2018).

Screening Assays and Animal Models for Ocular Infectious Diseases and Uveitis

Infectious keratitis can be caused by various bacterial strains and fungi like *fusarium solani*. Obviously bacterial strains and fungi can be grown in vitro and effects of potential antibiotics and antifungal agents determined by direct application to these infectious agents (Jett et al. 1997). Researchers have also developed ex-vivo models using rabbit and human anterior eye segments (Pinnock et al. 2017) and animal models of infectious keratitis using mice and rabbits (Zhang et al. 2017b; Zhu et al. 2017).

The study of uveitis and drugs to treat this condition are best studied using animal models of the disease involving rats (Pepple et al. 2018) and mice (Chen et al. 2015). The different methods of inducing experimental autoimmune uveitis, panuveitis, and posterior uveitis are well described and reviewed by Bansal et al. (2015).

There is also an AIRE knockout mouse model of posterior uveitis and another Tg-knockout mouse model that reflects Birdshot retinochoroidopathy (Bansal et al. 2015).

Refractive Disorders/Errors

The most common types of refractive disorders, that result from misalignment of the light focusing on the retina, include nearsightedness (myopia), farsightedness (hyperopia), astigmatism, and presbyopia. While astigmatism results from uneven surface of the cornea, presbyopia develops due to stiffness of the lens. The Latter can lead to formation of cataracts of the lens. While in general refractive disorders can be corrected with eyeglasses, contact lenses, Lasik or PRK, a recent advancement includes the ability of EV-06 (a lipoic acid synthase modulator/S-adenosylmethionine decarboxylase stimulator) to temporarily change the fluidity of the lens to overcome presbyopia. On the other hand, myopia has been far more difficult to address since it primarily affects children, although due to the increasing use of computers/tablets and handheld wireless phones and a reduction in time spent outdoors the incidence and prevalence of myopia is steadily rising. Asian children experience myopia disproportionately greater than Caucasian children. Recent projections indicate that >50% of the world's population will have myopia by 2050, and more children and young adults will be affected. Since high myopia causes vision loss due to myopic macular degeneration, myopia and it comorbidities (cataracts, retinal detachment, and glaucoma) may become the leading cause of irreversible blindness worldwide. While t.o. dosing of atropine solution/ointment appears promising for treating myopia, additional modalities are urgently needed to stem the tide of myopia development around the globe.

Animal Models of Myopia

For sake of brevity, only myopia will be dealt with here. Several animal models using chicks, rats, mice, guinea pigs, and monkeys have been developed over the years (reviewed by Schaeffel and Feldkaemper 2015). Form deprivation (suturing of eyelids or patching the eye) and lens-induced methods are the most common and effective ways to induce myopia, and several classes of drugs have been tested for their efficacy in slowing down and/or preventing myopia. By far the most effective treatment observed in numerous animal species is the use of t.o. antimuscarinic agents (Schaeffel and Feldkaemper 2015; Jiang et al. 2018) that are able to reduce axial length within the globe. A recent study demonstrated that PG FP-receptor agonist, latanoprost (30 μg peribulbarly injected daily for 4 weeks) reduced form-deprivation-induced myopia in guinea pigs by 41% (Yang et al. 2018). Interestingly, peribulbarly injected FP-receptor antagonist, AL-8810 0.5 μg/day for 4 weeks, actually induced myopia in naïve animals (Yang et al. 2018).

Conclusions

Of the many ocular disorders, glaucoma, AMD, DR, and myopia represent the major sight-threatening diseases that afflict millions of people on the planet. Even irritating ocular disorders that do not necessarily directly cause blindness such as dry eye and seasonal allergic conjunctivitis affects millions on an annual basis. Since sight is such a precious sense, much effort has been expended to find suitable treatment modalities for the various eye diseases discussed above. It will be evident from the above discourse that while majority of the drugs discovered, developed, and launched for ocular utility represented agonists or antagonist of several GPCRs, that the newer drug classes are antibodies, gene therapeutics, growth factor proteins, cellular therapies, and miniature devices. This agnostic approach to combating ocular diseases is very encouraging and it is hoped that combination products using these foundational elements will be even more productive in helping the patients who suffer from the diseases of the eye. It is also hoped that diagnostics encompassing various biomarkers and devices, and new technologies such as adaptive optics and OCT-coupled with angiography will prove helpful in our deeper understanding of the ocular disease processes and thus lead to superior preventative measures in the future.

References and Further Reading

Alexander SPH, Christopoulos A, Davenport AP et al (2017a) The concise guide to PHARMACOLOGY 2017/18: G protein-coupled receptors. Br J Pharmacol 174:S17–S129

Alexander SPH, Striessnig J, Kelly E et al (2017b) The concise guide to PHARMACOLOGY 2017/18: voltage-gated ion channels. Br J Pharmacol 174: S160–S194

Alexander SPH, Kelly E, Marrion NV et al (2017c) The concise guide to PHARMACOLOGY 2017/18: transporters. Br J Pharmacol 174:S360–S446

Alexander SPH, Fabbro D, Kelly E et al (2017d) The concise guide to PHARMACOLOGY 2017/18: enzymes. Br J Pharmacol 174:S272–S359

Al-Zamil WM, Yassin SA (2017) Recent developments in age-related macular degeneration: a review. Clin Interv Aging 12:1313–1330

Ambati J, Ambati BK, Yoo SH et al (2003) Age-related macular degeneration: etiology, pathogenesis, and therapeutic strategies. Surv Ophthalmol 48:257–293

American Academy Ophthalmology – Eye Health Statistics – American Academy of Ophthalmology. https://www.aao.org/newsroom/eye-health-statistics

Anderson OA, Finkelstein A, Shima DT (2013) A2E induces IL-1β production in retinal pigment epithelial cells via the NLRP3 inflammasome. PLoS One 8: e67263

Bansal S, Barathi VA, Iwata D et al (2015) Experimental autoimmune uveitis and other animal models of uveitis: an update. Indian J Ophthalmol 63:211–2118

Batlle JF, Fantes F, Riss I et al (2016) Three-year follow-up of a novel aqueous humor microshunt. J Glaucoma 25: e58–e65

Baudouin C, Rolando M, Benitez Del Castillo JM, et al (2018) Reconsidering the central role of mucins in dry eye and ocular surface diseases. Prog Retin Eye Res. pii: S1350-9462 (18) 30062-4

Cavet ME, DeCory HH (2018) The role of nitric oxide in the intraocular pressure lowering efficacy of latanoprostene bunod: review of nonclinical studies. J Ocul Pharmacol Ther 34:52–60

Chader GJ (2002) Animal models in research on retinal degenerations: past progress and future hope. Vis Res 42:393–399

Chen H-H, Namil A, Severns B et al (2014) In vivo optimization of 2,3-diaminopyrazine Rho Kinase inhibitors. Bioorg Med Chem Lett 24:1875–1879

Chen J, Qian H, Horai R et al (2015) Mouse models of experimental autoimmune uveitis: comparative analysis of adjuvant-induced vs spontaneous models of uveitis. Curr Mol Med 15:550–557

Chen L, Bai Y, Zhao M (2016) TLR4 inhibitor attenuates amyloid-b-induced angiogenic and inflammatory factors in ARPE-19 cells: implications for age-related macular degeneration. Mol Med Rep 13:3249–3256

Chen LJ, Ito S, Kai H et al (2017) Microfluidic co-cultures of retinal pigment epithelial cells and vascular endothelial cells to investigate choroidal angiogenesis. Sci Rep 7:3538

Clouzeau C, Godefrov D, Riancho L et al (2012) Hyperosmolarity potentiates toxic effects of benzalkonium chloride on conjunctival epithelial cells in vitro. Mol Vis 18:851–863

Cordeiro MF, Normando EM, Cardoso MJ et al (2017) Real-time imaging of single neuronal cell apoptosis in patients with glaucoma. Brain 140:1757–1767

Crider JY, Sharif NA (2001) Functional pharmacological evidence for EP_2 and EP_4 prostanoid receptors in immortalized human trabecular meshwork and N-PCE cells. J Ocul Pharmacol Ther 17:35–46

Da Cruz L, Fynes K, Georgiadis O et al (2018) Phase 1 clinical study of an embryonic stem cell-derived retinal pigment epithelium patch in age-related macular degeneration. Nat Biotechnol. https://doi.org/10.1038/nbt.4114

Di Y, Chen XL (2018) Effects of LY294002 on the function of retinal endothelial cells in vitro. Int J Ophthalmol 11:1447–1450

Dismuke WM, Sharif NA, Ellis DZ (2009) Human trabecular meshwork cell volume decrease by NO-independent soluble guanylate cyclase activators YC-1 and BAY-58-2667 involves the BK_{Ca} ion channel. Invest Ophthalmol Vis Sci 50:3353–3359

Dismuke WM, Sharif NA, Ellis DZ (2010) Endogenous regulation of human Schlemm's canal cell volume by nitric oxide signaling. Invest Ophthalmol Vis Sci 51:5817–5824

Dogru M, Kojima T, Simsek C et al (2018) Potential role of oxidative stress in ocular surface inflammation and dry eye disease. Invest Ophthalmol Vis Sci 59: DES163–DES168

Du H, Xiao X, Stiles T et al (2016) Novel mechanistic interplay between products of oxidative stress and components of the complement system in AMD pathogenesis. Open J Ophthalmol 6:43–50

Farjood F, Vargis E (2018) Novel devices for studying acute and chronic mechanical stress in retinal pigment epithelial cells. Lab Chip 18:3413–3424

Fuwa M, Toris CB, Fan S et al (2018) Effects of a novel selective EP2 receptor agonist, omidenepag isopropyl, on aqueous humor dynamics in laser-induced ocular hypertensive monkeys. J Ocul Pharmacol Ther 34:531–537

Galloway CA, Dalvi S, Hung SC et al (2017) Drusen in patient-derived hiPSC-RPE models of macular dystrophies. Proc Natl Acad Sci U S A 114:E8214–E8223

Galloway CA, Dalvi S, Shadforth AMA et al (2018) Characterization of human iPSC-RPE on a prosthetic Brusch's membrane manufactured from silk fibroin. Invest Ophthalmol Vis Sci 59:2792–2800

Gamache DA, Wei ZY, Weimer LK et al (2002) Corneal protection by the ocular mucin secretagogue 15(S)-HETE in a rabbit model of desiccation-induced corneal defect. J Ocular Pharmacol Ther. 18:349–361

Gomes PJ (2014) Trends in prevalence and treatment of ocular allergy. Curr Opin Allergy Clin Immunol 14:451–456

Griffin BW, Klimko P, Crider JY et al (1999) AL-8810: a novel $PGF_{2\alpha}$ analog with selective antagonist effects at the FP prostaglandin receptor. J Pharmacol Exp Ther 290:1278–1284

Groneberg DA, Bielory L, Fischer A et al (2003) Animal models of allergic and inflammatory conjunctivitis. Allergy 58:1101–1113

Hagan S, Fyfe MCT, Ofen-Frimpong B et al (2018) Narrow spectrum kinase inhibitors demonstrate promise for the treatment of dry eye disease and other ocular inflammatory disorders. Invest Ophthalmol Vis Sci 59:1443–1453

Hartman RR, Kompella UB (2018) Intravitreal, subretinal, and suprachoroidal injections: evolution of microneedles for drug delivery. J Ocular Pharmacol Ther 34:141–153

He S, Stankowska DL, Ellis DZ et al (2018) Targets of neuroprotection in glaucoma. J Ocul Pharmacol Ther 34:85–106

Hellberg MR, McLaughlin MA, Sharif NA et al (2002) Identification and characterization of the ocular hypotensive efficacy of travoprost, a potent and selective FP prostaglandin receptor agonist, and AL-6598, a DP prostaglandin receptor agonist. Surv Ophthalmol 47 (Suppl 1):S13–S33

Hernandez-Zimbron LF, Zamora-Alvarado R, Ochoa-Del la Paz L et al (2018) Age-related macular degeneration: new paradigms for treatment and management of AMD. Oxid Med Cell Longev: 2018:8374647

Hines-Beard J, Wesley S, Bond WS et al (2016) Virus-mediated EpoR76E gene therapy preserves vision in a glaucoma model by modulating neuroinflammation and decreasing oxidative stress. J Neuroinflammation 13:39

Hollander H, Makarov F, Stefani FH et al (1995) Evidence of constriction of optic axons at the lamina cribrosa in the normotensive eye in humans and other mammals. Ophthalmic Res 27:296–309

Holló G, Topouzis F, Fechtner RD (2014) Fixed-combination intraocular pressure-lowering therapy for glaucoma and ocular hypertension: advantages in clinical practice. Expert Opin Pharmacother 15:1737–1747

Hovakimyan M, Ramoth T, Lobler M et al (2012) Evaluation of protective effects of trehalose on desiccation of epithelial cells in three dimensional reconstructed human corneal epithelium. Curr Eye Res 37:082–089

Ishikawa M, Jin D, Sawada Y et al (2015) Future therapies of wet age-related macular degeneration. J Ophthalmol: 2015:138070

Jett BD, Hatter KL, Huycke MM et al (1997) Simplified agar plate method for quantifying viable bacteria. Biotechniques 23:648–650

Jiang L, Zhang S, Chen R et al (2018) Effects of the tyrosinase-dependent dopaminergic system on refractive error development in guinea pigs. Invest Ophthalmol Vis Sci 59:4631–4638

Jonas JB, Aung T, Bourne RR et al (2017) Glaucoma. Lancet 390:2183–2193

Kashani A, Lebkowski JS, Rahhal FM et al (2018) A bioengineered retinal pigment epithelial monolayer for advanced, dry age-related macular degeneration. Sci Transl Med 10:eaao4097, 10 pages

Katoli P, Sharif NA, Sule A et al (2010) NPR-B natriuretic peptide receptors in human corneal epithelium: mRNA, immunohistochemical, protein and biochemical pharmacology studies. Mol Vis 16:1241–1252

Kelly CR, Williams GW, Sharif NA (2003) Real-time intracellular Ca^{2+}-mobilization by travoprost acid, bimatoprost, unoprostone and other analogs via endogenous mouse, rat and cloned human FP prostaglandin receptors. J Pharmacol Exp Ther 304:238–245

Kirihara T, Taniguchi T, Yamamura K et al (2018) Pharmacologic characterization of omidenepag isopropyl, a novel selective EP2 receptor agonist, as an ocular hypotensive agent. Invest Ophthalmol Vis Sci 59:145–153

Lambert NG, Singh MK, ElShelmani H et al (2016) Risk factors and biomarkers of age-related macular degeneration. Prog Retin Eye Res 54:64–102

Liu F, Ding X, Li J et al (2016) Aqueous humor cytokine profiling in patients with wet AMD. Mol Vis 22:352–361

Liu CH, Wang Z, Ye S et al (2017) Animal models of ocular angiogenesis: from development to pathologies. FASEB J 31:4665–4681

Lin CW, Sherman B, Moore LA et al (2018) Discovery and preclinical development of netarsudil, a novel ocular hypotensive agent for the treatment of glaucoma. J Ocul Pharmacol Ther 34:40–51

Maguire MG, Martin DF, Ying G-S et al (2016) Five-year outcomes with anti-vascular endothelial growth factor treatment of neovascular age-related macular degeneration. The comparison of age-related macular degeneration treatments trials. Ophthalmology 123:1751–1761

Marquart ME (2011) Animal models of bacterial keratitis. J Biomed Biotechnol 2011:680642

Marshall LL, Roach JM (2016) Treatment of dry eye disease. Consult Pharm 31:96–106

McNally S, O'Brien CJ (2014) Metabolomics/proteomics strategies used to identify biomarkers for exfoliation glaucoma. J Glaucoma 23(8 Suppl 1):S51–S54

Mead B, Ahmed Z, Tomarev S (2018) Mesenchymal stem cell-derived small extracellular vesicles promote neuroprotection in a genetic DBA/2J mouse model of glaucoma. Invest Ophthalmol Vis Sci 59:5473–5480

Messmer EM (2015) The pathophysiology, diagnosis, and treatment of dry eye disease. Dtsch Arztebl Int 112:71–81

Negelhout TJ, Gamache DA, Roberts L et al (2005) Preservation of tear film integrity and inhibition of corneal injury by dexamethasone in a rabbit model of lacrimal gland inflammation-induced dry eye. J Ocular Pharmacol Ther 21:139–148

NEI (2014) Eye Disease Statistics Fact Sheet – National Eye Institute – NIH. https://nei.nih.gov/sites/default/.../NEI_Eye_Disease_Statistics_Factsheet_2014_V10.pdf

Nguyen QD, Merrill PT, Sepah YJ et al (2018) Intravitreal sirolimus for the treatment of noninfectious uveitis: evolution through preclinical and clinical studies. Ophthalmology 125:1984–1993

O'Brien TP (2013) Allergic conjunctivitis: an update on diagnosis and management. Curr Opin Allergy Clin Immunol 13:543–549

Ohia SE, Njie-Mbye YF, Robinson J et al (2018) Serotonin-2B/2C receptors-mediate bovine ciliary muscle contraction: role in IOP regulation. J Ocul Pharmacol Ther 34:70–75

Offord E, Sharif NA, Mace K et al (1999) Immortalized human corneal epithelial cells for ocular toxicity and inflammation studies. Invest Ophthalmol Vis Sci 40:1091–1101

Olivares AM, Althoff K, Chen GF et al (2017) Animal models of diabetic retinopathy. Curr Diab Rep 17:93. (17 pages)

Osborne NN, Álvarez CN, del Olmo Aguado S (2014) Targeting mitochondrial dysfunction as in aging and glaucoma. Drug Discov Today 19:1613–1622

Patil R, Xu S, Rusinko A, Feng Z et al (2016) Rapid identification of novel inhibitors of aquaporin-1 channel by high-throughput screening. Chem Biol Drug Des 87:794–805

Pepple KL, Wilson L, Van Gelder RN (2018) Comparison of aqueous and vitreous lymphocyte population from two rat models of experimental uveitis. Invest Ophthalmol Vis Sci 59:2504–2511

Pinnock A, Shivshetty N, Roy S et al (2017) Ex vivo rabbit and human corneas as models for bacterial and fungal keratitis. Graefes Arch Clin Exp Ophthalmol 255:333–342

Qin J, Rinella N, Zhang Q et al (2018) OCT angiography and cone photoreceptor imaging in geographic atrophy. Invest Ophthalmol Vis Sci 59:5985–5992

Quigley HA, McKinnon SJ, Zack DJ et al (2000) Retrograde axonal transport of BDNF in retinal ganglion cells is blocked by acute IOP elevation in rats. Invest Ophthalmol Vis Sci 41:3460–3466

Ramachandran C, Patil RV, Sharif NA et al (2011) Effect of elevated intracellular cAMP on actomyosin contraction in bovine trabecular meshwork cells. Invest Ophthalmol Vis Sci 52:1474–1485

Rickman CB, Farsiu S, Toth CA et al (2013) Dry age-related macular degeneration: mechanism, therapeutic targets, and imaging. Invest Ophthalmol Vis Sci 54:ORSF68–ORSF80

Schaeffel F, Feldkaemper M (2015) Animal models in myopia research. Clin Exp Optom 98:507–517

Sharif NA (2017) Ocular hypertension and glaucoma: a review and current perspectives. Int. J. ophthalmol Vis Sci. 2:22–36

Sharif NA (2018a) *i*Drugs and *i*Devices discovery and development- preclinical assays, techniques and animal model studies for ocular hypotensives and neuroprotectants. J Ocular Pharmacol Ther 34:7–39

Sharif NA (2018b) Glaucomatous optic neuropathy treatment options: the promise of novel therapeutics, techniques and tools to help preserve vision. Neural Regen Res 13:1145–1150

Sharif NA, Klimko P (2019) FP prostaglandin-receptor antagonists: discovery, pharmacological characterization and therapeutic utility. Br J Pharmacol. https://doi.org/10.1111/bph.14335

Sharif NA, May JA (2011) Potential for serotonergic agents to treat elevated intraocular pressure and glaucoma: focus on $5HT_2$ receptor agonists. Expert Rev Ophthalmol 6:105–120

Sharif NA, Xu SX (1999) Human retina contains polyamine-sensitive [^{3}H]-ifenprodil binding sites: implications for neuroprotection? Br J Pharmacol 83: 236–240.

Sharif NA, Xu S, Yanni JM (1994) Histamine receptor-subtype affinities, selectivities and potencies of emedastine, a novel H_1-selective antagonist, and other ocularly employed antihistamines. Drug Dev Res 33:448–453

Sharif NA, Xu SX, Miller ST et al (1996) Characterization of the ocular anti-allergic and anti-histaminic effects of olopatadine, a novel drug for treating ocular allergic diseases. J Pharmacol Exp Ther 278:1251–1260

Sharif NA, Kelly CR, Crider JY et al (2003) Ocular hypotensive FP prostaglandin (PG) analogs: PG receptor subtype binding affinities and selectivities, and agonist potencies at FP and other PG receptors in cultured cells. J Ocul Pharmacol Ther 19:501–515

Sharif NA, McLaughlin MA, Kelly CR (2007) AL-34662: a potent, selective, and efficacious ocular hypotensive serotonin-2 receptor agonist. J Ocular Pharmacol Ther 23:1–13

Sharif NA, Katoli P, Scott D et al (2014) FR-190997, a non-peptide bradykinin B_2-receptor partial agonist, is a potent and efficacious intraocular pressure lowering agent in ocular hypertensive cynomolgus monkeys. Drug Dev Res 5:211–223

Sharif NA, Williams GW, Davis TL (2000) Pharmacology and autoradiography of human DP prostanoid receptors using [^{3}H]-BWA868C, a DP receptor-selective antagonist radioligand. Br J Ophthalmol 131:1025–1038

Shen W, Lee S-R, Yam M et al (2018) A combination therapy targeting endoglin and VEGF-A prevents sub-retinal fibro-neovascularization caused by induced Muller cell disruption. Invest Ophthalmol Vis Sci 59:6075–6088

Simo R, Stitt AW, Gardner TW (2018) Neurodegeneration in diabetic retinopathy: does it really matter? Diabetologia 61:1902–1912

Smedowski A, Liu X, Pietrucha-Dutczak M et al (2016) Predegenerated Schwann cells–a novel prospect for cell therapy for glaucoma: neuroprotection, neuroregeneration and neuroplasticity. Sci Rep 6:23187

Stern ME, Pflugfelder SC (2017) What have we learned from animal models of dry eye. Int Ophthalmol Clin 57:109–118

Tezel G, Yang X, Luo C et al (2010) Oxidative stress and the regulation of complement activation in human glaucoma. Invest Ophthalmol Vis Sci 51:5071–5082

Tham Y-C, Li X, Wong TY et al (2014) Global prevalence of glaucoma and projections of glaucoma burden through 2040. Ophthalmology 121:2081–2090

Thomas D, Papadopoulo O, Doshi R et al (2000) Retinal ATP and phosphorus metabolites: reduction by hypoxia and recovery with MK-801 and diltiazem. Med Sci Res 28:87–91

Tien T, Zhang J, Muto T et al (2017) High glucose induces mitochondrial dysfunction in retinal Müller cells: implications for diabetic retinopathy. Invest Ophthalmol Vis Sci 58:2915–2921

Tsirouki T, Dastiridou A, Symeonidis C et al (2018) A focus on the epidemiology of uveitis. Ocul Immunol Inflamm 26:2–16

Venugopalan P, Wang Y, Nguyen T et al (2016) Transplanted neurons integrate into adult retinas and respond to light. Nat Commun 7:10472

Vitoux MA, Kessal K, Baudouin C et al (2018) Formaldehyde gas exposure increases inflammation in an in vitro model of dry eye. Toxicol Sci 1(160):108–117

Wang W, Lo ACY (2018) Diabetic retinopathy: pathophysiology and treatments. Int J Mol Sci 19:1816, 14 pages

Wang JW, Woodward DF, Stamer WD (2013) Differential effects of prostaglandin E2-sensitive receptors on contractility of human ocular cells that regulate conventional outflow. Invest Ophthalmol Vis Sci 54:4782–4790

Waterbury LD, Galindo D, Villanueva L et al (2011) Ocular penetration and anti-inflammatory activity of ketorolac 0.45% and bromfenac 0.09% against lipopolysaccharide-induced inflammation. J Ocul Pharmacol Ther 27:173–179

Waugh N, Loveman E, Colquitt J et al (2018) Treatments for dry age-related macular degeneration and Stargardt disease: a systematic review. Health Technol Assess 22:1–168

Weinreb RN, Aung T, Medeiros FA (2014) The pathophysiology and treatment of glaucoma: a review. JAMA 311:1901–1911

WHO (2018) Blindness and vision impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. WHO priority eye diseases. https://www.who.int/blindness/causes/priority/en/

Wiederholt M, Thieme H, Stumpff F (2000) The regulation of trabecular meshwork and ciliary muscle contractility. Prog Retin Eye Res 19:271–295

Willcox MDP (2011) Review of resistance of ocular isolates of Pseudomonas aeruginosa and staphylococci from keratitis to ciprofloxacin, gentamicin and cephalosporins. Clin Exp Optom 94:161–168

Williams PA, Harder JM, Foxworth NE et al (2017) Vitamin B_3 modulates mitochondrial vulnerability and prevents glaucoma in aged mice. Science 355:756–760

Xu G, Weinreb RN, Leung CK (2014) Optic nerve head deformation in glaucoma: the temporal relationship between optic nerve head surface depression and retinal nerve fiber layer thinning. Ophthalmology 121:2362–2370

Yang S, Zhao J, Sun X (2016) Resistance to anti-VEGF therapy in neovascular age-related macular degeneration: a comprehensive review. Drug Des Devel Ther 10:1857–1867

Yang J, Pan M, Reinech PS et al (2018) Prostaglandin F2α receptor modulation affects eye development in guinea pigs. Basic Clin Pharmacol Toxicol 123:263–270

Yanni JM, Miller ST, Gamache DA et al (1997) A comparison of topical ocular anti-allergy drugs: effects on human conjunctival mast cell mediator release. Ann Allergy Asthma Immunol 79:541–545

Yanni JM, Sharif NA, Gamache DA et al (1999) A current appreciation of sites for pharmacological intervention in allergic conjunctivitis: effects of new topical ocular drugs. Acta Ophthalmol Scand 77:33–37

Yucel YH, Zhang Q, Gupta N et al (2000) Loss of neurons in magnocellular and parvocellular layers of the lateral geniculate nucleus in glaucoma. Arch Ophthalmol 118:378–384

Zhang S, Yu N, Zhang R et al (2016) Interleukin-17A induces IL-1β secretion from RPE cells via NLRP3. Invest Ophthalmol Vis Sci 57:312–319

Zhang TT, Muto J, Kim D et al (2017a) High glucose induced mitochondrial dysfunction in retinal Muller cells: implications for diabetic retinopathy. Invest Ophthalmol Vis Sci 58:2915–2921

Zhang Z, Abdel-Razek O, Wang G (2017b) The mouse model for ocular surface staphylococcus aureus infection. Curr Protoc Mouse Biol 7:55–63

Zhu H, Kochevar IE, Behlau I et al (2017) Antimicrobial blue light therapy for infectious keratitis: ex vivo and in vivo studies. Invest Ophthalmol Vis Sci 58:586–593

Clinical Pharmacology of Tinnitus: Design and Evaluation

8

Agnieszka J. Szczepek

Contents

Abstract

Tinnitus – the perception of phantom sound – is estimated to seriously affect the quality of life of about 3% of the entire world population, making it an attractive target for pharmacotherapy. However, none of the so far conducted clinical trials with the use of pharmacological substances could be called a thrilling success. There are multiple reasons for this, which are discussed in this chapter. Moreover, a comprehensive overview of factors that should be taken under consideration when designing clinical pharmacological study for tinnitus is presented in an anticipation to help design trials producing meaningful clinical data and identifying clinically relevant substances effective in tinnitus treatment.

Purpose and Rationale

The purpose of this chapter was to review and to present contemporary information regarding the design of clinical trials for tinnitus. In addition, the aim was to distinguish the two main different directions that are being developed in tinnitus pharmacology, namely, the treatment of tinnitus percept and the treatment of tinnitus-related distress.

A. J. Szczepek (✉)
Department of ORL, Head and Neck Surgery, Charité University Hospital, Berlin, Germany
e-mail: agnes.szczepek@charite.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_61

Although the American Academy of Otolaryngology-Head and Neck Surgery recommended *against* the contemporary medial therapy of tinnitus (antidepressants, anticonvulsants, anxiolytics, or intratympanic medications (Tunkel et al. 2014)), the need to explore pharmacological intervention remains. The method: data from clinical studies that were performed between 2006 and 2016 and were analyzed in a recent systematic review (Hall et al. 2015, 2016) were included in the present study. Sixty-five studies that used pharmacological approach for tinnitus treatment were extracted. In these particular 65 studies, the pharmacological substances used, the outcome domains, and the outcome measure instruments were analyzed, presented, and discussed.

Tinnitus

Tinnitus Characteristics

Tinnitus is a *symptom* of various diseases that manifests itself as a sound percept without an external source (Jastreboff 1990). *Objective tinnitus* can be heard by other persons (often only with an aid of amplification) because it is caused by internal bodily noises (e.g., pulse that could be heard if the diseased blood vessel is located in the proximity of the ear). *Subjective tinnitus* is a phantom sound heard exclusively by the affected person. Illnesses capable of inducing subjective tinnitus include, but are not limited to, the head and neck injuries and all diseases that induce hearing impairment (middle ear inflammation, meningitis, otosclerosis, Meniere's disease, presbyacusis, ototoxicity, noise-induced hearing loss, vestibular schwannomas, meningioma, intracranial pressure, atherosclerosis, diabetes, and other diseases) (Baguley et al. 2013). The grade of hearing loss has been suggested to correlate with the grade of tinnitus impairment (Mazurek et al. 2010). Recently, cochlear synaptopathies that cause *hidden hearing loss* are being considered as a possible reason underlying tinnitus (Guest et al. 2017; Liberman and Kujawa 2017). Unfortunately, in many cases, the direct cause of tinnitus remains unknown making causative therapy approaches difficult.

Box 1

Tinnitus often associates with hearing loss (Henry et al. 2014), but pharmacological intervention for hearing loss has a very small therapeutic time window, as the auditory hair cells are postmitotic, and in mammals, they are unable to regenerate (Seymour and Pereira 2015). Successful attempts of therapy against noise-induced hearing loss included using steroids, magnesium, coenzyme Q10, or D-methionine (Sakat et al. 2016). Importantly, all of the clinical studies were either using protective approach (prior to noise exposure) or an intervention immediately after noise exposure. Non-pharmacological approach of treating hearing loss (hearing aids, implantable hearing aids, and cochlear implants) is effective not only in restoring the ability to hear but also in reducing tinnitus-related distress (Olze et al. 2011; Ramos Macias et al. 2015).

There are several clinical terms used in the descriptive diagnostics of tinnitus that relate to the duration or to the severity of tinnitus – the most common ones are presented in Table 1.

Regardless of the specific cause of tinnitus, the common denominator for all types of tinnitus is the *activation of auditory cortex under acoustically sterile conditions*. This activation results from the stimulation of auditory pathway that may take place in various parts of auditory circuits, starting from the periphery and ending in the auditory cortex.

Influence of Tinnitus on Nonauditory Systems

In addition to activating the auditory system, processed acoustic signals stimulate also other structures. A good example is the reaction of a central nervous system to nonverbal acoustic stimuli, such as music (Koelsch 2014). The sound of music activates not only the auditory but also the somatosensory, autonomic, vestibular, and limbic systems (Fig. 1).

Table 1 Various types of tinnitus used in clinical descriptions.

Type of tinnitus		Remarks
Constant	Intermittent	Refers to the presence of tinnitus (continuous and noncontinuous)
Distressing (decompensated)	Non-distressing (compensated)	Refers to the psychological effect of tinnitus on the affected person
Acute	Chronic	Refers to the duration of tinnitus, where tinnitus is considered acute when occurring no longer than 3 months and chronic when longer than 3 months. In some countries or societies, the duration of tinnitus is regarded as chronic when longer than 6 or even 12 months
Peripheral	Central	Refers to the anatomical place (but not the cause!) of tinnitus origin, where "peripheral tinnitus" is considered to originate from cochlea and "central tinnitus" from anywhere between the cochlear nucleus and auditory cortex
With mental comorbidities	Without mental comorbidities	Refers to comorbid mental conditions such as anxiety and depression of phobias
Unilateral	Bilateral	Refers to the affected side

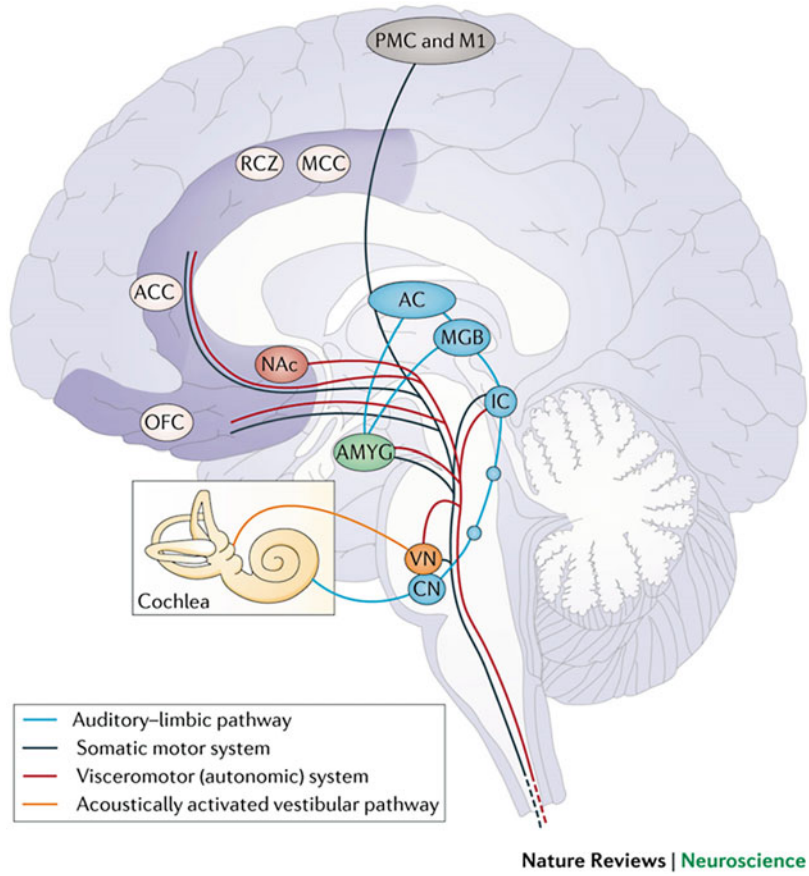


Fig. 1 The main pathways underlying autonomic and muscular responses to music. The auditory cortex (AC) also projects to the orbitofrontal cortex (OFC) and the cingulate cortex (projections not shown). Moreover, the amygdala (AMYG), OFC, and cingulate cortex send numerous projections to the hypothalamus (not shown) and thus also exert influence on the endocrine system, including the neuroendocrine motor system. *ACC*, anterior cingulate cortex; *CN*, cochlear nuclei; *IC*, inferior colliculus; *M1*, primary motor cortex; *MCC*, middle cingulate cortex; *MGB*, medial geniculate body; *NAc*, nucleus accumbens; *PMC*, premotor cortex; *RCZ*, rostral cingulate zone; *VN*, vestibular nuclei (Reprinted with permission from Springer Nature from (Koelsch 2014))

Likewise, simple auditory stimuli such as the sound of chimes or a beep were shown to stimulate several structures in the central nervous system in addition to the auditory brain (Georgiewa et al. 2016). Interestingly, neuroimaging of tinnitus patients determined an increased cortical activity in the auditory brain of tinnitus patients (Arnold et al. 1996) as well as different reactions to the sound of the non-auditory areas (Georgiewa et al. 2016) suggesting *altered neuronal connectivity in the tinnitus brain* (Leaver et al. 2016).

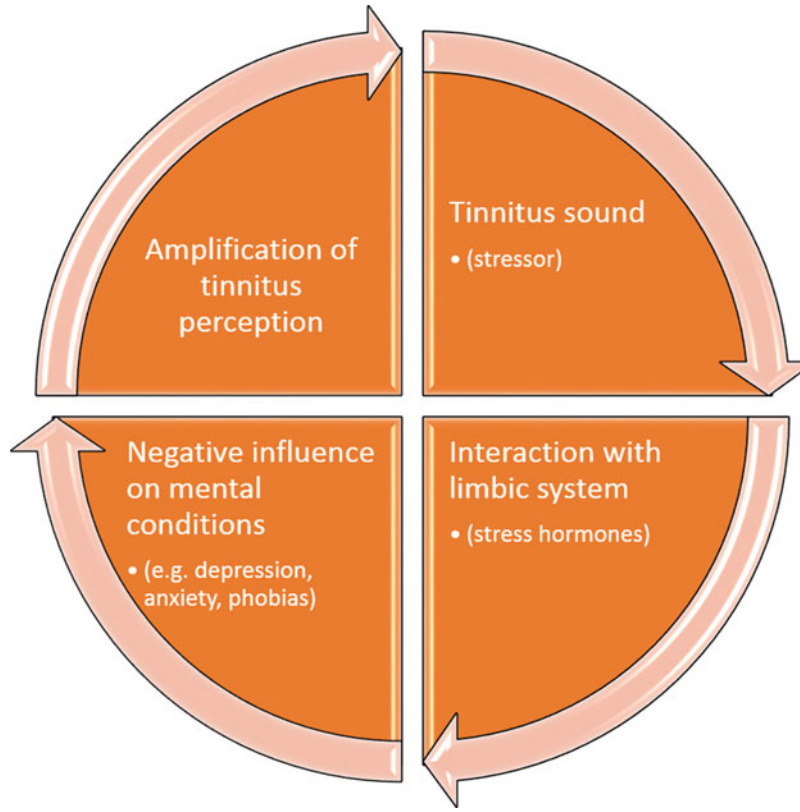


Fig. 2 Tinnitus as stressor that amplifies other mental conditions as well as its own related distress

Box 2
Acoustic signals induce auditory and non-auditory brain stimulation.

Acoustic stimulation can be viewed as an auditory signal that induces neuronal auditory and nonauditory reactions, of which the evoked emotional reactions can be of positive or negative nature. In addition, sound may occasionally induce responses from the autonomic nervous system. In people *suffering* from tinnitus, the phantom sound perceived by the auditory system comes to be *negatively labeled* in terms of emotional load. Because of this, the *sound of tinnitus becomes a stressor* for the affected person. As such, it activates hypothalamus-pituitary-adrenal (HPA) axis inducing the release of stress hormones and provoking tinnitus-specific stress reactions that are referred to as *tinnitus-induced distress*. These reactions include but are not limited to nervousness, insomnia, problems with concentrations, and other secondary and tertiary responses to stress. Some of the reactions may induce the development or an aggravation of already existing mental conditions, such as anxiety (Pattyn et al. 2016) or depressive symptoms (Hoare et al. 2011). In turn, these conditions may worsen tinnitus percept and tinnitus-related distress (Fig. 2).

Multidisciplinary Aspects of Tinnitus

Tinnitus is a sensation of a sound and as such it often takes the tinnitus sufferers to the office of audiologist. However, the profession of audiologist as a health-care professional is known only in some countries (e.g., the United Kingdom, Sweden, the USA, Canada, Australia, Malaysia, India, or Portugal), whereas in other countries

(e.g., Germany, Austria, Poland, France, or Czech Republic), people experiencing tinnitus or other hearing problems are examined by the ORL specialists and then optionally referred to other medical professionals. Persons suffering from a long-term (chronic) tinnitus, who in addition to perceiving a phantom sound react to it in a negative emotional way, are often referred to clinical psychologists. In addition, patients with bothersome tinnitus who have comorbid metal conditions may be attended to by psychiatrist, while patients with so-called somatosensoric tinnitus (Haider et al. 2017) will be attended to also by a physiotherapist. As a result, numerous health professionals deal with tinnitus patients: general practitioners, ORL specialists, audiologists, psychologists, psychiatrists, cardiologists, neurologists, physical therapists, and dentists (Tunkel et al. 2014). Because of distinct education and partitioned competences, although focused on tinnitus treatment, these specialists will have different clinical expertise on the subject. During the design of clinical study, one needs to take this under consideration when involving health practitioners.

Contemporary Studies Involving Pharmacological Interventions for Tinnitus

Recent systematic review that analyzed outcome domains and instruments used to measure outcome of clinical trials for tinnitus has identified 228 trials meeting the review criteria and performed between the years 2006 and 2016 (Hall et al. 2016). Of these trials, 65 involved pharmacological agents. The drugs could be generally split into two categories (for the precise listing, see Table 2):

- **Targeting the auditory pathway** (substances aiming at the auditory pathway glutamate receptors, substances blocking sodium or potassium pumps)
- **Targeting the tinnitus-related distress and/or comorbid disorders** (substances aiming at serotonin uptake, dopamine receptors, antagonists at the μ-opioid receptor, etc.)

Only about a 40% of the drugs used in tinnitus-related clinical trials aim at the tinnitus percept via targeting the auditory pathway, while the remaining 60% are directed against the comorbid diseases and/or tinnitus-related distress (Fig. 3).

The Design of Pharmacological Intervention for Tinnitus

The design of clinical intervention for tinnitus depends on several factors, such as sample homogeneity, choice of pharmacological target, or effect measured (Table 3). Sample homogeneity could be achieved by choosing one subtype of tinnitus in an age- and gender-matched group of patients. However, despite the attempts to standardize tinnitus diagnostic and classification procedure (Crummer and Hassan 2004; Langguth et al. 2011), there is still lack of internationally acknowledged and scientifically and clinically verified tinnitus subtypes.

Pharmacological trials for tinnitus use standard trial design (parallel, crossover, blinding, etc.). However, several crucial factors need to be taken under consideration when designing tinnitus trial:

1. That tinnitus is a subjective symptom
2. That the acoustic properties of tinnitus are measured with subjective methods
3. That perceiving tinnitus must not mean suffering from tinnitus
4. That all the accepted means to measure the degree of tinnitus-induced distress (or degree of suffer) are subjective
5. That targeting the disease, which presumably caused tinnitus, must not necessarily target tinnitus itself

Three further issues are of vast importance when designing the pharmacological trial for tinnitus:

Table 2 Drugs used in the analyzed pharmacological trials in the period between 2006 and 2016[a]

Drug	Number of trials	Mode of action
Acamprosate	1	The mechanism of action of acamprosate is unknown and controversial. Targets NMDA
Alprazolam	2	A potent, short-acting anxiolytic of the benzodiazepine class – a minor tranquilizer
AM-101	6	Esketamine hydrochloride, an N-methyl-D-aspartate (NMDA) receptor antagonist
Atorvastatin	1	Atorvastatin works by inhibiting HMG-CoA reductase, an enzyme found in the liver tissue that plays a key role in the production of cholesterol in the body
AUT00063	1	Small-molecule modulator of Kv3 potassium channels
Betahistine dihydrochloride	2	Betahistine has a very strong affinity as an antagonist for histamine H_3 receptors and a weak affinity as an agonist for histamine H_1 receptors
BGG492A	1	Competitive antagonist of the AMPA and kainate receptors
Cannabis	1	Cannabinoid is one of a class of diverse chemical compounds that acts on cannabinoid receptors in cells that alter neurotransmitter release in the brain
Carbamazepine	1	Carbamazepine is a blocker of voltage-gated sodium channels that binds to activated voltage-gated sodium channels, preventing repetitive and sustained firing of an action potential
Caroverine	2	Acts as an N-type calcium channel blocker, competitive AMPA receptor antagonist, and noncompetitive NMDA receptor antagonist (Arnold et al. 1996). It also has potent antioxidant effects (Baguley et al. 2013)
Cilostazol	1	Cilostazol is a phosphodiesterase inhibitor with therapeutic focus on cyclic adenosine monophosphate (cAMP)
Cinnarizine	1	Cinnarizine is an antihistamine and a calcium channel blocker; it is also known to promote cerebral blood flow
Cyclobenzaprine	1	Cyclobenzaprine is a muscle relaxer medication used to relieve skeletal muscle spasms and associated pain in acute musculoskeletal conditions
D-cycloserine	1	Is an antibiotic used to treat tuberculosis and target the glycine-binding site of N-methyl-D-aspartate (NMDA) receptors in humans
Deanxit	1	Deanxit is made up of two components: **flupentixol** 0.5 mg (or flupenthixol, an antipsychotic) and **melitracen** 10 mg (a tricyclic antidepressant)
Dexamethasone	1	Corticosteroid medication
Escitalopram	1	An antidepressant of the selective serotonin reuptake inhibitor (SSRI) class
Fluoxetine	1	An antidepressant of the selective serotonin reuptake inhibitor (SSRI) class
Fluvoxamine	1	Selective serotonin reuptake inhibitor (SSRI) and σ_1 receptor agonist
Gabapentin	5	Mimics the chemical structure of the neurotransmitter gamma-aminobutyric acid (GABA)
Ginkgo biloba	3	A possible treatment for dementia and Alzheimer's disease, possibly improving cerebral circulation
Hangekobokuto	2	A lignan isolated from the bark, seed cones, and leaves of trees belonging to the genus *Magnolia*. Used as analgesic and to treat anxiety and mood disorders

(continued)

Table 2 (continued)

Drug	Number of trials	Mode of action
Lidocaine	2	Lidocaine alters signal conduction in neurons by blocking the fast voltage-gated Na+ channels in the neuronal cell membrane responsible for signal propagation
Lyophilized powder of enzymolyzed honeybee larvae	1	Unknown
Magnesium	1	NMDA antagonist
Melatonin	4	N-Acetyl-5-methoxy tryptamine: a hormone that is produced by the pineal gland in animals and regulates sleep and wakefulness
Memantine	1	NMDA antagonist
Naltrexone	1	Naltrexone and its active metabolite 6β-naltrexol are antagonists at the μ-opioid receptor
Neramexane	6	A drug related to memantine (Arnold et al. 1996), which acts as an NMDA antagonist (Baguley et al. 2013) and has neuroprotective effects
Paroxetine	1	An antidepressant of the selective serotonin reuptake inhibitor (SSRI) class
Piribedil	1	D_2 and D_3 receptor agonist
Pramipexole	1	Dopamine agonist
Prednisolone	1	Steroid medication
Q10	2	Antioxidant, part of mitochondrial respiratory chain
Simvastatin	1	Simvastatin inhibits 3-hydroxy-3-methylglutaryl (HMG) coenzyme A reductase
Trazodone	1	Serotonin antagonist and reuptake inhibitor (SARI) class
Vardenafil	1	PDE5 inhibitor
Vestipitant	2	Selective antagonist for the NK1 receptor (substance P receptor)
Zinc	1	NMDA antagonist

[a]Based on the supplementary data from (Hall et al. 2016)

Target Selection

General target selection discriminates between tinnitus percept and tinnitus-induced distress. One-third of the previous studies targeted tinnitus-induced distress, and roughly equal number targeted tinnitus percept (Fig. 4). The rest of the studies either have not officially stated their target or targeted other domains, such as quality of life or comorbid symptoms. In addition, the molecular identity of the target needs to be specified. Few genetic studies that were performed to identify possible candidate genes associated with tinnitus rather than doing just that pointed the need of better phenotyping or subtyping of tinnitus (Vona et al. 2017).

Sample Selection (Inclusion and Exclusion Criteria and Sample Size)

Stringent inclusion and exclusion criteria supported by up-to-date diagnostic criteria should be set to assure sample homogeneity. The sample should be composed based on gender distribution, age, duration of tinnitus (in the analyzed trials, duration of tinnitus ranged from less than 3 months up to a year; 19 trials either did not report or reported ambiguously the tinnitus duration (Hall et al. 2016)), possible cause of tinnitus, the degree of tinnitus-induced distress (39 of 65 trials in the past did not report this (Hall et al. 2015)), and the degree of hearing loss. In addition, comorbid mental conditions

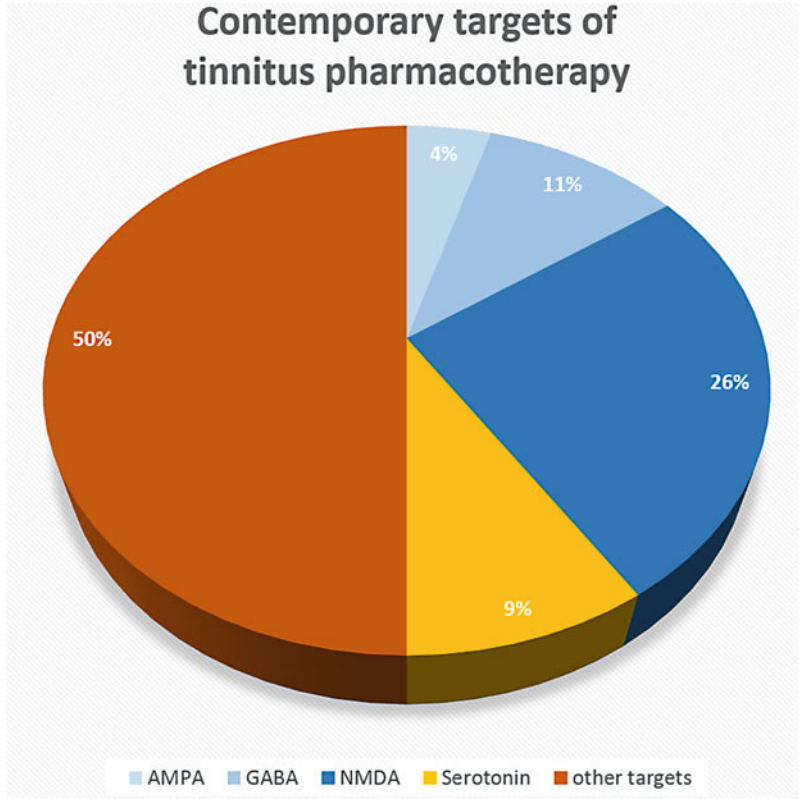


Fig. 3 Targets of pharmacotherapy (2006–2016)

Table 3 Factors particularly important during the design of clinical trial for tinnitus

Sample homogeneity	Age
	Gender
	Duration of tinnitus
	Presence of tinnitus
	Cause of tinnitus (known/unknown)
	In case of known cause – central or peripheral origin
	Comorbid conditions
	Equal degree of psychological effect of tinnitus on the affected individual
Choice of pharmacological target	The cause of tinnitus
	Tinnitus-related distress
	Comorbid conditions
Observed effect	Choice of outcome measures (domains measured, tools used for measurement)
	Clinical significance of the measured changes
	Time course of the treatment
	Time course of the follow-up

and other conditions must be taken under account. The method for calculation of sample size should be stated (in the trials analyzed, the sample size varied from 10 to 821 subjects, and the method of calculation was stated only in 9 of 65 trials (Hall et al. 2016)).

Choice of Methods to Measure the Trial Outcome

In the pharmacological tinnitus trials analyzed in the past 10 years (2006–2016), various outcome measures were used (Hall et al. 2016) and included psychometric questionnaires, numerical scales, and audiometric measurements (Fig. 5). The psychometric questionnaires were predominantly used when targeting the tinnitus-induced distress, whereas audiometric methods and numerical scales were used when the primary target consisted of tinnitus percept (Meikle et al. 2007).

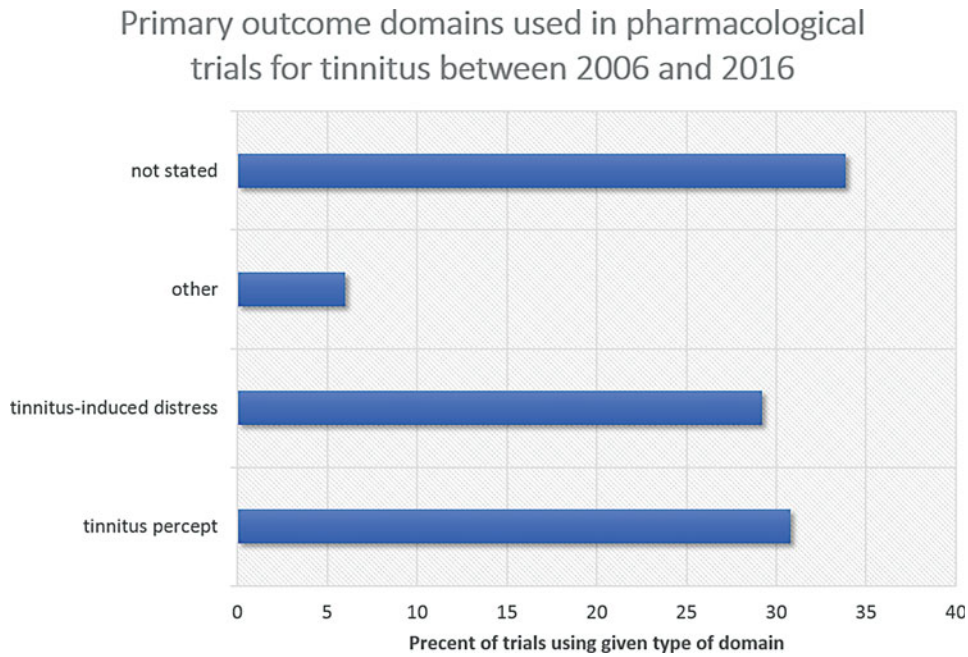


Fig. 4 Primary outcome domains used in the past trials (2006–2016)-based on supplementary data from (Hall et al. 2016).

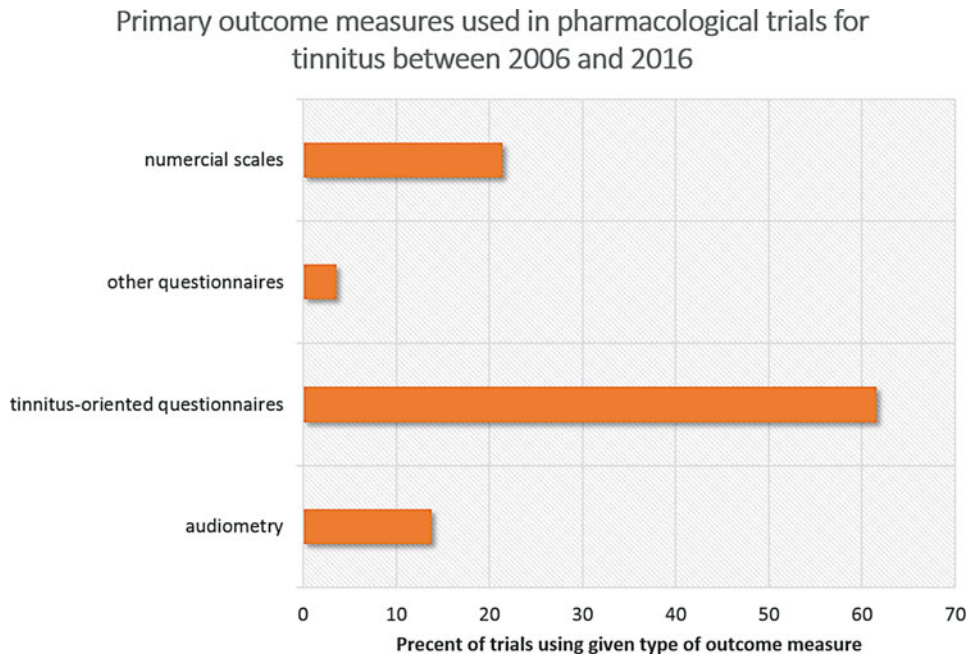


Fig. 5 Primary outcome measures used in the past trials (2006–2016)-based on supplementary data from (Hall et al. 2016).

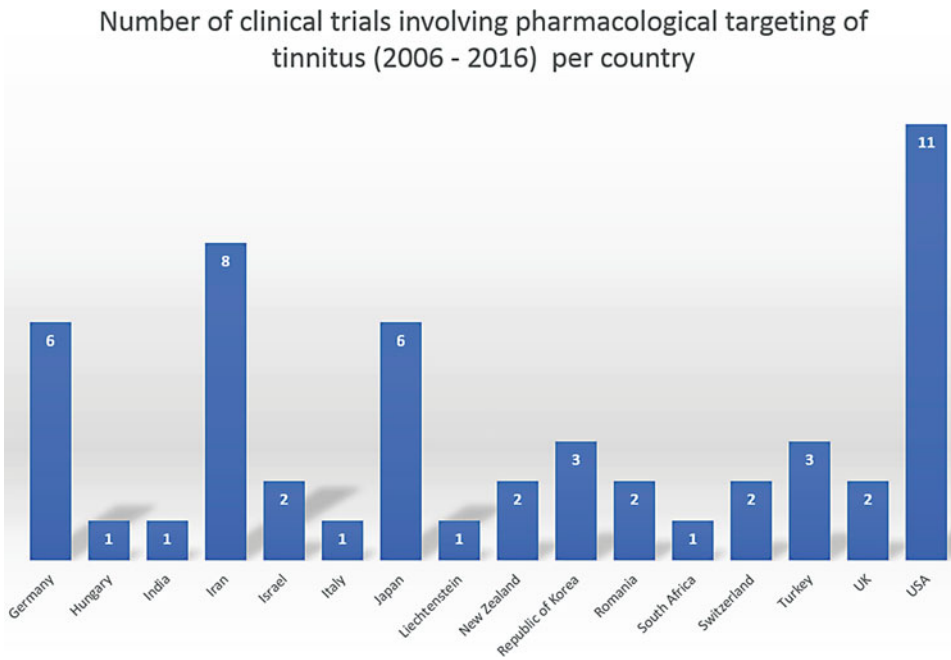


Fig. 6 Countries in which the clinical trials were conducted in the past (2006–2016)-based on supplementary data from (Hall et al. 2016).

There are several psychometric instruments that measure tinnitus-related distress. However, one needs caution when choosing the instrument, as they are not identical and often measure different domains with various sensitivities. In addition, the majority of tinnitus-related questionnaires are available in English and not in other languages. The most used tinnitus questionnaires worldwide in the past 10 years (Hall et al. 2016) include English version of Tinnitus Questionnaire (Hallam et al. 1988) and its German version (Hiller and Goebel 1992), Tinnitus Handicap Inventory (Newman et al. 1996) and Tinnitus Functional Index (Meikle et al. 2012). In addition, new questionnaires are being developed to address emerging issues, such as acceptance of tinnitus (Weise et al. 2013) To date, many questionnaires were validated and translated into other languages, e.g., Tinnitus Functional Index is presently available in German (Bruggemann et al. 2017), Swedish (Hoff and Kahari 2017), Polish (Wrzosek et al. 2016), and Dutch (Rabau et al. 2014), but it still remains to be offered in several other tongues, especially considering where the clinical pharmacological trials are being conducted (Fig. 6).

The study design in the past was predominantly randomized controlled (Fig. 6) (Hall et al. 2016), which is a general trend in the clinical research (Fig. 7).

Although clinical trials for tinnitus that were conducted in the past have not delivered a breakthrough in medical research (Beebe Palumbo et al. 2015; Plein et al. 2016; Savage and Waddell 2012), they delivered a lot of information that can be used to design an improved and well-focused trial, in which a variety of tinnitus phenotypes would be recognized (Fig. 8).

The therapeutic avenues that have in the past been explored for tinnitus include acupuncture, electromagnetic stimulation, hearing aids, hypnosis, psychotherapy, tinnitus masking devices, cognitive behavioral therapy, and tinnitus retraining therapy (Savage and Waddell 2014). For the design of the future trials, it should not be excluded that a drug therapy could be combined with one or more of the above approaches. In fact, the effectiveness of cognitive behavioral therapy

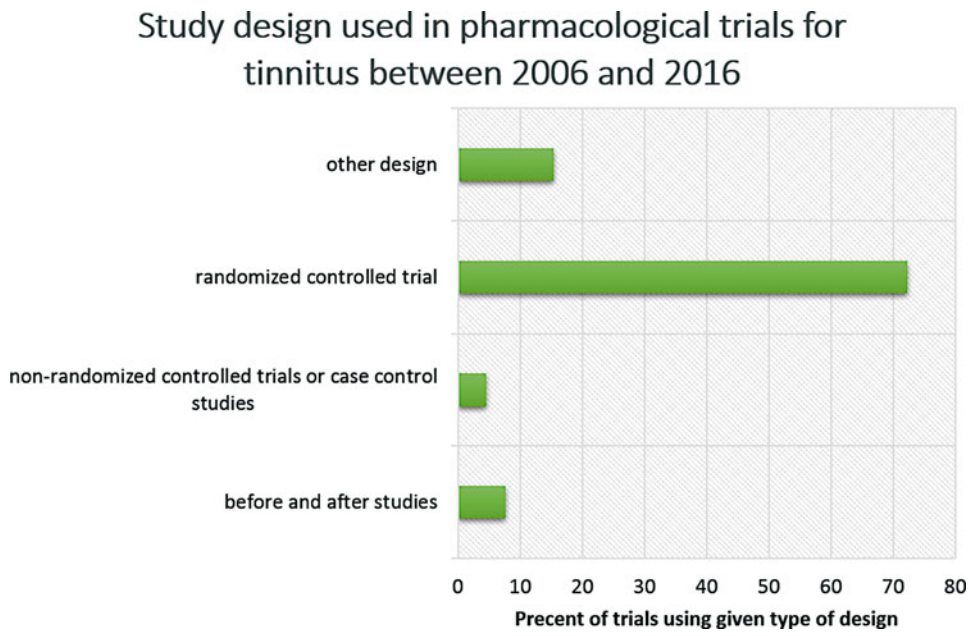


Fig. 7 Types of study design used in the past trials (2006–2016)-based on supplementary data from (Hall et al. 2016).

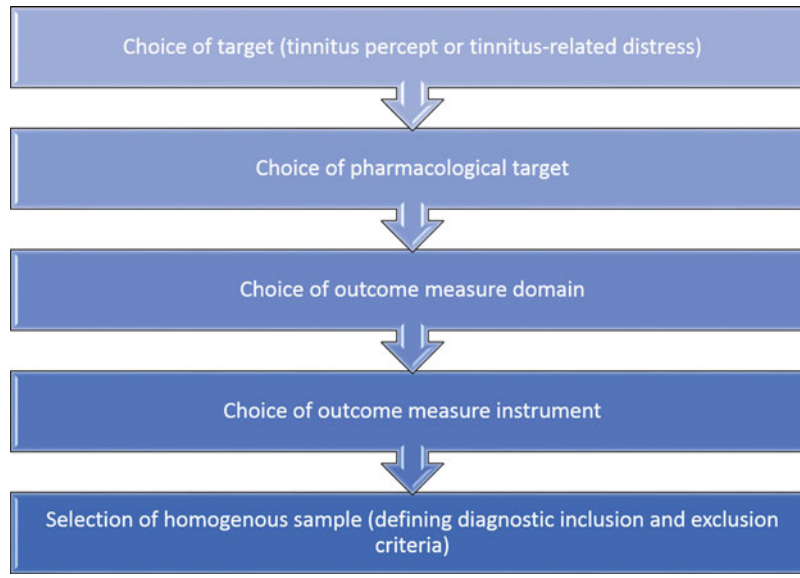


Fig. 8 General design scheme of pharmacological trial for tinnitus

in tinnitus management indicated by Cochran (Martinez-Devesa et al. 2010) and other systematic reviews (Hesser et al. 2011) could be a starting point for such combination therapy.

Taken together, the pharmacological trails were in the past anticipated to pinpoint a substance that would universally cure millions of people suffering from tinnitus. However, rather than providing a quick and uncomplicated solution, pharmacological trials uncovered the enormous diversity among tinnitus sufferers and a consequent need for rigorous tinnitus classification. Introduction of tinnitus taxonomy would improve the choice of inclusion and exclusion

criteria. This should also have positive impact on the selection of outcome domains and their measurement. Discovery or design of a universal medication against tinnitus, which would be comparable with a pain killer reducing discomfort of a headache as well as a stomachache or a toothache, is a Holy Grail of the clinical tinnitus research. However, improving the design of clinical pharmacological trials for tinnitus may result in obtaining partial answers and putting together step by step the 1000 pieces tinnitus puzzle.

References and Further Reading

Arnold W, Bartenstein P, Oestreicher E et al (1996) Focal metabolic activation in the predominant left auditory cortex in patients suffering from tinnitus: a PET study with [18F]deoxyglucose. ORL J Otorhinolaryngol Relat Spec 58:195–199

Baguley D, Mcferran D, Hall D (2013) Tinnitus. Lancet 382:1600–1607

Beebe Palumbo D, Joos K, De Ridder D et al (2015) The management and outcomes of pharmacological treatments for tinnitus. Curr Neuropharmacol 13:692–700

Brüggemann P, Szczepek AJ, Kleinjung T, Ojo M, Mazurek B (2017) [Validation of the German Version of Tinnitus Functional Index (TFI)]. Laryngorhinootologie 96(9):615–619

Crummer RW, Hassan GA (2004) Diagnostic approach to tinnitus. Am Fam Physician 69:120–126

Georgiewa P, Szczepek AJ, Rose M et al (2016) Cerebral processing of emotionally loaded acoustic signals by tinnitus patients. Audiol Neurootol 21:80–87

Guest H, Munro KJ, Prendergast G et al (2017) Tinnitus with a normal audiogram: relation to noise exposure but no evidence for cochlear synaptopathy. Hear Res 344:265–274

Haider HF, Hoare DJ, Costa RFP et al (2017) Pathophysiology, diagnosis and treatment of somatosensory tinnitus: a scoping review. Front Neurosci 11:207

Hall DA, Szczepek AJ, Kennedy V et al (2015) Current-reported outcome domains in studies of adults with a focus on the treatment of tinnitus: protocol for a systematic review. BMJ Open 5:e009091

Hall DA, Haider H, Szczepek AJ et al (2016) Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. Trials 17:270

Hallam RS, Jakes SC, Hinchcliffe R (1988) Cognitive variables in tinnitus annoyance. Br J Clin Psychol 27 (Pt 3):213–222

Henry JA, Roberts LE, Caspary DM et al (2014) Underlying mechanisms of tinnitus: review and clinical implications. J Am Acad Audiol 25:5–22. quiz 126

Hesser H, Weise C, Westin VZ et al (2011) A systematic review and meta-analysis of randomized controlled trials of cognitive-behavioral therapy for tinnitus distress. Clin Psychol Rev 31:545–553

Hiller W, Goebel G (1992) A psychometric study of complaints in chronic tinnitus. J Psychosom Res 36:337–348

Hoare DJ, Kowalkowski VL, Kang S et al (2011) Systematic review and meta-analyses of randomized controlled trials examining tinnitus management. Laryngoscope 121:1555–1564

Hoff M, Kahari K (2017) A Swedish cross-cultural adaptation and validation of the Tinnitus Functional Index. Int J Audiol 56:277–285

Jastreboff PJ (1990) Phantom auditory perception (tinnitus): mechanisms of generation and perception. Neurosci Res 8:221–254

Koelsch S (2014) Brain correlates of music-evoked emotions. Nat Rev Neurosci 15:170–180

Langguth B, Biesinger E, Del Bo L et al (2011) Algorithm for the diagnostic and therapeutic management of tinnitus. In: Møller AR, Langguth B, De Ridder D, Kleinjung T (eds) Textbook of tinnitus. Springer, New York, pp 381–385

Leaver AM, Seydell-Greenwald A, Rauschecker JP (2016) Auditory-limbic interactions in chronic tinnitus: challenges for neuroimaging research. Hear Res 334:49–57

Liberman MC, Kujawa SG (2017) Cochlear synaptopathy in acquired sensorineural hearing loss: manifestations and mechanisms. Hear Res 349:138–147

Martinez-Devesa P, Perera R, Theodoulou M et al (2010) Cognitive behavioural therapy for tinnitus. Cochrane Database Syst Rev:CD005233

Mazurek B, Olze H, Haupt H et al (2010) The more the worse: the grade of noise-induced hearing loss associates with the severity of tinnitus. Int J Environ Res Public Health 7:3071–3079

Meikle MB, Stewart BJ, Griest SE et al (2007) Assessment of tinnitus: measurement of treatment outcomes. Prog Brain Res 166:511–521

Meikle MB, Henry JA, Griest SE et al (2012) The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear Hear 33:153–176

Newman CW, Jacobson GP, Spitzer JB (1996) Development of the tinnitus handicap inventory. Arch Otolaryngol Head Neck Surg 122:143–148

Olze H, Szczepek AJ, Haupt H et al (2011) Cochlear implantation has a positive influence on quality of life, tinnitus, and psychological comorbidity. Laryngoscope 121:2220–2227

Pattyn T, Van Den Eede F, Vanneste S et al (2016) Tinnitus and anxiety disorders: a review. Hear Res 333:255–265

Plein CT, Harounian J, Floyd E et al (2016) A systematic review of eligibility and outcomes in tinnitus trials: reassessment of tinnitus guideline. Otolaryngol Head Neck Surg 154:24–32

Rabau S, Wouters K, Van De Heyning P (2014) Validation and translation of the Dutch tinnitus functional index. B-ENT 10:251–258

Ramos Macias A, Falcon Gonzalez JC, Manrique M et al (2015) Cochlear implants as a treatment option for unilateral hearing loss, severe tinnitus and hyperacusis. Audiol Neurootol 20(Suppl 1):60–66

Sakat MS, Kilic K, Bercin S (2016) Pharmacological agents used for treatment and prevention in noise-induced hearing loss. Eur Arch Otorhinolaryngol 273:4089–4101

Savage J, Waddell A (2012) BMJ Clin Evid. (0506)

Savage J, Waddell A (2014) BMJ Clin Evid. (0506)

Seymour ML, Pereira FA (2015) Survival of auditory hair cells. Cell Tissue Res 361:59–63

Tunkel DE, Bauer CA, Sun GH et al (2014) Clinical practice guideline: tinnitus executive summary. Otolaryngol Head Neck Surg 151:533–541

Vona B, Nanda I, Shehata-Dieler W et al (2017) Genetics of tinnitus: still in its infancy. Front Neurosci 11:236

Weise C, Kleinstauber M, Hesser H et al (2013) Acceptance of tinnitus: validation of the tinnitus acceptance questionnaire. Cogn Behav Ther 42:100–115

Wrzosek M, Szymiec E, Klemens W et al (2016) Polish translation and validation of the tinnitus handicap inventory and the tinnitus functional index. Front Psychol 7:1871

Clinical Aspects in Sleep Disorders and Apnea

9

Thomas Penzel and Ingo Fietze

Contents

Abstract

Sleep disorders are frequently reported complaints. Insomnia and hypersomnolence are symptoms often reported by patients and study participants. Sleep disorders with clinical consequences are not as common as complaints and symptoms might suggest. Sleep medicine is a new discipline which has developed its own curricula and physician specialization. Sleep medicine has developed a classification of sleep disorders with a manual with definitions and severity criteria. This classification will become part of the ICD-11 currently developed. The classification defines insomnia, sleep-related breathing disorders, central disorders of hypersomnolence, circadian rhythm sleep-wake disorders, parasomnias, sleep-related movement disorders, and other sleep disorders.

Diagnostic procedures include validated questionnaires; daytime testing of alertness and sleepiness; home recording of sleep-wake behavior, activity, and physiological signals; and finally a sleep laboratory investigation, cardiorespiratory polysomnography, with all signals recorded which change during normal and pathological sleep. Quantitative assessment of sleep, sleep stages, arousals from sleep, and vegetative functions during sleep is

T. Penzel (✉) · I. Fietze
Interdisciplinary Sleep Medicine Center, Charitécentrum für Herz- Kresilauf- und Gefäßmedizin CCM11, Charité – Universitätsmedizin Berlin, Berlin, Germany
e-mail: thomas.penzel@charite.de; ingo.fietze@charite.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_41

well-established, and normative values including age as modifier are well described.

Sleep disorders are recognized as risk factors for many other medical and mental disorders. Sleep disorders impair performance and may be perceived as early aging. Untreated sleep disorders cause costs at all levels of health care and need to be recognized and treated as appropriate.

Sleep disorders are a target to clinical pharmacology by being recognized and potentially excluded in any pharmacological trial. And sleep disorders are subject to drug discovery and development.

Purpose and Rationale

Sleep disorders have a high prevalence in the population. Prevalence is reported to be between 10% and 30% in the general population (Ohayon 2011). According to a health survey by the Robert-Koch Institute in 1998, about 14% of the German male population and 27% of the German female population complain about frequent or moderate insomnia (suffering from not sleeping) (Penzel et al. 2005). The survey did not check for sleep disorders according to medical definitions nor did the survey use validated and approved. Only three questions could be related to sleep disorders. These were "complaints about insomnia," "need too much sleep," and "being tired." Questions were rather unspecific. This had been recognized by many sleep researchers, and they initiated their own large sample surveys on sleep disorders (Ohayon and Zulley 2001). Assessment was performed on sleep dissatisfaction and sleep duration and more specific complaints on sleep problems. These extensive computer-driven interviews had a mean duration of 48 min per person. Definitely this is unpractical for general health surveys. A follow-up health survey by the Robert-Koch Institute between 2008 and 2011 contained much more specific questions compared to the first one in 1998 (Schlack et al. 2013). Some new questions were derived from the Pittsburgh Sleep Quality Index (Buysse et al. 1989), a well-validated questionnaire used in sleep medicine. The survey assessed the same complaints as previously and in addition more specific questions on sleep duration, problems in initiating and maintaining sleep, frequency of these complaints, perceived quality of sleep, and use of sleeping medication, all during the past 4 weeks. Based on the answers, the authors were able to estimate a prevalence of 5.7% for insomnia in Germany. This survey yielded the most recent and reliable prevalence data for Germany.

However in the German health insurance system, sleep disorders play a very small role. Whereas depression, the third most common diagnosis, is responsible for 5.6% of all days on sickness leave, sleep disorders are only listed for 0.26% of days for sickness leave (DAK report 2017). Still days for sickness leave doubled over the last decade. This reflects very well an important fact. Sleep disorders remain not to be a reason for consulting a physician in the first line. Sleep disorders remain not to be a reason for requesting for sick leave. Many patients do not report sleep complaints in the first line but often as a secondary symptom. Still sleep complaints, being difficulties in initiating and maintaining sleep or in suffering from non-refreshing sleep, are very common. Therefore it is important to investigate sleep complains and sleep problems in detail.

The first assessment is, both in patients and in presumably healthy subjects, whether they suffer from sleep problems as secondary problems or from a genuine sleep disorder. The field of sleep medicine has developed over the last three decades and has defined its own classification of sleep disorders since 1979. The latest version of the classification is called the International Classification of Sleep Disorders (ICSD-3) version 3 (AASM 2014). These are listed in the following Table 1.

One major change from the second edition of ICSD to this third edition is that the many different insomnia subtypes as carefully differentiated in the second edition had now been pooled together to only three definitions. They are now called a disorder. Specifically these are chronic insomnia disorders, short-term insomnia, and other insomnia. A chronic insomnia disorder requires a duration of symptoms for at least

Table 1 Categories of the International Classification of Sleep Disorders (AASM 2014) together with their subgroups and number of disorder definitions

Group	Subgroups	Number of definitions
Insomnia	Isolated symptoms and normal variants	5
Sleep-related breathing disorders	Obstructive sleep apnea disorders Central sleep apnea syndromes Sleep-related hypoventilation disorders Sleep-related hypoxemia disorder Isolated symptoms and normal variants	19
Central disorders of hypersomnolence	Isolated symptoms and normal variants	9
Circadian rhythm sleep-wake disorder	–	7
Parasomnias	NREM-related parasomnias REM-related parasomnias Other parasomnias Isolated symptoms and normal variants	15
Sleep-related movement disorders	Isolated symptoms and normal variants	13
Other sleep disorder	–	1
Appendix A: Sleep-related medical and neurological disorders	–	6

3 months. Previously there was a distinction between primary and secondary insomnias. Secondary insomnias were related to a primary psychiatric, medical, or substance abuse disorder. Moreover, the primary insomnias were distinguished in several more subtypes. However it turned out that symptoms and consequences do not allow to distinguish all these subtypes clearly and consistently. The differentiation was difficult, if not impossible to achieve. As a consequence they were pooled together as evidence suggests that untreated insomnia may result in adverse comorbid conditions. Another major change is that many other secondary sleep disorders were taken out of the classification. Only a few very prominent secondary sleep disorders remained in Appendix A as sleep-related medical and neurological disorders. Another change was that all different environmental-induced sleep problems were now forced to be diagnosed as either a full-blown disorder of the other definitions including all defined criteria or the problems remain as "other sleep disorder." With this, the third edition became more focused and much more condense with clear definitions and severity criteria. A list of all diagnoses defined is presented in Table 2.

With this table it becomes clear that the codes being used comes from different sections from ICD 10. Most codes came from G47 and F51 sections. Some other diagnoses were distributed across several other chapters. The new classification, as presented in the ICSD third edition (AASM 2014) is now well based on pathophysiology and is now mature enough to be stable. Finally it will be incorporated in the new ICD 11, which is currently developed, as a chapter of its own, entitled "Sleep-Wake Disorders." The proposal for the new chapter for ICD 11 follows the structure as described in Table 2 and in the ICSD third edition.

Sleep-Related Breathing Disorders

The group of disorders which attracts most attention from the health-care system is sleep-related breathing disorders. Within all types of sleep-related breathing disorders, obstructive sleep apnea

Table 2 List of all sleep disorders as they are defined and specified with their ICD 10 code. Diagnoses names and codes were taken from AASM coding manual 2014

Group of sleep disorder	Diagnosis	ICD 10 code
Insomnia	Chronic insomnia disorder	F51.01
	Short-term insomnia disorder	F51.02
	Other insomnia disorder	F51.09
	Excessive time in bed	–
	Short sleeper	–
Sleep-related breathing disorders	Obstructive sleep apnea, adult	G47.33
	Obstructive sleep apnea, pediatric	G47.33
	Central sleep apnea with Cheyne-stokes breathing	R06.3
	Central sleep apnea due to a medical disorder without Cheyne-stokes breathing	G47.37
	Central sleep apnea due to high-altitude periodic breathing	G47.32
	Central sleep apnea due to a medication or substance	G47.39
	Primary central sleep apnea	G47.31
	Primary central sleep apnea of infancy	P28.3
	Primary central sleep apnea of prematurity	P28.4
	Treatment-emergent central sleep apnea	G47.39
	Obesity hypoventilation syndrome	E66.2
	Congenital central alveolar hypoventilation syndrome	G47.35
	Late-onset central hypoventilation with hypothalamic dysfunction	G47.36
	Idiopathic central alveolar hypoventilation	G47.34
	Sleep-related hypoventilation due to a medication or substance	G47.36
	Sleep-related hypoventilation due to a medical disorder	G47.36
	Sleep-related hypoxemia	G47.36
	Snoring	R06.83
	Catathrenia	–
Central disorders of hypersomnolence	Narcolepsy type 1	G47.411
	Narcolepsy type 2	G47.419
	Idiopathic hypersomnia	G47.11
	Kleine-Levin syndrome	G47.13
	Hypersomnia due to a medical disorder	G47.14
	Hypersomnia due to a medication or substance	F11 – F19
	Hypersomnia associated with a psychiatric disorder	F51.13
	Insufficient sleep syndrome	F51.12
	Long sleeper	–
Circadian rhythm sleep-wake disorder	Delayed sleep-wake phase disorder	G47.21
	Advanced sleep-wake phase disorder	G47.22
	Irregular sleep-wake rhythm disorder	G47.23
	Non-24-hour sleep-wake rhythm disorder	G47.24
	Shift work disorder	G47.26
	Jet lag disorder	G47.25
	Circadian sleep-wake disorder not otherwise specified	G47.20

(continued)

Table 2 (continued)

Group of sleep disorder	Diagnosis	ICD 10 code
Parasomnias	Disorders of arousal from NREM sleep	–
	Confusional arousals	G47.51
	Sleepwalking	F51.3
	Sleep terrors	F51.4
	Sleep-related eating disorder	G47.59
	REM sleep behavior disorder	G47.52
	Recurrent isolated sleep paralysis	G47.53
	Nightmare disorder	F51.5
	Exploding head syndrome	G47.59
	Sleep-related hallucinations	H53.16
	Sleep enuresis	N39.44
	Parasomnia due to a medical disorder	G47.54
	Parasomnia due to a medication or substance	F11 – F19
	Parasomnia unspecified	G47.50
	Sleep talking	–
Sleep-related movement disorders	Restless legs syndrome	G25.81
	Periodic limb movement disorder	G47.61
	Sleep-related leg cramps	G47.62
	Sleep-related bruxism	G47.63
	Sleep-related rhythmic movement disorder	G47.69
	Benign sleep myoclonus at sleep onset	G47.69
	Propriospinal myoclonus at sleep onset	G47.69
	Sleep-related movement disorder due to a medical disorder	G47.69
	Sleep-related movement disorder due to a medication or substance	F11 – F19
	Sleep-related movement disorder unspecified	G47.69
	Excessive fragmentary myoclonus	–
	Hypnagogic foot tremor and alternating leg muscle activation	–
	Sleep starts (hypnic jerks)	–
Other sleep disorders		G47.8
Appendix A: Sleep-related medical and neurological disorders	Fatal familial insomnia	–
	Sleep-related epilepsy	–
	Sleep-related headaches	–
	Sleep-related laryngospasm	–
	Sleep-related gastroesophageal reflux	–
	Sleep-related myocardial ischemia	–

disorders present the highest prevalence and are responsible for the highest direct costs in health-care systems worldwide, related to sleep disorders. Sleep apnea is a disorder with respiratory cessations during sleep. An apnea is counted if the duration is longer than 10 s. Most apnea events last for 30–50 s, but they may last more than a minute. In parallel with the respiratory cessation, oxygen saturation decreases due to no breathing. Apnea events end with a central nervous activation with parallel

increase in sympathetic tone resulting in increased heart rate and increased blood pressure during this so-called arousal. Apnea events are called obstructive if they are caused by a collapse of the upper airways when sleeping. Partial obstruction will result in hypopnea events with similar effects as apnea events. Therefore both types of events are counted together and are related to total sleep time; the apnea-hypopnea index (AHI) is the measure for severity. Central apnea events are characterized by a cessation of airflow, but the upper airways remain open. These are observed usually in patients with cardiac problems (i.e., heart failure). Some apnea events may have both an obstructive component with an obstruction of the upper airways and a component with now respiratory efforts, similar to a central apnea. These events are called mixed apnea events. These apnea events are observed in patients who start a sleep apnea therapy, and then this picture is called treatment-emergent apnea – as listed in the above list of diagnoses. While a few apnea events are observed in anybody during sleep, if the number of these events exceeds 5 events/hour of sleep (AHI $>$ 5 events/hour), this is diagnosed as mild sleep apnea. If the number of apneas exceeds 15 events/hour, this is moderate sleep apnea, and if more than 30 events/hour of sleep are found, this is severe sleep apnea.

In early epidemiological studies, the prevalence of obstructive sleep apnea (OSA) was estimated to be 4% in men and 2% in women (Young et al. 1993). Many large epidemiological studies including the investigation of sleep apnea were carried out since. Studies were performed in several countries worldwide. Over time these studies reported an increasing prevalence (Peppard et al. 2013). Peppard et al. reported obstructive sleep apnea (AHI $>$ 15 events/hour) in 10–17% of men depending on age and 3–9% in women depending on age. More recent studies have reported even higher prevalence. A population-based cohort study in Lausanne (HypnoLaus) reported moderate to severe sleep apnea (AHI $>$15 events/hour) in 49% of men and 23% of women in consecutive participants aged 40–85 years (Heinzer et al. 2015). With this high prevalence, a discussion on the definitions of sleep apnea and on the medical relevance of sleep apnea has been inaugurated.

Because apnea events occur during sleep, affected patients are not aware of this. Patients do not report apneas themselves. A bedpartner may have observed apnea events during sleep or the patient turning blue due to low oxygen when breathing ceases. Most patients with sleep apnea snore as a sign of higher upper airway collapsibility. During an apnea event, when airflow ceases, there is no snoring. Therefore snoring is usually intermittent. Apnea events are terminated by arousals. Arousals are most often visible in the sleep EEG as an increase in EEG frequency and as a shift toward higher sleep stages (Bonnet et al. 2007). Arousals with their central nervous activation open the upper airways to regain breathing and the activation interrupts sleep continuity (Eckert et al. 2014). Therefore patients are not able to reach deep sleep (sleep stage N3) if severely affected. In addition, light sleep (sleep stages N1 and N2) and REM sleep are interrupted by many arousals causing fragmented sleep. As a consequence of fragmented sleep patients experience and report unrefreshing sleep and are often sleepy during daytime (Eckert et al. 2014). Sleep apnea accompanied by excessive daytime sleepiness had been called obstructive sleep apnea syndrome (OSAS) in the past. Today, we know that excessive daytime sleepiness is a typical but not always reported consequence of sleep apnea and sleep physicians avoid the term ‘syndrome’. The numerous hypoxia events and reoxygenation after each apnea during sleep cause a stress to the endothelial system (Lavie 2003). The increase in sympathetic tone with each arousal causes vasoconstriction with increases in heart rate and blood pressure. These periodic changes result in additional stress to the vascular system during sleep. With these rapid and frequent changes, the normal regulation of sleep and sleep recreation becomes severely impaired. The normal lowering of heart rate and blood pressure and the lowering of vascular load cannot take place. The normal hormone secretion pattern during sleep is impaired. As a consequence these changes in physiology impose a risk factor for cardiovascular disorders such as

hypertension, cardiac arrhythmias, myocardial infarction, and stroke (Shahar et al. 2001). The association with hypertension is most obvious and had been reported early (Young et al. 1997). Sleep apnea is also a risk factor for diabetes and metabolic disorders (Resnick 2003). Patients experience this as early aging and lower performance.

Treatment of Sleep-Related Breathing Disorders

The therapy of choice for obstructive sleep apnea is nasal CPAP (continuous positive air pressure) (Mayer et al. 2017). The therapeutic mechanism is mechanical. It is a pneumatic stenting of the upper airways (Sullivan et al. 1981). This treatment is very effective if the nasal mask is tolerated by the patient (Sanders et al. 2008). The upper airways are opened during sleep by a continuous flow of room air resulting in a positive pressure of 4–15 cm H_2O in average. The exact pressure needed in a particular patient is titrated in the sleep laboratory. The pressure needed is more or less constant for a particular patient. The pressure needed varies by 1–3 cm H_2O depending on sleep stage and body position. Usually higher pressure are needed during REM sleep and sleeping supine. Since couple of years, machines are available which perform a continuous assessment of the upper airway obstruction by a high-frequency oscillation method. High frequency means 20 Hz and the superimposed oscillation is generated by a loudspeaker creating air vibrations on top of the supplied air pressure. Sensing reflecting oscillations allow to determine the collapse of the upper airways. Then pressure is increased automatically until the airways open. These machines determine automatically the required air pressure and are called APAP machines, automatic titrating CPAP (Morgenthaler et al. 2008). For patients who cannot tolerate to exhale against an increased pressure, another group of machines was developed, which lower the air pressure as soon as the patient wants to expire. This lowering of pressure can be set by the sleep physician, and it is usually 4–8 cm H_2O lower than the inspiratory pressure. These machines are called BPAP or (Bilevel PAP). The time for inspiration and expiration may be set to limit if needed. All these machines do not present a classical ventilation therapy because there is just a nasal mask, there is no volume control, and airflow is continuous. However the distinction between more sophisticated modes, like a time-controlled BPAP and conventional ventilation, becomes more fuzzy. Due to the high number of patients with obstructive sleep apnea and with prescribed CPAP or APAP devices, this is an economically very important market with all consequences for health-care and health insurance systems.

Sleep apnea therapy with all kind of ventilator devices require using the device each night and wearing a mask during each night. Some patients cannot tolerate the nasal mask. Then therapy adherence declines and as a consequence treatment efficiency declines as well. Patients are seeking for alternative treatments. Several alternative therapies for sleep apnea were developed. The most common and popular alternative therapy is the use of mandibular advancement devices (MAD). The oral devices produce a protrusion of the lower jaw by an average of 8 mm up to 15 mm. The devices can be used only if teeth are healthy and stable. The protrusion has to be titrated to reach maximum effectiveness in terms of apnea reduction snoring reduction and while avoiding pain and discomfort during the night and next day. The devices open the upper airways somewhat more. Accordingly they may turn apnea events into hypopnea events and definitely reduce snoring. Depending on the severity of upper airway collapse, this treatment may reduce the number of apnea and hypopnea events to zero. Most often a number of events stay. On average in large patient groups of any sleep apnea severity, a reduction of AHI by 50% is reported. In selected patient groups, results may be better. Today MAD is recommended if the CPAP mask is not tolerated or if the upper airway collapse is not very severe. The upper airway collapse that is not very severe is observed best by checking the effective CPAP pressure during CPAP titration trials. If the effective CPAP

pressure is low, say 6 cm H_2O, then there is a high chance of having a very effective MAD treatment.

Procedures

Diagnostic procedures in sleep medicine start first with an assessment of complaints and symptoms regarding sleep behavior and sleep habits. Then follows an assessment of complaints and symptoms. In order to standardize this assessment, a large number of well-established and validated questionnaires are in use. The questionnaires are then followed by clinical investigations to assess clinical features associated with the sleep disorder. A recording of a few characteristic physiological signals at home follow to assess respiration and oxygen saturation during sleep. The last and final procedure in the diagnosis of sleep disorders and sleep-disordered breathing in particular is polysomnography (PSG) in a sleep laboratory or sleep center. This requires trained personnel for the recording and for the interpretation of the recorded data. A sleep center is able to diagnose all disorders as specified above.

Questionnaires

A much used questionnaire covering many aspects with a focus on insomnia problems is the Pittsburgh Sleep Quality Index (PSQI) mentioned earlier already (Buysse et al. 1989). Another general and frequently used questionnaire is the Epworth sleepiness scale (ESS) (Johns 1991). This questionnaire consists of eight questions describing situations where one might fall asleep. The situations become increasingly severe and unwanted. Each question is answered with a likelihood of the situation ranging from zero to three. The maximum score is 24 then. Today a threshold of 11 is regarded as being sleepy more than normal. This questionnaire is used much in patients with central disorders of hypersomnolence and in patients with sleep-related breathing disorders. In patients with sleep-related breathing disorders, the ESS is recognized as not being sensitive and not being predictive for the diagnosis of sleep apnea. This is mainly due to the fact that sleepiness is not always a sign of sleep-related breathing disorders and that patients with sleep-related breathing disorders are not necessarily sleepy (Qaseem et al. 2014). However the association between sleep-related breathing disorders and sleepiness is so intuitive that the application of the ESS in these patients remains to be very popular.

If a patient complains about problems initiating and maintaining sleep, the Insomnia Severity Index (ISI) is used (Morin et al. 2011). This short questionnaire is now used in many pharmacological studies to assess the severity and interventional improvement of insomnia. In clinical sleep centers, this index is used to document insomnia and its severity. Sleep centers use the ISI often in combination with Beck Depression Inventory (BDI) in order to distinguish between depression and depression-induced insomnia and insomnia as the primary complaint.

For the assessment of a periodic leg movement syndrome (PLMS) and restless legs (RLS) as part of the sleep-related movement disorders, a specific questionnaire has been developed. This RLS-DI questionnaire is used in many studies and in clinical assessment of PLMS and RLS. This is a very useful and well-validated tool (Walters et al. 2003).

Sleep-related breathing disorders have a high prevalence as previously stated (Peppard et al. 2013). There is a body of evidence that sleep-related breathing disorders have cardiovascular consequences (Shahar et al. 2001; Marin et al. 2005). Therefore there is a clinical need to identify sleep apnea and treat patients if they suffer from moderate and severe sleep apnea (Mayer et al. 2017). Furthermore, evidence showed that surgical interventions of many kinds applied to patients with sleep-related breathing disorders result in much higher postsurgical complication rates. Therefore anesthesia guidelines recommend to assess sleep-related breathing disorders prior to elective surgical interventions. In pharmacological trials, there may be need to exclude sleep-related breathing disorders on a fast and reliable basis. This assessment should be simple and should have a high sensitivity with adequate

specificity. Exactly for this purpose, several questionnaires were developed and tested. A systematic review and meta-analysis of questionnaires had been performed and sensitivity and specificity had been listed (Abrishami et al. 2010). A more recent meta-analysis has resulted in a clinical guideline and can be used to review the range of studies using questionnaires and other methods as well (Qaseem et al. 2014). A questionnaire for screening for sleep apnea by anesthesiologists prior to surgical interventions had been presented and validated. This is the STOP-BANG questionnaire (Chung et al. 2008). Each letter stands for a question or finding. The STOP-BANG questionnaire asks for snoring (S), tiredness (T), observed and reported apnea events (O), being treated for high blood pressure (P), BMI higher than 35 kg/m^2 (B), age higher than 50 years (A), neck circumference higher than 40 cm (N), and male gender (G). Each positive answer adds a point. The risk for sleep apnea is increased if the score is 3 and higher (Ong et al. 2010). Sensitivity and specificity are given in Table 3. This questionnaire is regarded as having high sensitivity and specificity compared to other questionnaires. Another questionnaire developed for family physicians is the Berlin questionnaire (Netzer et al. 1999). This questionnaire was developed as a tool for general physicians to assess the risk for sleep apnea (compare Table 3). A very recent development is the NoSAS score which claims to be better than the STOP-BANG and the Berlin questionnaire (Marti-Soler et al. 2016). This questionnaire does not simply use yes/no questions but carefully assigns points to the same items used by the other questionnaires. Other research groups have developed clinical scores which combine clinical findings often seen in sleep apnea and a few questions to one apnea score (Flemons et al. 1994). In principle these scores are similar to the STOP-BANG questionnaire.

The questionnaires described above were selected because they have a worldwide distribution and validated translation into many languages. Of course there are more, sometimes regionally popular, questionnaires. In order to become considerably better than questionnaires, which means to achieve a higher specificity, a recording of physiological signals over the night is required. This is done with medical devices allowing long-term recording, often denoted as 'sleep lab' meaning 'polysomnography'. A simplified family of devices is denoted as polygraphy or home sleep apnea testing (HSAT). Practically, this method is now very much used in clinical practice to diagnose sleep apnea and to initiate treatment because it does not require a hospital or supervised bed. This method is described with more detail below.

Polysomnography

The reference method for the diagnosis of sleep-related breathing disorders is cardiorespiratory polysomnography. This method includes a quantitative assessment of sleep, respiration, vegetative functions, movements, and behavior. All physiological functions involved in sleep and in sleep disorders are quantitatively assessed. The method and its assessment are well described in the standardized manual of the AASM (Berry et al. 2016). The signals which track a physiological function for the duration of a night are listed in Table 4. Sensor methods were evaluated over the past and are now found to be reliable and satisfactory for good results. Electrode types, sensor types, and number of signals are well standardized, and polysomnographic equipment is available from several companies worldwide. Polysomnography is used both for research and clinical purposes. The biggest differences between systems are found in the flexibility of hardware and the accompanying software for the management of additional sensors and signals. This distinguishes systems which can be used for clinical purpose only and systems which can be used for research as well. Good software provides a good automatic analysis of signals which is used for a pre-evaluation in clinical work and which provides tools for quantitatively assessing signals for research purposes.

The sleep recording needs to be evaluated visually according to the rules specified in the AASM manual (Berry et al. 2016). The recorded signals are displayed in 30 s epochs and then evaluated or

Table 3 Questionnaires for sleep-related breathing disorders. The values are given as percentages and stem from two studies (Silva et al. 2011 und Pereira et al. 2013). The "clinical score" is composed of snoring, age, blood pressure, and male gender (Flemons et al. 1994)

	AHI > = 5			AHI > = 15			AHI > = 30		
	Sensitivity	Specificity	Positive predictive value	Sensitivity	Specificity	Positive predictive value	Sensitivity	Specificity	Positive predictive value
ESS >10				39.0	71.4	64.8	46.1	70.4	68.7
Berlin questionnaire	86	25	91.7	91	28	73.4	89	18	45.9
STOP-BANG	90	42	93.7	93	28	73.9	96	21	48.6
NoSAS				79	69	47			
Clinical score	33	83	95.0	35	78	77.5	36	72	50.0
Home sleep testing	87	67	96.2	77	95	97.1	50	93	84.8

Table 4 Functions and associated signals assessed during polysomnography

Function	Signal	Sensor
Sleep	Electroencephalogram EEG	Electrophysiological electrode: 3–6 standardized positions
	Electrooculogram EOG	Electrophysiological electrode: Left and right eyes
	Electromyogram mental EMG	Electrophysiological electrode: 2 leads
Respiration	Airflow	Temperature-sensing sensors at the nose and mouth Pressure sensing (nasal prongs) Pneumotachograph with full face mask
	Respiratory effort	Inductive plethysmography with two belts Esophageal pressure
	Oxygen SaO_2 or pO_2	Pulse oximetry sensor Transcutaneous O_2
	Carbon dioxide CO_2% or pCO_2	End-tidal CO_2 from expired air with ultrared absorption sensor Transcutaneous CO_2
	Snoring	Microphone: Supraglottis or room
	Respiratory movement (indirect)	Radar/microwave technology Matt in the bed or mattress sensors ECG-derived respiration
Cardiac	Electrocardiogram	Electrophysiological electrode: 1 lead
Movement	Electromyogram of legs and arms	Electrophysiological electrode
	Body	Actigraphy or simple movement sensor
Behavior	Video	Room camera
	Voice	Room microphone
	Body position	3-D acceleration sensor
Additional options	Body core temperature	Rectal or ear probe, thermo-pill
	Gastric pH	Antimony sensor probe
	Electrodermal activity	Resistance probe, no standard
	Blood pressure	Finger photoplethysmography, pulse transit time as surrogate
	Pulse wave	Optical pulse plethysmography/pulse oximetry sensor, no standard sensors

scored (Fig. 1). The sleep EEG with EOG and EMG is used to score sleep stages in categories wake (W) with alpha waves in the EEG; high muscle tone; light sleep (N1), which is transitional sleep with less than 50% of alpha waves lower muscle tone and rolling eye movements; light sleep with specific EEG patterns such as sleep spindles and K-complexes (N2); and slow-wave sleep or deep sleep with more than 20% of time per epoch with delta waves of a frequency between 0.5 Hz and 2 Hz and an amplitude of at least 75 microvolt (N3). The rapid-eye-movement sleep or REM-sleep is denoted as R and is characterized by mixed EEG waves, low muscle tone, and rapid eye movement. This sleep stage is associated with frequent reports of dreaming. The AASM manual sets rules how to score short awakenings, termed as arousals from sleep. These arousals are scored based on EEG activations and reflect cortical activation only (Bonnet et al. 2007). The AASM manual also sets rules how to classify obstructive and central apnea events, obstructive and central hypopnea events, and hypoventilation. The manual sets rules for scoring movement events to classify leg movements and periodic leg movements. A separate chapter defines all the parameters which are needed for a polysomnography report. These are specified below.

Cardiorespiratory polysomnography is attended by trained sleep nurses who can attach or readjust sensors when contact problems occur during the night. Sleep nurses attend all night in a sleep center to be available for patient calls as

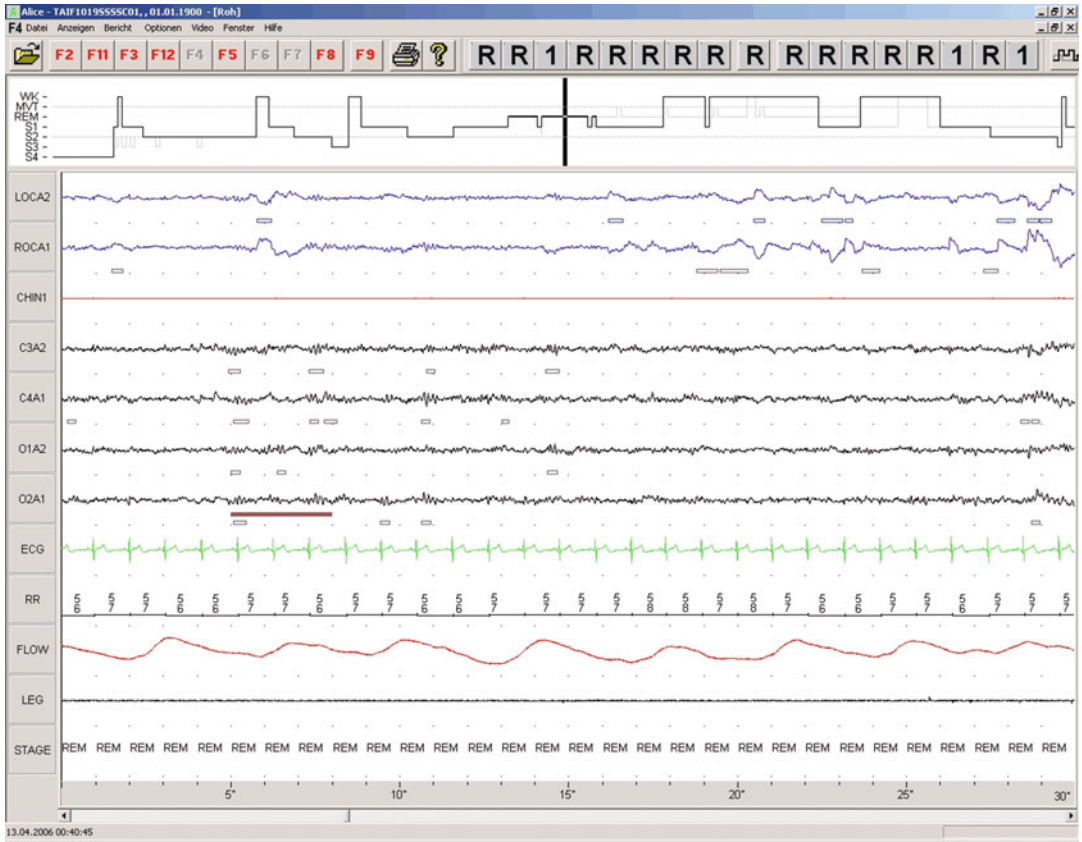


Fig. 1 Example of polysomnography with a 30 s epoch of REM sleep. Rapid eye movements can be observed in the EOG tracings (LOCA2 and ROCA1). A low muscle tone is visible (CHIN1). The EEG shows in four leads (C3A2, C4A1, O1A2, O2A1) mixed frequency low amplitude activity which is characteristic for REM sleep

well. Even if simpler methods like questionnaires or home sleep apnea testing are used today, as soon as sleep-disordered breathing may be more complex or as soon as comorbidities are present in a particular patient, a polysomnography is required for a full assessment of the disorder (Qaseem et al. 2014; Mayer et al. 2017).

Home Sleep Apnea Testing – Polygraphy

In order to reduce costs and expand the capacities for the diagnosis of sleep-disordered breathing, many portable monitoring devices for sleep apnea were developed over the last two decades. A systematic review and meta-analysis for these devices had been compiled (El Shayeb et al. 2014). A review combined with a clinical guideline has been released for the management of sleep-disordered breathing as well (Qaseem et al. 2014). The portable devices have achieved high sensitivity and specificity up to an extent that they are used for out-of-center diagnosis of sleep-related breathing disorders in the majority of patients under certain conditions (Collop et al. 2011). There is an ambition to diagnose or at least recognize sleep apnea even simpler than that. Short tests during daytime or questionnaires are desirable. A number of tests had been developed and questionnaires as well. The questionnaires had been described above. The guideline by the American college of physicians (ACP) evaluated the modalities for the diagnosis of sleep apnea and concluded with recommendations in how far portable monitoring can be used and in how far questionnaires, described above, can be used for sleep-related breathing disorders (Qaseem et al.

2014): "The ACP recommends a sleep study for patients with unexplained daytime sleepiness. The ACP recommends polysomnography for diagnostic testing in patients suspected of obstructive sleep apnea. ACP recommends portable sleep monitors in patients without serious comorbidities as an alternative to polysomnography where polysomnography is not available for diagnostic testing." A more recent position statement of the American Academy of Sleep Medicine states that home sleep apnea testing (HSAT) has to be prescribed by a physician, based on medical history and face-to-face examination (Rosen et al. 2017). The raw data of HSAT must be reviewed by certified sleep physician. An automatic analysis of raw data is not sufficient. A HSAT is not recommended to screen asymptomatic populations. HSAT systems are used worldwide. In some regions of the world, mainly in Europe, they are called Polygraph (PG), a term derived from "polysomnography" but without the "somno" component because polygraphy does not record sleep with EEG, EOG, and EMG.

To characterize and classify systems for the home sleep apnea testing, four levels had been introduced first (Flemons et al. 2003). When a cardiorespiratory polysomnography, as defined above, is meant, then this is called a level 1 diagnosis, which is the highest level with highest quality and with most effort made. Due to the requirement of attending personnel, this is performed in a sleep lab setting usually. If the same system is used at home or under experimental conditions outside a sleep lab, without attending personnel, then this is called a level 2 sleep assessment. A level 3 recording device describes a system which can record respiration (respiratory flow, respiratory movement, oxygen saturation), heart rate, and body position with usually 4–6 channels (Flemons et al. 2003). These systems were developed to record sleep apnea at home or elsewhere outside a sleep center. These are called HSAT or PG today. Simpler systems, such as an actigraphy, an oximetry with heart rate, and a respiratory flow recording, are called level 4 systems. They record typically 1–3 channels only. Developments during the last 10 years with sophisticated signal analysis from few signals allow to detect sleep apnea from few signals. Therefore another classification system had been introduced, the SCOPER system, which is based on assessment of functions instead of simply counting signals (Collop et al. 2011). The SCOPER system checks the validated assessment of sleep (S), cardiovascular functions (C), oxygen saturation (O), body position (P), respiratory effort (E), and respiratory flow (R). In order to see whether a specific system fulfills the criteria, validation studies are examined. With this it might be possible that a sophisticated analysis of pulse wave recording may reveal sleep stages in terms of wake, non-REM, REM sleep, heart rate, and respiratory effort to detect apnea events, all with one signal. Therefore this new SCOPER system is more appropriate to evaluate and assess usefulness of modern systems to diagnose sleep-related breathing disorders. A class of systems which did profit much from this new SCOPER classification where those making use of arterial peripheral tone based on finger pulse wave recordings (Yalamanchali et al. 2013). These systems use the finger pulse wave with a sophisticated analysis based on pulse amplitude changes and pulse-to-pulse interval changes over time instead of directly recording airflow or respiratory movement. The Watch-PAT system derives respiration, apnea events, both obstructive and central, and estimates non-REM and REM sleep from the pulse timing behavior. HSAT systems are those who fulfill the requirements to record sleep-related breathing disorders with sufficient sensitivity and specificity. Using the conditions described above, HSAT systems now become the diagnostic tool of first choice to diagnose sleep apnea unless comorbidities that are present in sleep apnea need to be excluded (Rosen et al. 2017).

Evaluation (Parameters and Statistical Evaluation)

The AASM manual explains how to evaluate sleep stages and all-related events. Beside this the manual specifies and defines parameters summarizing the evaluation and which should be included in a polysomnography report (Berry et al. 2016). The parameters and quantitative descriptions derived from polysomnography

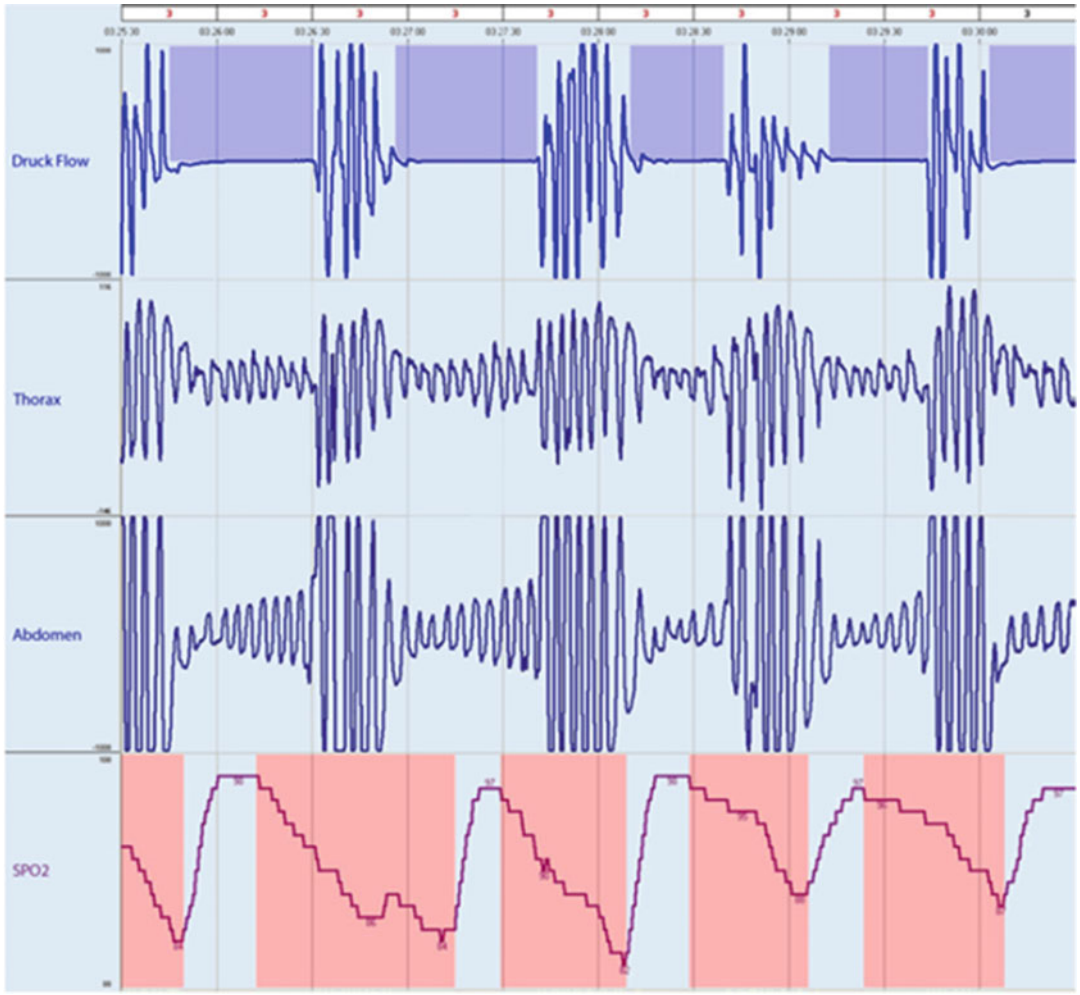


Fig. 2 A 5 min window with five obstructive apnea events is shown. The tracings from top to bottom present nasal pressure, thoracic movements, abdominal movements, oxygen saturation

relate to the following groups: global sleep metric, sleep stages, respiration, cardiovascular assessment, movement assessment, and further observations. Global sleep metrics are total recording time (TRT), the time between lights out and lights on, total sleep time (TST), and percent sleep efficiency, which is the ratio between TRT and TST. The sleep stages are specified by minutes in each sleep stage and percent related to TST, together with sleep latency, stage R latency, and wake after sleep onset. The arousals are quantified by total number and number related to TST, called as arousal index (ArI). Cardiac events are mainly noted as yes or no, and heart rate is given by highest, lowest, and mean values. Limb movements are given as total number and as an index related to TST, both with and without arousal. The respiratory parameters are a long list. This starts with number of apnea events, obstructive, central, mixed, hypopnea events, obstructive, and central and a sum of all events. An important step in the evaluation is to distinguish the type of respiratory events. Obstructive apnea events are those where a cessation of airflow is detected, but respiratory effort continues as recognized by movements in thoracic and abdominal belt measurements (Fig. 2). Central apnea events are those where a cessation of airflow is detected and no respiratory

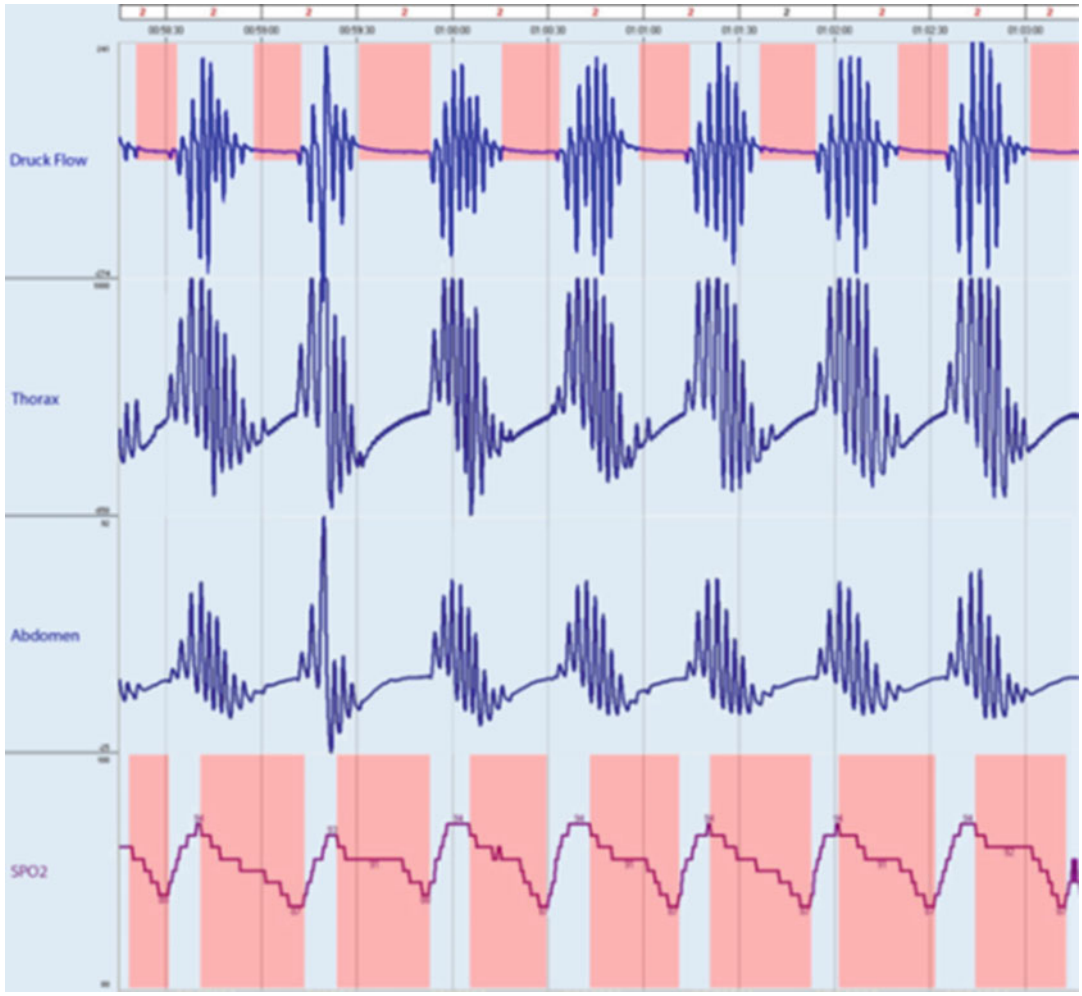


Fig. 3 A 5 min window with central apnea events forming a Cheyne-Stokes breathing pattern is shown. The tracings from top to bottom present nasal pressure, thoracic movements, abdominal movements, oxygen saturation

effort is found as recognized by movements in thoracic and abdominal belt measurement (Fig. 3). However both types of events are accompanied by drops in oxygen saturation (SpO_2). Usually obstructive apnea events are longer (30 to 60 seconds) compared to central apnea events (20 to 40 seconds). Correspondingly the drop in oxygen saturation, the desaturation is somewhat less in central apnea events. If the respiratory flow in cases of central sleep apnea forms a spindle like crescendo decrescendo pattern with a cycle length of 40 seconds or more, then this is called Cheyne-Stokes breathing (Fig. 3). This is typically found in patients with heart failure. The apnea-hypopnea index (AHI) is the sum of all events per TST. For the other event types, equivalent indices are calculated. As such there is an apnea index (AI) for obstructive, mixed, and central apneas. Then respiratory effort-related events and oxygen desaturation events are counted. The occurrence of hypoventilation, Cheyne-Stokes breathing, periodic breathing, and snoring is noted. Possibly additional oxygen saturation statistics can be presented.

The complete list of parameters with definitions is specified in the AASM manual for scoring of sleep and associated events (Berry et al. 2016).

Critical Assessment of the Sleep Recording

In the beginning of sleep research and sleep medicine polysomnography was the investigational tool for sleep centers. As sleep research moved into clinical sleep medicine, practical and economic aspects became more and more important. Today sleep apnea is mainly diagnosed using home sleep apnea testing (HSAT) as explained in more detail above. The role of polysomnography in the diagnosis of sleep apnea is decreasing (Mayer et al. 2017). This process is continuing and not finished (Hirshkowitz 2016). With technological developments, home sleep apnea testing becomes more sophisticated and as a consequence more sensitive and specific. This process does not only apply to sleep-related breathing disorders but also to other groups of sleep disorders. With new smartphone applications, it is possible to track sleep. However most smartphone applications were developed intuitively according to the thought: activity correlates to wake and no activity correlates to sleep. The lowest activity might correlate to very deep sleep. This does really reflect our knowledge about slow-wave sleep and REM sleep. Few smartphone apps had been validated. However this technological field is improving quickly. New apps make use of additional sensors like camera, noise, and external bed matts, and external pulse wave sensors. And new apps are validated against polysomnography. These new apps cross the line between simple gadgets for wellness and lifestyle to medical useful devices. It is possible that new apps are able to track sleep and wake and sleep problems over prolonged periods of time with adequate accuracy. However a good validation against polysomnography and considering the specific group of subjects being investigated is always needed. With these recent developments, the role of polysomnography in future may no longer be clinical routine recordings. But polysomnography may become a validation tool and a research tool again (Hirshkowitz 2016).

An important discussion is currently ongoing about the role of the AHI as a severity parameter for sleep apnea (Penzel et al. 2015). In view of the high prevalence as reported by some epidemiological studies (Heinzer et al. 2015) and in view of not showing any positive effect on cardiovascular mortality when treating patients with CPAP just according to their high AHI (McEvoy et al. 2016), there is some concerns that the AHI may not be the optimal parameter to express the severity of the disease. While this had been assumed for long and while it is clear that the pathophysiology is linked to the collapsibility of the upper airways (Eckert et al. 2014), the AHI was regarded as a simple and reliable surrogate for the severity of the disease. Previous studies did show that survival over 12 years is associated with the AHI and that an effective lowering of AHI by CPAP did result in much lower mortality and morbidity (Marin et al. 2005). Only with the recent studies, this parameter is challenged. Therefore a discussion is in place that previous studies did show such impressive beneficial results, because studies were performed on clinical populations seeking help in a hospital setting. And as soon as subjects without symptoms and without complaints are investigated and if they are treated according to AHI only, then these positive effects may disappear. There are thoughts that the AHI may be an indicator for an increased cardiovascular risk and that the elevated AHI must be seen together with the complete clinical picture with other symptoms and findings together. In order to clarify these different mechanisms, the model currently under development is that there are different phenotypes with obstructive sleep apnea (Penzel et al. 2015). Some subjects may develop sleep apnea as part of normal aging of the upper airways. Other subjects develop obstructive sleep apnea as an accompanying factor with obesity. Another group develops obstructive sleep apnea due to morphological retrognathia. Another group may suffer from a narrow upper airway or a more collapsible upper airway with excessive soft tissue. There may be also a group with neural deficits in respiratory regulation during sleep. A hypersensitivity to CO_2 may cause a hyperventilation and then an apnea as a compensation to hyperventilation. This compensation would be physiologic because during sleep, the respiratory system tolerates higher CO_2 levels. These concepts may form a

number of different phenotypes, which, until now, are not well defined and characterized.

Alternative Treatments for Sleep Apnea

The first-line treatment for obstructive sleep apnea is CPAP. CPAP with mechanical splinting of the upper airways has a very high effectiveness, if patient compliance with therapy is good. Since many patients cannot tolerate the nasal mask, the second choice for treatment is an individually fitted mandibular advancement device. Oral appliances or mandibular advancement devices are very popular and much better accepted by a majority of patients because they are more comfortable than a nasal mask. However this treatment is less effective (Schwartz et al. 2017). On average the AHI is lowered by 50%. Since the mandibular advancement device causes a protrusion of the lower jaw, a widening of the upper airway space is achieved. This widening is partial and may be enough to overcome the upper airway collapse. However it may be not enough in some subjects, and then apnea events are converted into hypopnea events. Hypopnea events might be converted into snoring. A conversion was not achieved because the widening by oral appliances was by far not enough. Accordingly there is a high variability across patient groups. Unfortunately the selection criteria for finding those patients who benefit most are not clear before treatment initiation. Some patients cannot tolerate an oral appliance in their mouth over the night. Other patients may not wear such a device due to their tooth conditions. If this therapy of second choice cannot be applied, then other alternatives need to be investigated.

Shortly after sleep apnea was discovered and about at the same time as CPAP was invented, a concept to open the upper airway by electrical stimulation of the corresponding nerve was patented. A nervus hypoglossus stimulating device was developed and first tested in models and soon thereafter applied to humans with obstructive sleep apnea (Schwartz et al. 2001). This pivotal trial in eight subjects proved the success of the concept. A few technological problems were detected when these eight subjects continued to use the nervus hypoglossus stimulation over an extended period of time. The principle was a sensing of inspiratory effort and a stimulation of the hypoglossus nerve during inspiration on one side only. The stimulation causes a protrusion of the tongue, and with this protrusion, the upper airways are widened and tissue becomes stiffer during the stimulation phase. A remote control and a timer in the stimulating device took care that the stimulation period was synchronized with the sleep period. It became clear that the positioning of the stimulating electrodes, the cuff around the nerve, had to be placed carefully on the right position in order to achieve maximum effects. The pivotal trial did show that widening by stimulation could lower AHI by 50% in average. Only a couple of years later, the idea was picked up again, and three competing companies brought devices into clinical practice to be evaluated in large trials in order to achieve FDA approval. Two companies did achieve FDA approval in the meantime. The large studies did fulfill safety, security, and efficiency expectations. If patients were carefully selected according to weight (BMI < 35 kg/m^2) and according to a reactive and oval-shaped upper airway, then the effectiveness of the stimulation treatment to lower AHI was somewhat higher than 50% (Strollo et al. 2014). The device is expensive compared to CPAP, and the procedure is invasive compared to CPAP. The titration is similar to CPAP, because during the titration night, the optimal electrodes for stimulation must be identified and the optimal current for stimulation must be set in order to achieve maximal widening and no arousal.

Patients with central apnea and Cheyne-Stokes respiration under conditions of heart failure may be treated for their heart failure first. If patients treated for heart failure according to guidelines and if Cheyne-Stokes respiration persists, then a trial of CPAP or other ventilation may be initiated if the LVEF is not below 45%. A new approach in these patients with central sleep apnea is to test pharyngeal nerve stimulation. However studies with this treatment are very small and long-term studies are missing. This phrenic nerve

stimulation therapy for Cheyne-Stokes respiration is seeking FDA approval now.

A considerable number of studies with various kinds of surgical interventions were performed in the past. Maxillofacial surgery was able to widen the upper airways effectively. If patients were carefully selected, then this procedure was highly effective. However only few patients do fulfill the restrictive selection criteria for this intense surgical procedure. Other ENT-related surgical procedures had the upper airways as a target. The therapeutic principle was always to remove obstacles in the upper airways to allow breathing during sleep. A much used method was the uvulopalatopharyngoplasty (UPPP) and also adenotonsillectomy. These surgical procedures, and their variants, had only partial success. The AHI was lowered by 30–50% depending on the study. In these cases only mechanical obstacles can be removed. The neural and functional components of upper airway collapse are impossible to be treated surgical. And unfortunately, a prediction on who might benefit more and who benefits less from surgical procedures could not be established despite many efforts.

Positional therapy is useful when sleep apnea has a major positional component. It is estimated that about 10% of patients with obstructive sleep apnea have a largely position dependent sleep apnea with more events or with a higher collapsibility when sleeping supine. These patients may benefit from positional trainers such as pillow vest or electronic devices which train the sleeping subject to avoid the supine position.

There are only few pharmacological therapy approaches in place. A good overview is provided by Gautier et al. (2017). Antihypertensive drugs may have some beneficial effects on sleep apnea. A few experimental trials could show only very small effects. Sleep apnea is associated with inflammatory processes and may potentially lead to atherosclerosis following hypoxia stress (Lavie 2003). Anti-inflammatory drug treatment has not been evaluated for effects on sleep apnea. Acetazolamide has been investigated in patients with sleep apnea both at normal and high altitude. A recent meta-analysis has investigated the effect of acetazolamide on sleep apnea in high altitude and found that acetazolamide does reduce apnea in terms of AHI but is more beneficial in healthy subjects than in subjects with sleep apnea (Liu et al. 2017). Anti-oxidative drugs may have a beneficial effect because they follow the same concept to prevent inflammatory consequences of the repeated intermittent hypoxia. Several studies are conducted now to systematically evaluate effects of several drugs. Another approach for drug applications is to influence local receptors of the upper airways (Wirth et al. 2013). However human trials have not been presented.

In summary pharmacological treatment will have a good chance as another alternative therapy. Especially in view of the change in concept for sleep apnea, there may be phenotypes where sleep apnea is not so severe but still annoying. If sleep apnea is reduced then potentially the corresponding increase in cardiovascular risk is reduced as well. Therefore we can expect beneficial effects of pharmacological treatment, if proven to be successful.

References and Further Reading

AASM (2014) International classification of sleep disorders. Third edition (ICSD-3). American Academy of Sleep Medicine, Darien

Abrishami A, Khajehdehi A, Chung FA (2010) Systemic review of screening questionnaires for obstructive sleep apnea. Can J Anesth 85:423–438

Berry RB, Brooks R, Gamaldo CE, Harding SM, Lloyd RM, Marcus CL, Vaughn BV, for the American Academy of Sleep Medicine (2016) The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 2.3. American Academy of Sleep Medicine, Darien

Bonnet MH, Doghramji K, Roehrs T, Stepanski EJ, Sheldon SH, Walters AS, Wise M, Chesson AL Jr (2007) The scoring of arousal in sleep: reliability, validity, and alternatives. J Clin Sleep Med 3:133–145

Buysse DJ, Reynolds ICF, Monk TH et al (1989) The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 28:193–213

Chung F, Yegneswaran B, Liao P, Vairavanathan S, lslam S, Khajehdehi A, Shapiro CM (2008) STOP questionnaire a tool to screen patients for obstructive sleep apnea. Anesthesiology 108:812–821

Collop NA, Tracy SL, Kapur V, Mehra R, Kuhlmann D, Fleishman SA, Ojile JM (2011) Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. J Clin Sleep Med 7:531–548

Eckert DJ, Younes MK (2014) Arousal from sleep: implications for obstructive sleep apnea pathogenesis and treatment. J Appl Physiol 116:302–313

El Shayeb M, Toper LA, Stafinski T, Pawluk L, Menon D (2014) Diagnostic accuracy of level 3 portable sleep tests versus level 1 polysomnography for sleep-disordered breathing: a systematic review and meta-analysis. Canad Med Assoc J 186:E25–E51. https://doi.org/10.1503/cmaj.130952

Flemons WW, Whitelaw WA, Brant R, Remmers JE (1994) Likelihood ratios for a sleep apnea clinical prediction rule. Am J Respir Crit Care Med 150:1279–1285

Flemons WW, Littner MR, Rowley JA et al (2003) Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American Thoracic Society. Chest 124:1543–1579

Gautier-Veyret E, Pépin JL, Stanke-Labesque F (2017) Which place of pharmacological approaches beyond continuous positive airway pressure to treat vascular disease related to obstructive sleep apnea? Pharmacol Ther S0163-7258(17):30297–30298. https://doi.org/10.1016/j.pharmthera.2017.12.006

Heinzer R, Marques-Vidal P, Marti-Soler H et al (2015) Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. Lancet Respir 3:310–318

Hirshkowitz M (2016) Polysomnography challenges. Sleep Med Clin 11:403–411

Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 14:540–545

Lavie L (2003) Obstructive sleep apnoea syndrome – an oxidative stress disorder. Sleep Med Rev 7:35–51

Liu HM, Chian IJ, Kuo KN, Liou CM, Chen C (2017) The effect of acetazolamide on sleep apnea at high altitude: a systematic review and meta-analysis. Ther Adv Respir Dis 11:20–29. https://doi.org/10.1177/1753465816677006

Marin JM, Carrizo SJ, Vicente E, Agusti AG (2005) Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. Lancet 365(9464):1046–1053

Marshall J, Hildebrandt S, Sydow H, Nolting HD, Burgart E, Woköck T (2017) Gesundheitsreport 2017. Medhochzwei, DAK-Gesundheit/Heidelberg

Marti-Soler H, Hirotsu C, Marques-Vidal P, Vollenweider P, Waeber G, Preisig M, Tafti M, Tufik SB, Bittencourt L, Tufik S, Haba-Rubio J, Heinzer R (2016) The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. Lancet Respir Med 4:742–748

Mayer G, Arzt M, Braumann B, Ficker JH, Fietze I, Galetke W, Maurer JT, Orth M, Penzel T, Pistner HP, Randerath W, Rösslein M, Sitter H, Stuck BA (2017) S3-Leitlinie Nicht erholsamer Schlaf/Schlafstörungen - Kapitel "Schlafbezogene Atmungsstörungen". Somnologie 20 (Suppl s2):S97–S180

McEvoy RD, Antic NA, Heeley E, Luo Y, Ou Q, Zhang X, Mediano O, Chen R, Drager LF, Liu Z, Chen G, Du B, McArdle N, Mukherjee S, Tripathi M, Billot L, Li Q, Lorenzi-Filho G, Barbe F, Redline S, Wang J, Arima H, Neal B, White DP, Grunstein RR, Zhong N, Anderson CS, SAVE Investigators and Coordinators (2016) CPAP for prevention of cardiovascular events in obstructive sleep apnea. N Engl J Med 375:919–931

Morgenthaler TI, Aurora RN, Brown T et al (2008) Practice parameters for the use of autotitrating continuous positive airway pressure devices for titrating pressures and treating adult patients with obstructive sleep apnea syndrome: an update for 2007. An American Academy of Sleep Medicine report. Sleep 31:141–147

Morin CM, Belleville G, Bélanger L, Ivers H (2011) The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 34:601–608

Netzer NC, Stoohs RA, Netzer CM (1999) Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 131:485–491

Ohayon MM (2011) Epidemiological overview of sleep disorders in the general population. Sleep Med Rev 2:1–9

Ohayon MM, Zulley J (2001) Correlates of global sleep dissatisfaction in the German population. Sleep 24:780–787

Ong TH, Raudha S, Fook-Chong S et al (2010) Simplifying STOP-BANG: use of a simple questionnaire to screen for OSA in an Asian population. Sleep Breath 14:371–376

Penzel T, Peter H, Peter JH, Becker HF, Fitze I, Fischer J, Mayer G, Podszus T, Raschke F, Riemann D, Schäfer T, Sitter H (2005) Gesundheitsberichterstattung des Bundes. Heft 27: Schlafstörungen. Robert-Koch Institut

Penzel T, Schöbel C, Fietze I (2015) Revise respiratory event criteria or revise severity thresholds for sleep apnea definition. J Clin Sleep Med 11:1357–1359

Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM (2013) Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol 177:1006–1014

Pereira EJ, Driver HS, Stewart SC et al (2013) Comparing a combination of validated questionnaires and level III portable monitor with polysomnography to diagnose and exclude sleep apnea. J Clin Sleep Med 9:1259–1266

Qaseem A, Dallas P, Owens DK, Starkey M, Holty JE, Shekelle P (2014) Diagnosis of obstructive sleep apnea in adults: a clinical practice guideline from the American college of physicians. Ann Intern Med 161:210–220

Resnick HE, Redline S, Shahar E, Gilpin A, Newman A, Walter R, Ewy GA, Howard BV, Punjabi NM, Sleep Heart Health Study (2003) Diabetes and sleep

disturbances: findings from the Sleep Heart Health Study. Diabetes Care 26:702–709

Rosen IM, Kirsch DB, Chervin RD, Carden KA, Ramar K, Aurora RN, Kristo DA, Malhotra RK, Martin JL, Olson EJ, Rosen CL, Rowley JA, American Academy of Sleep Medicine Board of Directors (2017) Clinical use of a home sleep apnea test: an American Academy of sleep medicine position statement. J Clin Sleep Med 13:1205–1207

Sanders MH, Montserrat JM, Farre R et al (2008) Positive pressure therapy: a perspective on evidence-based outcomes and methods of application. Proc Am Thorac Soc 5:161–172

Schlack R, Hapke U, Maske U, Busch MA, Cohrs S (2013) Häufigkeit und Verteilung von Schlafproblemen und Insomnie in der deutschen Erwachsenenbevölkerung. Bundesgesundheitsblatt 6:740–748

Schwartz AR, Bennett ML, Smith PL, De Backer W, Hedner J, Boudewyns A, Van de Heyning P, Ejnell H, Hochban W, Knaack L, Podszus T, Penzel T, Peter JH, Goding GS, Erickson DJ, Testerman R, Ottenhoff F, Eisele DW (2001) Therapeutic electrical stimulation of the hypoglossal nerve in obstructive sleep apnea. Arch Otolaryngol Head Neck Surg 127:1216–1223

Schwartz M, Acosta L, Hung YL, Padilla M, Enciso R (2017) Effects of CPAP and mandibular advancement device treatment in obstructive sleep apnea patients: a systematic review and meta-analysis. Sleep Breath. https://doi.org/10.1007/s11325-017-1590-6

Shahar E, Whitney CW, Redline S, Lee ET, Newman AB, Nieto FJ, O'Connor GT, Boland LL, Schwartz JE, Samet JM (2001) Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. Am J Respir Crit Care Med 163:19–25

Silva GE, Vana KD, Goodwin JL et al (2011) Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. J Clin Sleep Med 7:467–472

Strollo PJ Jr, Soose RJ, Maurer JT, de Vries N, Cornelius J, Froymovich O, Hanson RD, Padhya TA, Steward DL, Gillespie MB, Woodson BT, Van de Heyning PH, Goetting MG, Vanderveken OM, Feldman N, Knaack L, Strohl KP, STAR Trial Group (2014) Upper-airway stimulation for obstructive sleep apnea. N Engl J Med 370:139–149

Sullivan CE, Issa FG, Berthon-Jones M, Eves L (1981) Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. Lancet 1(8225):862–865

Walters AS, LeBrocq C, Dhar A et al (2003) Validation of the international restless legs syndrome study group rating scale for restless legs syndrome. Sleep Med 4:121–132

Wirth KJ, Steinmeyer K, Ruetten H (2013) Sensitization of upper airway mechanoreceptors as a new pharmacologic principle to treat obstructive sleep apnea: investigations with AVE0118 in anesthetized pigs. Sleep 36:699–708

Yalamanchali S, Farajian V, Hamilton C, Pott TR, Samuelson CG, Friedman M (2013) Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. JAMA Otolaryngol Head Neck Surg 139:1343–1350

Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S (1993) The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 328:1230–1235

Young T, Peppard P, Palta M, Hla KM, Finn L, Morgan B, Skatrud J (1997) Population-based study of sleep-disordered breathing as a risk factor for hypertension. Arch Intern Med 157:1746–1752

Pharmacodynamic Evaluation: Diabetic Methodologies

10

Juergen Sandow

Contents

J. Sandow (✉)
Centre of Pharmacology, Frankfurt-Main University,
Glashuetten, Germany
e-mail: j-sandow@t-online.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_33

Abstract

Here, the new aspects of diabetes methodology will be reviewed with reference to injection therapy, using analog insulins, GLP1 agonists (incretin mimetics), and the coformulation of incretin mimetics and insulins in the same solution. The options of diabetes therapy with sensor-augmented insulin pumps and the clinical pharmacology of bihormonal insulin pumps dispensing insulin and glucagon as required will be presented (bionic pumps). The specific aspect of advanced insulin pump therapy is the use of algorithms which control the insulin infusion rate and initiate shutdown of insulin infusion when impending hypoglycemia is detected. Bionic pumps (bihormonal) may also provide a stable glucagon solution for fine-tuning of glucose regulation and emergency release, in case of hypoglycemia developing rapidly or inadvertently, in particular during sleep at night time.

Introduction

Here, the new aspects of diabetes methodology will be reviewed with reference to injection therapy, using analog insulins, GLP1 agonists (incretin mimetics), and the coformulation of incretin mimetics and insulins in the same solution. The options of diabetes therapy with sensor-augmented insulin pumps and the clinical pharmacology of bihormonal insulin pumps dispensing insulin and glucagon as required will be presented (bionic pumps). The specific aspect of advanced insulin pump therapy is the use of algorithms which control the insulin infusion rate and initiate shutdown of insulin infusion when impending hypoglycemia is detected. Bionic pumps (bihormonal) may also provide a stable glucagon solution for fine tuning of glucose regulation and emergency release, in case of hypoglycemia developing rapidly or inadvertently, in particular during sleep at night time.

The insulin therapy in type I diabetic patients is mandatory, and early insulin therapy is now extended to the combination with incretin mimetics. This will decrease the insulin dose, avoid hypoglycemia, and reduce body weight reduction when possible. In advanced type II diabetic patients, early insulin therapy is recommended for efficacy during the initial period of restoring glucose control. Injections may be more acceptable and patient friendly when using the combination of incretin mimetics plus insulin. When required for efficacy, this injection therapy is supported by orally active antidiabetic drugs (OADs).

Clearly, at advanced stages of type II diabetes, injection therapy is mandatory. It is essential to achieve metabolic control within a clinically acceptable timeframe (target 6 months) and avoid therapeutic inertia. Failure to proceed from prolonged use of orally active antidiabetic drugs (OADs) to injection therapy – at least for the urgent initial correction of glucose control – is often addressed as therapeutic inertia (Harris et al. 2010; Khunti et al. 2013; Khunti and Millar-Jones 2017). The long delay when relying on the procedure of diet and change of lifestyle is harmful for the patient. Learning how to cope with diabetes in an effective manner is an important aspect (patient education). To implement early and effective support by pharmacotherapy is essential (The DCCT Research Group 1993), in particular to start injection therapy when HbA1c is of the order of 8.0 or above. In the previous edition (Becker 2011), clinical pharmacology of insulin preparations was discussed extensively, including attempts to find therapy other than by injection (e.g., insulin therapy by inhalation). Here, we will focus on efficient injection therapy and state-of-the-art insulin pumps.

There has been a very significant contribution of clinical pharmacology studies to rapid and highly effective exploration of insulin pens (coformulation of insulin and GLP1 agonists) and sensor-controlled insulin pumps. Glucose concentrations are controlled continuously (CSGM), replacing glucometers and repeated daily finger sticks by glucose sensors for immediate access and storage of 24-h glucose profiles. Insulin pumps are coupled to a sensor for automated adaptation to a target glucose range (hybrid-Aid systems), and "time in range" is an important parameter of system performance when developing new algorithms (Fig. 1).

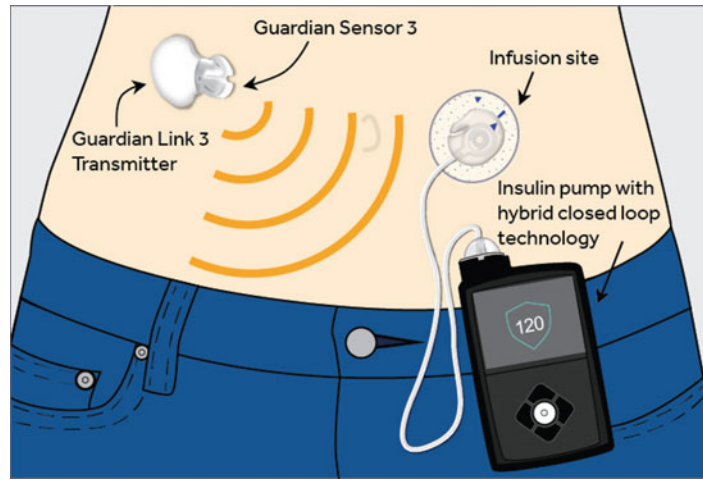


Fig. 1 State-of-the-art insulin pump controlled by sensor and guardian link (cablefree transmitter)

The interest of clinical pharmacology is moving from profiling of new antidiabetic compounds (preferably orally active) to patient-friendly injection therapy with glucose sensors and modern insulin pumps (Blak et al. 2012). In simple terms, "fear of injection" which was a major concern of diabetic patients before changing to insulin is no longer of practical relevance. Very extensive clinical pharmacology studies are required for the "intelligent pumps" with algorithms which adapt the insulin infusion rate to changes in monitored glucose concentrations. This is essential for automated dose adaptation. The patient is no longer under stress to perform the necessary changes at frequent intervals. The key concern of insulin-induced hypoglycemia (Briscoe and Davis 2006; Buckingham et al. 2008) is now under better control, in particular when bihormonal pumps become available (e.g., Ilet pump), which deliver fast-acting insulins and a stable glucagon solution at short intervals (mini bolus). Bihormonal pumps will provide solutions for improving dose adaptation in children with diabetes (Patterson et al. 2009) and in adults with rapidly changing insulin requirements, e.g., in sports and strenuous exercise. Automated shutdown of insulin pumps during impending hypoglycemia was an important step forward (e.g., Minimed 670G) enabled by algorithm control. In particular, at nighttime, it was helpful and reassuring, to have a monitoring system alerting parents and relatives ("mySentry") even when sleeping in an adjacent room close to the child. An alert function is now part of the Freestyle Libre II glucose sensor system.

The key aspect in type II diabetes is the change from basal analog insulin to GLP1 agonists (incretin mimetics, incretin analogs) for injection once or twice per day (QD), and formulations for injection once per week (QW). Treatment is initiated early in type II patients, at an advanced stage. This development was enabled by GLP1 agonists derived from natural gastrointestinal hormones (incretins), modified for prolonged action due to enzyme resistance. The long-term adherence to injection therapy is a very important characteristic (Asche et al. 2011) of the new GLP1 agonists (Chandran et al. 2015), and it represents the major improvement for advanced stages of type II diabetes. The key is to prevent progression of comorbidities associated closely with type I and type II diabetes, in particular cardiovascular comorbidity and end-stage renal disease with dependence on dialysis.

A further step forward is to reduce the dose of insulin required for effective and consistent metabolic control. This is achieved by combining injection of GLP1 agonists with analog insulins in the same solution for injection, a pharmaceutical development addressed as coformulation (Kalra and Gupta 2015). Current examples are iDeg-Lira and iGlar-Lixi frequently used in clinical medicine. Here is the most attractive potential for patients with type II diabetes at an advanced stage. The term "fear of

needle" is disappearing because insulin pens and similar pens for coformulation solutions use small and thin needles which are almost devoid of pain of injection and enable injection at different sites, e.g., arms or abdomen with the same efficiency. In simple terms, extended clinical pharmacology studies have become more acceptable to test persons and diabetic patients because of modern application devices (pain-free injection with longer intervals between injections, e.g., once per week). Practical lifetime of glucose sensors is extended from 3 days to 1 week or longer; there are now advanced sensors for subcutaneous implantation with a practical lifetime of 6 months (Senseonics, Kropff et al. 2017). The sensors provide an alarm in case of impending hypoglycemia (e.g., Freestyle Libre II), which is of particular relevance at nighttime.

Current insulin pumps are controlled by glucose sensors to detect the risk of impending hypoglycemia.

These sensor-augmented pumps (SAP) shut down for a pre-defined time period when hypoglycemia is detected (e.g., Minimed system) (Fig. 2).

Continuous subcutaneous glucose monitoring (CSGM) with sensors provides minute to minute data for the patient to be monitored, e.g., on the glucose meter or recorded continuously on a smartphone (Pankowska et al. 2005). In clinical studies, digital recording and storage of profiles is essential for submission to FDA and EMA in product applications.

Fig. 2 Sensor-augmented pumps (SAP) shut down when hypoglycemia is detected (Minimed system)

Clinical Insulin Therapy

Within about 10 years, there has been a major development in the aspects of clinical insulin therapy. It is well understood that prolonged treatment with orally active antidiabetic agents may not provide the intended correction of glucose regulation in due time (therapeutic inertia); the associated diabetes-related comorbidities identified by cardiovascular risk factors and lipid profile may persist indefinitely. The frequently invoked "fear of injections "is no longer relevant because insulin pens and more advanced injection devices are patient friendly and easily accepted by patients. Development and progression of comorbidities is under excellent control, provided that the decision for early injection therapy is implemented early on. Adherence to long-term injection therapy is readily established due to once per week (QW) coformulations of GLP1 agonists (Barnett 2013).

This review will therefore focus on early injection therapy in type II diabetes patients and on improvements for injection therapy of type I diabetes patients. Correction of glucose profile should be achieved whenever possible within 6 months after starting treatment, preferably with insulins when HbA1c is >8.0 and rising (American Diabetes Association ADA – Standards of Care 2019, European guidelines and related guidelines of International Diabetes Federation IDF). The methods applied in clinical pharmacology have adapted accordingly from characterizing the pharmacokinetic-pharmacodynamic profile of a new compound (PK-PD) toward phase 2 and extended phase 3 trials. Very often there is the need for a comparative design, including frequently used products as comparators (ADA 2019). This development follows the requirements of regulatory agencies (FDA, EMA) to submit comparative trials.

The need for early insulin therapy in type I diabetes, in particular in children and adolescents, is well understood. Clinical efficacy is achieved

with long-acting and short-acting analog insulins for basal insulin support and prandial insulin control, in a similar manner as for biosynthetic human insulin (Sandow et al. 2015; Landgraf and Sandow 2016; Mbanya et al. 2017). Glucose monitoring by parents and family members is much improved by the change to continuous subcutaneous glucose monitoring (CSGM) as an essential element. Clinical pharmacology has focused on the evaluation of suitable devices (sensor-augmented insulin pumps and bihormonal insulin pumps delivering insulin and stable glucagon solution as required) rather than on extending the range of analog insulin products. In particular attempts to replace injections by, e.g., insulin inhalation, were not successful and may have limited relevance in the future. Highly effective but very costly solutions like intraperitoneal insulin pumps with filling by subcutaneous ports are available, but cost and complexity is prohibitive. Most suitable and effective devices for children are insulin pumps, supported by CSGM sensors (sensor-augmented), and automatic shutdown of insulin pumps is enabled when glucose is decreasing rapidly (e.g., Minimed 630G and 670G).

The critical factor in clinical safety is hypoglycemia, for instance, under stress and increased glucose requirements in sports, particularly hypoglycemia when it remains unrecognized at nighttime (Briscoe and Davis 2006; Buckingham et al. 2008; Lipska et al. 2017).

The critical limitation in patients with type II diabetes is fear of injection which often delays the decision for therapy, at an advanced stage when only insulin therapy is effective to establish and maintain long-term glucose control. Once started, insulin therapy may be combined with orally active antidiabetic drugs (OAD combination therapy), or insulin may be injected together with a GLP1 agonist (insulin plus incretin mimetics as a coformulation).

A particular problem for clinical pharmacology studies is the selection of patients at an advanced state of obesity, where results of clinical pharmacology in obesity may differ markedly from those with regular body weight and composition (Vilsbøll et al. 2012). Comorbidities are addressed in the study design, e.g., by selection of patients with cardiovascular disease and increased risk factors (Mannucci and Monami 2017; Dalsgaard et al. 2017), patients with impaired kidney function (Mann et al. 2017), and other groups with diabetes-associated comorbidities (e.g., diabetic retinopathy).

The current state of clinical pharmacology for diabetes therapy with injectable products has changed markedly during the period from 2005 to 2019. The change was initiated by insulin pumps with controlled subcutaneous delivery, using fast-acting insulins to enhanced subcutaneous absorption (Haidar et al. 2013). The most critical problem of insulin-induced hypoglycemia is in part resolved by automatic suspension of pump action when algorithms detect a fast decrease in glucose concentrations and impending hypoglycemia (Hovorka et al. 2010; Kumareswaran et al. 2014). Clinical pharmacology studies are directed toward algorithm controlled sensor-augmented insulin pumps. Clearly, there are essential aspects of patient-friendly therapy, avoiding the dangers of hypoglycemia, providing devices for subcutaneous insulin administration based on highly efficient algorithms for dosage control, at the same time offering a patient-friendly device and long-term patient satisfaction (Elleri et al. 2013; Luijf et al. 2013; Fig. 3).

This development critically depends on improving sensors for glucose concentrations (Leelarathna et al. 2014; Tauschmann and Hovorka 2017).

Fig. 3 Small insulin pump (t-slim) with Dexcom sensor and smartphone monitoring

Fig. 4 Freestyle Libre sensor with reading device, 14 days use period. Hypoglycemia alert available

Sensors are now developed to an amazing level of precision and – at the same time – patient convenience. The route of administration for insulin remains subcutaneous injection or infusion. Intraperitoneal administration requires devices which are very expensive and difficult to handle. Patient acceptance has become the key word of new developments. The starting position is generally a glucose sensor which operates for 2 weeks (e.g., Freestyle Libre I or II) and removes the need for multiple daily finger pricks previously required for effective fine tuning of insulin dosage (Fig. 4).

For basal therapy, there is a wide selection of basal insulins to provide 24-h coverage with the single injection and increased insulin concentrations for patients with significant clinical obesity. In general terms, access to insulin in industrialized countries represents the challenge of selecting wisely from a wide range of options. Pricing of insulins and coverage by insurance is a general problem not addressed here. Health economics have clearly established the benefits of early insulin therapy, in particular confirming the durability of effects obtained and the efficient prevention of progression of diabetes-associated comorbidities. The need for hospital admissions due to insulin-related complications is now markedly reduced.

In terms of clinical insulin therapy, the development of biosimilar insulins has contributed to methods of biosynthesis but not to methods of clinical pharmacology evaluation, except that antigenicity has become a more important characteristic. Studies on the antigenicity of biosimilar glargine products are required even for a similar product from different production sites (North America or Europe). The clinical pharmacology of recombinant human insulin products has been reviewed in detail; they are characterized by clamp studies and in a few instances by direct studies for their PK-PD profiles in healthy subjects (Frank and Chance 1993; Sandow et al. 2015; Landgraf and Sandow 2016). For initial insulin substitution, basal analog insulins are preferred for convenience of handling and superior time-action-profile. Insulin degludec (Heise et al. 2012; Haahr and Heise 2014; Marso et al. 2017) was selected for the first coformulation product in solution together with the GLP1 agonist, liraglutide (Neumiller and Campbell 2009). As an example for the clinical benefit of coformulations, recent studies with iDegLira have confirmed the relevance of this approach in patients with type II diabetes (Buse et al. 2009, 2010, 2013, 2014). The dose of insulin may be reduced in coformulation because of the synergistic action of the GLP1 agonist on basal and prandial glucose secretion, the risk of clinically significant hypoglycemia is reduced, and weight reduction may be established as an additional benefit of the new therapeutic approach (depending on the GLP1 agonist selected in the coformulation).

Simple Insulin Pumps

Here is a different approach, how to adapt insulin dosing at short intervals to the rapidly changing glucose requirements in particular of younger people, in sports, in competition of athletes, and to the challenges of modern business life with an

Fig. 5 OmniPod patch pump with reader and monitoring device, *wearing of the patch pump on arm or leg permissible for convenience*

increasing stress load and frequent travelling. Simple pumps deliver a steady infusion rate and were first developed for pediatric indications (Maniatis et al. 2001; Ahern et al. 2002; Willi et al. 2003; Weinzimer et al. 2008). They require manual adjustment of the insulin delivery rate, their effect on glucose is monitored by the conventional procedure of frequent finger sticks. More convenient glucose sensors for continuous subcutaneous glucose monitoring (CSGM) offer a reading of the actual subcutaneous glucose concentrations, e.g., on a smartphone or dedicated glucose meter.

The current generation of insulin pumps (sensor-augmented) is using glucose sensors to detect the risk of impending and developing hypoglycemia. When signals indicate a rapidly decreasing plasma glucose concentration, the pump shuts down for a pre-defined time period (Minimed 670G). Continuous subcutaneous glucose monitoring (CSGM) with sensors provides minute to minute data for the patient, data to be monitored, e.g., on a glucose meter or view continuously on a smartphone. This procedure also initiates recording of glucose profiles in clinical studies.

A specific development has been the technology of patch pumps, which are very small and inconspicuous (Anhalt and Bohannon 2010; Buckingham et al. 2018); they require a minimum of tubing. Access to subcutaneous blood glucose data is directly from the pump, and electronic data transfer to the specific glucometer is very convenient (OmniPod Hybrid Closed-Loop System). Patch pumps are a welcome option for young people because they are inconspicuous and provide reliable insulin support with an acceptable amount of initial training (Fig. 5).

Glucose Sensors

Glucose sensors are currently developed to the stage of being active for 6 months after subcutaneous insertion of the suitable small device (Kropff et al. 2017). This requires currently a brief surgical procedure, which is well accepted by diabetic patients.

At a practical level, subcutaneous glucose sensors have been available and very efficiently used for about 20 years. They detect dangers of hypoglycemia, control the rate of devices for subcutaneous insulin administration and provide input for highly efficient algorithms of dosage control. Early warning when hypoglycemia is developing has been the critical and important component in diabetes methodology, in particular during nighttime in children, when devices provide an early alert. A major step forward was the mySentry™ system coupled to an insulin pump (Kaiserman et al. 2013, Cengiz 2013). This system provided a warning for parents and relatives even when children were sleeping in a separate room (MiniMed Paradigm REAL-Time system). Progress of adapting insulin delivery critically depends on coupled sensors for glucose concentrations, because they achieve an amazing level of precision and markedly improve patient convenience (Thabit et al. 2015). This is addressed as sensor-augmented dosage control. The route of insulin administration will remain subcutaneous injection

or infusion whenever the effective rate control of pumps is required. Intraperitoneal insulin pumps are very expensive and difficult to handle. The concept of an "artificial pancreas" has been fully implemented with subcutaneous infusion and fast-acting insulins.

When considering patient acceptance as the key element of new developments, the starting position for improving insulin therapy is to replace or supplement conventional glucose meter by a glucose sensor (Koschinsky and Heinemann 2001) which operates for 2 weeks (e.g., Freestyle Sensor, Weinstein et al. 2007; Garg et al. 2009). This provides immediate information of glucose concentrations (e.g., Freestyle Libre sensor) and enhances effective fine tuning of insulin dosage by the patient. Long-term glucose sensors are currently developed up to six months use (Senseonics); they require subcutaneous insertion of a suitable small device (Kropff et al. 2017, Senseonics).

At a practical level, subcutaneous glucose sensors are now very efficiently coupled to insulin pumps with an ever-increasing convenience and reliability (Weinstein et al. 2005; Wilson et al. 2007; Wolpert 2008; O'Connell et al. 2009). With algorithms for subcutaneous insulin dosing, modern pumps apply fast-acting insulins because their insulin absorption profile provides an option for reliable 24-h control, in particular for children with diabetes. Diabetes methodology is concentrating on well-tuned controlled delivery devices. The cost will remain limiting for years to come for reasons of price and lack of reimbursement; extensive clinical studies are required to confirm the long-term advantage and safety for diabetes patients at risk. The most advanced device using aspart insulin and in the second pump dasiglucagon is in the process of applying for an FDA product license (Ilet pump, developed by the group of Damiano based on experience with children and adults, Castle et al. 2019).

Briefly, insulin pumps with sensor control are being evaluated in extensive clinical studies for dose adaptation and have been widely introduced into clinical medicine. One overreaching aspect is early detection of hypoglycemia at night and during exercise, the most critical aspect of long-term efficient insulin therapy.

Insulin Pumps with Glucose Sensors

Glucose sensors are coupled to insulin pumps with an ever-increasing convenience and reliability to control insulin dosing (Tamborlane et al. 2008; Bailey et al. 2009; Garg et al. 2009; Aye et al. 2010; Bergenstal et al. 2010; Banerji and Dunn 2013; Castle et al. 2019). Sensor-augmented pumps (SAP) apply fast-acting insulin, e.g., aspart insulin, because the absorption profile is well characterized and provides an option for reliable 24-h control, in particular for children with diabetes (Boyne et al. 2003; Plotnick et al. 2003; Sulli and Shashaj 2003; Swan et al. 2008, 2009; Hirsch et al. 2008a, b; Kordonouri et al. 2010; JDRF 2009). Diabetes methodology is now concentrating to a high extent on well-tuned delivery devices even though the cost will remain limiting for years to come (JDRF 2010). Unfortunately, sensor-augmented pumps (SAP) may not be available in developing countries for many years for reasons of cost and lack of reimbursement (access to insulin has been widely discussed; there is now the issue of access to insulin pumps).

Interestingly, clinical pharmacology studies with insulin pumps are easier to perform than the classical glucose clamps studies because the devices comprise subcutaneous glucose sensors which immediately register and show the 24-h glucose profiles. At the time of using the pump, the sensors provide a record of the 24-h profile when suitable software is implemented. The sensors also provide hypoglycemia alert and up-to-date online information for the user (e.g., Freestyle Libre II) when required at short intervals, e.g., during exercise (Fig. 6).

Pumps controlled by coupled glucose sensors detect the risk of impending and developing hypoglycemia. A rapidly decreasing glucose signal will shut down the pump for a pre-defined time period. Such insulin pumps have been available for some time (Minimed 670G). The glucose sensors for continuous subcutaneous glucose monitoring (CSGM) provide detail for the patient to be viewed immediately on the glucose meter and when coupled to a smartphone (Mastrototaro et al. 2008). They were found particularly effective in patients with type I diabetes and are now

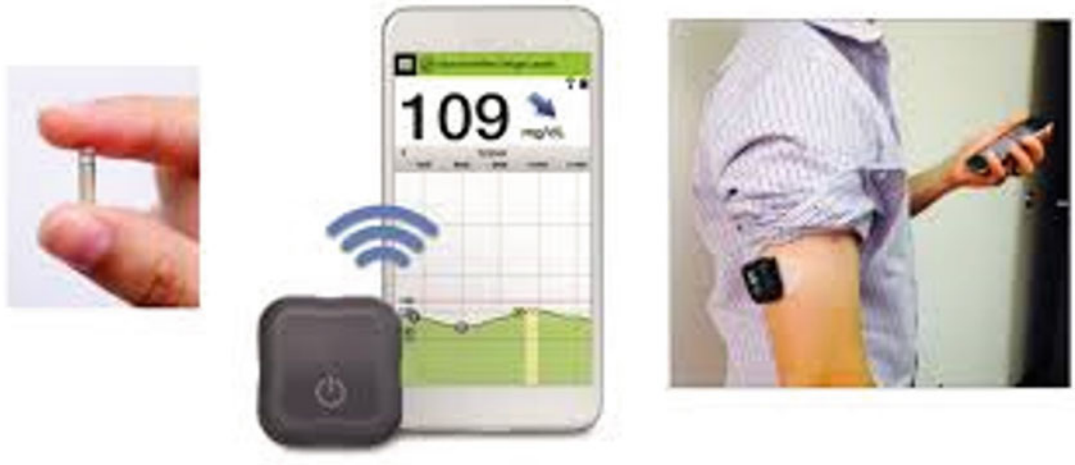


Fig. 6 Long-term glucose sensor with transmitter and smartphone control *(*Senseonic*)*

increasingly also applied in type II diabetes, to provide flexibility of dosage adjustment (Bailey et al. 2009).

Clinical Studies with Bihormonal Insulin Pumps

The bihormonal pumps (aka bionic pumps) deliver insulin by controlled infusion and – as a security process – dispense glucagon from a second pump in the case of impending hypoglycemia, which is detected by an algorithm (risk management) and immediately communicated to the patient by an alert signal. Bionic pumps establish a high level of precision, they adapt the insulin infusion rate due to glucose monitoring, and they keep the serum glucose within the pre-defined range. The critical measurement of the efficiency is addressed as "time in range."

On a higher level of precision, bionic pumps continuously adapt the effect of the insulin infusion by small doses of glucagon, and they deliver a bolus dose of glucagon in case of impending hypoglycemia (emergency protocol). Specific studies have been performed about the effect of exercise on glucose control, to confirm that algorithms will adapt the dispensing of glucagon during an acute decrease of plasma glucose caused by targeted exercise and by stressful sports. Technical issues remain to be resolved, mainly the availability of stable glucagon solutions, e.g., dasiglucagon instead of GlucaGen solution which needs to be prepared freshly every day in a clinical trial. There are two products of stable glucagon solution which will be available shortly after FDA approval. An emergency medication for future use is being developed for nasal application of glucagon powder (Basquimi™).

Early clinical studies were performed with two separate t-slim pumps, one for a fast-acting insulin (usually insulin aspart) and the other for the glucagon solution freshly prepared every day. For a very detailed review of the initial studies, refer to El-Khatib et al. (2014). Early studies in adult patients were followed by comparable studies in children. In particular, adaptation of the system to exercise-induced acute changes in insulin secretion and serum glucose is important. This was followed by extended field studies (Castle et al. 2018; El-Khatib et al. 2017).

Glucagon (Emergency and Bionic Pumps)

There has been considerable interest in a stable glucagon solution for injection in emergency situations, e.g., diabetic patients with loss of consciousness in particular at night, due to profound hypoglycemia. This is often associated with cramps and convulsions that may resemble neurologic

Fig. 7 Hypoglycemia emergency set of glucagon for immediate reconstitution

Fig. 8 Bihormonal pump with two separate syringes for dispensing insulin and glucagon. Controlled by a CSGM sensor, data are transmitted to the pump and processed by an algorithm to keep glucose concentration within a predetermined range. Glucagon is dispensed when impending hypoglycemia is recognized by the algorithm

disease (e.g., epilepsy). Currently available emergency medication is a lyophilized glucagon preparation (GlucaGen), which needs to be reconstituted before injection by family members (Fig. 7).

In clinical pharmacology studies, the glucagon solution was prepared freshly every day immediately before application in bihormonal insulin pumps. There is a need for a stable glucagon solution which maintains its activity for several weeks, under suitable storage conditions (preferably at room temperature, when used in a bihormonal pump). The clinical pharmacology studies were performed with insulin-induced hypoglycemia in the presence of rescue medication, by glucose infusion. The initial configuration for early clinical pharmacology studies was by using two separate pumps, e.g., t-slim pumps (Cengiz et al. 2011; Russell et al. 2016; El-Khatib et al. 2014, 2017; Castle et al. 2018).

Glucagon solutions prepared freshly every 24 h were carefully evaluated. The pharmaceutical characteristics of these solutions have been reported in detail (Wilson and Castle 2018). The glucagon analog dasiglucagon with clinical utility confirmed in phase 2 and 3 clinical trials (Hoevelmann et al. 2018) is currently applied as a stable solution in phase 3 trials. An advanced product is the first bihormonal pump system delivering both insulin and glucagon (Ilet pump) by an algorithm based on minute-to-minute data for subcutaneous glucose (Jacobs et al. 2014; Bakhtiani et al. 2014; Russell et al. 2016; El-Khatib et al. 2014, 2017; Castle et al. 2018). The Ilet pump system is in final evaluation by FDA for product licensing (June 2019) (Fig. 8).

GLP1 Agonists (Incretin Mimetics, Peptide Analogs)

Here is the next step of diabetes methodology. Pharmacodynamic evaluation has moved forward to characterize peptide hormones analogs (GLP1 agonists) which can be used in type II diabetes for injection therapy and – more conveniently – to be combined with basal analog insulin to reduce the risk of hypoglycemia and enhance weight reduction. The endogenous system (incretin hormones) was established and characterized in detail by extensive research. The critical characteristic of the GLP1 agonists is resistance to inactivation by endogenous enzymes (dipeptidyl-peptidase DPP). Modified DNA biosynthesis is required to obtain clinically relevant prolongation of half-life. Diabetes therapy operates with biosynthetic peptides derived from incretin hormones acting on the gastrointestinal tract in response to food ingestion (Drucker and Nauck 2006; Drucker et al. 2017; Mueller et al. 2017). The GLP1 agonists (incretins) are a new group of drugs for injection in type II diabetes (Zander et al. 2002). They are also applied in type I diabetes as an enhancement of insulin therapy, by preference after suitable development of coformulations with basal insulins, at a reduced dose of insulin. The critical advantage of such coformulations is reduction of the insulin dose and

low risk of hypoglycemia (Kalra et al. 2016; Kalra and Gupta 2016; Madsbad 2016). For GLP1 agonists, a clinically relevant prolongation of half-life has enabled the clinical use of depot formulations for injection once per week. Two of these coformulations have shown their clinical advantage by improving patient acceptance (dulaglutide and semaglutide). For diabetic patients, there are now regimens for injection once or twice per day injection (QD) and advanced regimens by injections once per week (QW). The group of GLP1 agonists is often addressed as "incretin mimetics." The response of insulin release and suppression of glucagon release to food ingestion is the characteristic difference when compared with the established insulin therapy. The clinically relevant products for therapy are short-acting GLP1 agonists for once or twice daily s.c. injection (e.g., exenatide, liraglutide, and lixisenatide QD), and the convenient group of GLP1 agonists for injection of 1 week (e.g., dulaglutide and semaglutide QW) . This group is of predominant clinical interest because patients prefer a regimen of infrequent injections supplemented by orally active antidiabetic drugs when required for enhanced efficacy (GLP1-A plus OAD). A very important specific aspect is that GLP1 agonists *are* used in clinical regimens which reduce the insulin dose and lower the risk of insulin-induced hypoglycemia. The current state of clinical studies with GLP1 agonists has been reviewed (Htike et al. 2017; Sharma et al. 2018) for the formulations for QD and QW injection (including exenatide, liraglutide, lixisenatide, dulaglutide and semaglutide; albiglutide was discontinued.). There is now an interesting and relevant selection of approved products for clinical use by patients and diabetes centers. Access to these products depends on country-specific conditions of availability and reimbursement (health economics); the advantages are clearly documented in several reviews (Banerji and Dunn 2013).

One characteristic requirement of the FDA and EMA has been for cardiovascular outcome and diabetes-associated comorbidities (Bethel et al. 2018; Stark Casagrande et al. 2013; Mann et al. 2017; Boyle et al. 2018). This development started with analog insulins and has been extended to the GLP1 agonists group (Marso et al. 2016; Hayward et al. 2015; Muskiet et al. 2018; Bahtiyar et al. 2018; Boyle et al. 2018; Bethel et al. 2018; Peterson and Barry 2018; Dalsgaard et al. 2017; Mannucci and Monami 2017; Pfeffer et al. 2015). The large range of clinical studies with GLP1 agonists may be summarized by differentiating single compounds for injection once per day (QD) or once per week (QW) and coformulations which contain the basal insulin together with a matched GLP1 agonist (iDegLira and iGlarLixi).

Exenatide

Exenatide is the protagonist of the new development in type II diabetes, to replace basal insulin therapy and reduce the risk of hypoglycemia. Exenatide is applied alone or supported by orally active medication (OADs) (Abdul-Ghani et al. 2015). Exenatide was approved in 2005. Early studies with daily injections were followed by development of once per week depot injection (Drucker et al. 2008; Best et al. 2009, Bydureon™). Clinical pharmacology studies supporting product development and subsequent very extensive clinical studies (Ahren 2011; Buse et al. 2009, 2010a, 2010b, 2011, 2013, 2014) are characterized by extending the range of observations to effects on cardiovascular tolerance and efficacy as well as diabetes-associated comorbidities – with a particular focus on weight reduction in diabetes-associated obesity. These studies follow requirements of regulatory agencies (FDA, EMA). The clinically relevant differences between established products are small, which is reassuring for countries where access is limited. Comparative trials mainly focus on promotional aspects, whereas the level of efficacy is generally of a very similar order, and specific safety issues remain to be monitored in patients with established risk factors.

Liraglutide

The clinical pharmacology of liraglutide has been reviewed in terms of pharmacokinetic profile and efficacy and safety when compared with, e.g., exenatide (Heise et al. 2012; Haahr and Heise

2014). Liraglutide was approved in 2009–2010. Clinical development of liraglutide is characterized by extensive comparative trials (Ahmann et al. 2015; Alatorre et al. 2017; Marso et al. 2016), the comparative clinical studies included important GLP1 agonists (exenatide, lixisenatide) and has prepared the field for coformulation with basal analog insulins (Anderson and Trujillo 2016). The injection once per day was a significant advantage (Victoza™). An interesting characteristic of liraglutide was the focus on weight reduction as a specific indication separate from the insulin field (Saxenda™). Saxenda may be used in patients with clinically significant obesity in the absence of diabetes. The clinical profile of liraglutide was a significant advantage for development of a coformulation with insulin degludec (iDegLira); this coformulation has been leading the field for some time (Gough et al. 2015).

Lixisenatide

In a similar manner as with liraglutide, the clinical advantages of lixisenatide (Barnett 2013; Werner 2014) were established and confirmed in comparative trials with established products (Rosenstock et al. 2016a, 2016b, 2016c; Pfeffer et al. 2015; Davies et al. 2017). Lixisenatide was approved in Europe in 2013 and in the United States in 2016 (Okere et al. 2018).

The results initiated the coformulation of insulin glargine and lixisenatide. The clinical role of such core formulations is outstanding (Kalra and Gupta 2016; Kalra et al. 2016); there are currently the best application of GLP1 agonist-related peptides (Drucker et al. 2017). The established characteristics are more effective lowering of HbA1c during the test period of 3 months of 6-month, clinical safety with regard to reduced incidence of confirmed hypoglycemia events, and acceptable gastrointestinal tolerance with decreasing symptoms nausea within the first 4 to 8 weeks of application. As with comparable formulations for daily injection, local tolerance and absence of significant local reactions at the injection site is well established (Fonseca et al. 2012) and a wide range for selection of the injection site is well documented (e.g. upper arm, leg, abdominal region). Treatment is initiated with a reduced daily dose and up titrated to the final dose depot injections as a general principle for formulations with initial gastrointestinal effects which limit tolerance.

Dulaglutide

Dulaglutide currently provides the most convenient once per week depot injection (QW) (Dungan et al. 2014; Kuritzky 2014). Dulaglutide was approved 2014. The critically important characteristic of using once per week (QW) injection has markedly improved adherence to therapy; the treatment is much more readily accepted by the patient (Alatorre et al. 2017) and maintained for prolonged time periods. This has also initiated a clinical development where dulaglutide is selected as the first treatment option of type II patients when their HbA1c control needs to be normalized at an early time, to prevent or reduce diabetic comorbidities. In this respect, patient education has become a very important part of injection therapy; many patients need to go on about the importance of preventing progress of diabetes-associated comorbidities. Interestingly, this is not a question of simply lowering the HbA1c to the normal range, but there have been many specific investigations about diabetes-associated comorbidities, early recognition of prediabetes and e.g. early treatment of lipid disorders (Fig. 9).

Semaglutide

Here is currently the next GLP1 agonist for injection once per week, with an extended half-life confirmed by dose escalation pharmacokinetic studies (Kapitza et al. 2017; Sorli et al. 2017). In line with other new antidiabetic medications for injection (FDA guideline), extensive studies on cardiovascular outcome and effects on kidney function were performed (Marso et al. 2016; Monami et al. 2017; Dicembrini et al. 2017;

Fig. 9 Dulagutide is a GLP1 agonist for subcutaneous injection once per week. Treatment is started with 0.75 mg/ml to adapt for effects on gastrointestinal function (tolerance) and subsequently continued with the full dose of 1.5 mg/ml. The important clinical advantage may be reduction of body weight

Abdul-Ghani et al. 2017; Jensen et al. 2018; Home 2019).

Semaglutide was approved on 2017 for QW injection (Davies et al. 2017). It is the first instance in this group of peptides subsequently developed as a clinical formulation and approved in 2019 for QD oral administration (Granhall et al. 2019)

Insulins and GLP1 Agonists (Coformulations)

Clinical pharmacology of pharmaceutical formulations became an important aspect in the improving treatment of type II diabetes, development started with premixed insulins (which were more difficult to handle) followed by coformulation of two analog insulins in a single solution, which provides effective glycemic coverage with a minimum of daily injections (Diamant et al. 2014; Alatorre et al. 2017; Frias et al. 2018). Studies were performed with basal insulin degludec and rapid-acting insulin aspart (degludec/aspart, aka iDegAsp) for injection in a single syringe (Kalra 2014; Kalra and Gupta 2015). This coformulation followed the concept of premixed insulins; it was easier to handle, but the risk of hypoglycemia when increasing dosage was not reduced.

That situation improved considerably when coformulations became available for type II in a single pen for injection with the solution for injection containing two soluble peptide hormones at a constant ratio of analog insulin and GLP1 agonist (Kalra and Gupta 2016; Kalra et al. 2016). Clinical pharmacology studies were performed with the coformulations of insulin degludec – liraglutide (iDegLira) and insulin glargine – and lixisenatide (iGlarLixi) (Gough et al. 2015; Billings et al. 2018; Vilsboll et al. 2016; Rosenstock et al. 2016a, b). The coformulations are now well established in therapy of type II diabetes (Nuffer et al. 2018, Valentine et al. 2017; Wysham et al. 2018).

Summary and Outlook

Within about 10 years, early therapy of diabetes mellitus type I and type II has improved markedly by injection of coformulations including a GLP1 analog and by the use of algorithm-controlled insulin pumps. The change in therapy and acceptance of more complex technology is due to recognition of the global epidemic that diabetes mellitus has reached. There is now an understanding of advanced early diagnosis and the need for screening. The options for early and effective therapy are particularly obvious for type II diabetes. Diabetes methodology in clinical pharmacology has changed from the assessment of insulin products (analog insulins with fast absorption and long-acting basal insulins) to the clinical assessment of GLP1 agonists (incretin mimetics) and coformulations with basal insulin analogs. Research on diabetes-associated comorbidities has established the potential for prevention and retardation when fully efficient early therapy is initiated (ADA 2018) and maintained (Leite et al. 2014; Merger et al. 2016; Petrosyan et al. 2017; Lehrke and Marx 2017; Wielgosz et al. 2018; Franch Nadal et al. 2017; An et al. 2019; Richter et al. 2018). The approach by diet and

lifestyle is clearly not sufficient in advanced diabetes I and II; early injection therapy may be the future option for reducing comorbidities in a patient-friendly and acceptable procedure.

Progress in the field of insulin pumps within about 10 years is very impressive due to advanced technology, even though access to modern insulin pumps is limited due to cost, and the technical complexity of long-term therapy. Patch pumps have advanced markedly (Omnipod™); they may remain a small albeit useful segment due to restrictions of intensive handling. The critical step forward is introduction of bihormonal (bionic) pumps in patients who are at particular risk or need intensive therapy without the complication of hypoglycemia. Stable glucagon solutions will be available in the near future (dasiglucagon), and it is remarkable that a nasal glucagon single dose preparation for emergency is now available (Baqsimi™). Proof of concept for the bionic insulin pumps is impressive due to recognition and adaptation of exercise induced changes in glucose control which have previously been a problem in particular for adaptation to sports and intensive exercise. Final FDA approval for such pumps is well advanced (Ilet pump).

In summary, the most impressive progress has been made with injectable coformulations which may help to overcome therapeutic inertia early after diagnosis of advanced stages of type II diabetes.

References and Further Reading

Abdul-Ghani MA et al (2015) Renal sodium-glucose cotransporter inhibition in the management of type 2 diabetes mellitus. Am J Physiol Renal Physiol 309(11): F889–F900

Abdul-Ghani M, DeFronzo RA, Del Prato S et al (2017) Cardiovascular disease and type 2 diabetes: has the dawn of a new era arrived? Diabetes Care 40 (7):813–820

ADA (2019) American Diabetes Association, Standards of Care 2019. https://professional.diabetes.org/content-page/practice-guidelines-resources

Ahern JA, Boland EA, Doane R et al (2002) Insulin pump therapy in pediatrics: a therapeutic alternative to safely lower HbA1c levels across all age groups. Pediatr Diabetes 3(1):10–15

Ahmann AJ, Rodbard HW, Rosenstock J et al (2014) Efficacy and safety of liraglutide vs. placebo when added to basal insulin analogs in patients with type 2 diabetes (LIRAADD2BASAL). Diabetes 63:A87

Ahren B (2011) The future of incretin-based therapy: novel avenues – novel targets. Diabetes Obes Metab 13:158–166

Alatorre C, Fernández Landó L, Yu M et al (2017) Treatment patterns in patients with type 2 diabetes mellitus treated with glucagon-like peptide-1 receptor agonists: higher adherence and persistence with dulaglutide compared with once-weekly exenatide and liraglutide. Diabetes Obes Metab 19(7):953–961

American Diabetes Association (2018) 3. Comprehensive medical evaluation and assessment of comorbidities: Standards of Medical Care in Diabetes-2018 Diabetes Care 41(Suppl 1):S28–S37

An J, Le QA, Dang T (2019) Association between different types of comorbidity and disease burden in patients with diabetes. J Diabetes 11(1):65–74

Anhalt H, Bohannon NJ (2010) Insulin patch pumps: their development and future in closed-loop systems. Diabetes Technol Ther 12(Suppl 1):S51–S58

Asche C, LaFleur J, Conner C (2011) A review of diabetes treatment adherence and the association with clinical and economic outcomes. Clin Ther 33(1):74–109

Aye T, Block J, Buckingham B (2010) Toward closing the loop: an update on insulin pumps and continuous glucose monitoring systems. Endocrinol Metab Clin N Am 39(3):609–624

Bahtiyar G, Pujals-Kury J, Sacerdote A (2018) Cardiovascular effects of different GLP-1 receptor agonists in patients with type 2 diabetes. Curr Diab Rep 18 (10):92

Bailey T, Zisser H, Chang A (2009) New features and performance of a next-generation SEVEN-day continuous glucose monitoring system with short lag time. Diabetes Technol Ther 11(12):749–755

Banerji MA, Dunn JD (2013) Impact of glycemic control on healthcare resource utilization and costs of type 2 diabetes: current and future pharmacologic approaches to improving outcomes. Am Health Drug Benefits 6(7):382–392

Barnett AH (2013) Complementing insulin therapy to achieve glycemic control. Adv Ther 30(6):557–576

Becker HA (2011) Pharmacodynamic evaluation: diabetes methodologies. In: Vogel HG et al (eds) Drug discovery and evaluation: methods in clinical pharmacology. Springer, Berlin/Heidelberg, pp 457–481

Bergenstal RM, Tamborlane WV, Ahmann A, et al (2010) Effectiveness of sensor-augmented insulin pump therapy in type 1 diabetes

Best JH, Boye KS, Rubin RR et al (2009) Improved treatment satisfaction and weight-related quality of life with exenatide once weekly or twice daily. Diabet Med 26:722–728

Bethel MA, Patel RA, Merrill P et al (2018) Cardiovascular outcomes with glucagon-like peptide-1 receptor agonists in patients with type 2 diabetes: a meta-analysis. Lancet Diabetes Endocrinol 6(2):105–113

Billings LK, Doshi A, Gouet D et al (2018) Efficacy and safety of IDegLira versus basal-bolus insulin therapy in patients with type 2 diabetes uncontrolled on

metformin and basal insulin: the DUAL VII randomized clinical trial. Diabetes Care 41:1009–1016
Blak BT, Smith HT, Hards M et al (2012) Optimization of insulin therapy in patients with type 2 diabetes mellitus: beyond basal insulin. Diabet Med 29:e13–e20
Boyle JG, Livingstone R, Petrie JR (2018) Cardiovascular benefits of GLP-1 agonists in type 2 diabetes: a comparative review. Clin Sci (Lond) 132(15):1699–1709
Boyne MS, Silver DM, Kaplan J et al (2003) Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor. Diabetes 52(11):2790–2794
Briscoe VJ, Davis SN (2006) Hypoglycemia in type 1 and type 2 diabetes: physiology, pathophysiology, and management. Clin Diabetes 24:115–121
Buckingham WDM, Lecher T et al (2008) Duration of nocturnal hypoglycemia before seizures. Diabetes Care 31(11):2110–2112
Buckingham BA, Forlenza GP, Pinsker JE et al (2018) Safety and feasibility of the OmniPod hybrid closed-loop system in adult, adolescent, and pediatric patients with type 1 diabetes using a personalized model predictive control algorithm. Diabetes Technol Ther 20 (4):257–262
Buse JB, Rosenstock J, Sesti G et al (2009) Liraglutide once a day versus exenatide twice a day for type 2 diabetes: a 26-week randomised, parallel-group, multinational, open-label trial (LEAD-6). Lancet 374:39–47
Buse JB, Sesti G, Schmidt SE et al (2010) Switching to once-daily liraglutide from twice-daily exenatide further improves glycemic control in patients with type 2 diabetes using oral agents. Diabetes Care 33:1300–1303
Buse JB, Drucker DJ, Taylor KL, Kim T et al (2010b) DURATION-1 Study Group. DURATION-1: exenatide once weekly produces sustained glycemic control and weight loss over 52 weeks. Diabetes Care 33(6):1255–1261
Buse JB, Sesti G, Schmidt SE et al (2010c) Switching to once-daily liraglutide from twice-daily exenatide further improves glycemic control in patients with type 2 diabetes using oral agents. Diabetes Care 33:1300–1303
Buse JB, Nauck M, Forst T et al (2013) Exenatide once weekly versus liraglutide once daily in patients with type 2 diabetes (DURATION-6): a randomised, open-label study. Lancet 381:117–124
Buse JB, Vilsbøll T, Thurman J et al (2014) Contribution of liraglutide in the fixed-ratio combination of insulin degludec and liraglutide (IDegLira). Diabetes Care 37:2926–2933
Castle JR, Elander M (2019) Long-term safety and tolerability of dasiglucagon, a stable-in-solution glucagon analogue. Diabetes Technol Ther 21(2):94–96
Castle JR, Youssef JE, Wilson LM et al (2018) Randomized outpatient trial of single- and dual-hormone closed-loop systems that adapt to exercise using wearable sensors. Diabetes Care 41(7):1471–1477
Cengiz E (2013) Analysis of a remote system to closely monitor glycemia and insulin pump delivery – is this the beginning of a wireless transformation in diabetes management? J Diabetes Sci Technol 7(2):362–364
Cengiz E, Tamborlane WV (2009) A tale of two compartments: interstitial versus blood glucose monitoring. Diabetes Technol Ther 11(Suppl 1):S11–S16
Cengiz E, Sherr JL, Weinzimer SA et al (2011) New-generation diabetes management: glucose sensor-augmented insulin pump therapy. Expert Rev Med Devices 8(4):449–458
Chance RE, Frank BH (1993) Research, development, production, and safety of biosynthetic human insulin. Diabetes Care 16(Suppl 3):133–142
Chandran A, Bonafede MK, Nigam S et al (2015) Adherence to insulin pen therapy is associated with reduction in healthcare costs among patients with type 2 diabetes mellitus. Am Health Drug Benefits 8(3):148–158
Chase HP, Beck R, Tamborlane W et al (2005) A randomized multicenter trial comparing the GlucoWatch biographer with standard glucose monitoring in children with type 1 diabetes. Diabetes Care 28(5):1101–1106
Chase HP, Beck RW, Xing D et al (2007) Continuous glucose monitoring in youth with type 1 diabetes: 12-month follow-up of the Juvenile Diabetes Research Foundation continuous glucose monitoring randomized trial. Diabetes Technol Ther 12(7):507–515
Clarke WL, Cox D, Gonder-Frederick LA et al (1987) Evaluating clinical accuracy of systems for self-monitoring of blood glucose. Diabetes Care 10 (5):622–628
Dalsgaard NB, Brønden A, Vilsbøll T, Knop FK (2017) Cardiovascular safety and benefits of GLP-1 receptor agonists. Expert Opin Drug Saf 16(3):351–363
Davies M, Pieber TR, Hartoft-Nielsen ML, Hansen OKH, Jabbour S, Rosenstock J (2017) Effect of oral semaglutide compared with placebo and subcutaneous semaglutide on glycemic control in patients with type 2 diabetes: a randomized clinical trial. JAMA 318 (15):1460–1470
Diamant M, Nauck MA, Shaginian R, Malone JK et al (2014) Glucagon-like peptide 1 receptor agonist or bolus insulin with optimized basal insulin in type 2 diabetes. Diabetes Care 37(10):2763–2773
Dicembrini I, Nreu B, Scatena A, Andreozzi F, Sesti G, Mannucci E, Monami M (2017) Microvascular effects of glucagon-like peptide-1 receptor agonists in type 2 diabetes: a meta-analysis of randomized controlled trials. Acta Diabetol 54(10):933–941
Drucker DJ, Nauck MA (2006) The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. Lancet 368(9548):1696–1705
Drucker DJ, Buse JB, Taylor K et al (2008) Exenatide once weekly versus twice daily for the treatment of type 2 diabetes: a randomised, open-label, non-inferiority study. Lancet 372:1240–1250
Drucker DJ, Habener JF, Holst JJ (2017) Discovery, characterization, and clinical development of the glucagon-like peptides. J Clin Invest 127(12):4217–4227
El-Khatib FH, Russell SJ, Magyar KL, Sinha M, McKeon K, Nathan DM, Damiano ER (2014) Autonomous and continuous adaptation of a bihormonal bionic pancreas in adults and adolescents with type 1 diabetes. J Clin Endocrinol Metab 99(5):1701–1711

El-Khatib et al (2017) Home use of a bihormonal bionic pancreas versus insulin pump therapy in adults with type 1 diabetes: a multicentre randomised crossover trial. Lancet 389:P369–P380

Elleri D, Allen JM, Kumareswaran K et al (2013) Closed-loop basal insulin delivery over 36 hours in adolescents with type 1 diabetes: randomized clinical trial. Diabetes Care 36(4):838–844

Fonseca VA, Alvarado-Ruiz R, Raccah D et al (2012) Efficacy and safety of the once-daily GLP-1 receptor agonist lixisenatide in monotherapy: a randomized, double-blind, placebo-controlled trial in patients with type 2 diabetes (GetGoal-Mono). Diabetes Care 35:1225–1231

Franch Nadal J, Mata Cases M, Mauricio Puente D (2016) Epidemiology and clinical management of type 2 diabetes mellitus and associated comorbidities in Spain (e-Management study). Med Clin (Barc) 147(Suppl 1):1–7

Frias J, Puig Domingo M, Meneghini L, Napoli R, Liu M, Soltes Rak E, Aroda VR (2018) More patients reach glycaemic control with a fixed-ratio combination of insulin glargine and lixisenatide (iGlarLixi) than with basal insulin at 12 weeks of treatment: a post hoc time-to-control analysis of LixiLan-O and LixiLan-L. Diabetes Obes Metab 20(9):2314–2318

Garg SK, Smith J, Beatson C et al (2009) Comparison of accuracy and safety of the SEVEN and the navigator continuous glucose monitoring systems. Diabetes Technol Ther 11(2):65–72

Gough SC, Bode BW (2015) Woo VC, et al 2015 one-year efficacy and safety of a fixed combination of insulin degludec and liraglutide in patients with type 2 diabetes: results of a 26-week extension to a 26-week main trial. Diabetes Obes Metab 17(10):965–973. PMID: 25980900

Granhall C, Donsmark M, Blicher TM et al (2019) Safety and pharmacokinetics of single and multiple ascending doses of the novel oral human GLP-1 analogue, oral semaglutide, in healthy subjects and subjects with type 2 diabetes. Clin Pharmacokinet 58(6):781–791

Guja C, Miulescu R (2017) Semaglutide – the "new kid on the block" in the field of glucagon-like peptide-1 receptor agonists? Ann Transl Med 5(23):475–480

Haahr H, Heise T (2014) A review of the pharmacological properties of insulin degludec and their clinical relevance. Clin Pharmacokinet 53:787–800

Haidar A, Legault L, Dallaire M et al (2013) Glucose-responsive insulin and glucagon delivery (dual-hormone artificial pancreas) in adults with type 1 diabetes: a randomized crossover controlled trial. CMAJ 185 (4):297–305

Harris SB, Kapor J, Lank CN et al (2010) Clinical inertia in patients with T2DM requiring insulin in family practice. Can Fam Physician 56:e418–e424

Heise T, Nosek L, Bottcher SG et al (2012) Ultralong-acting insulin degludec has a flat and stable glucose-lowering effect in type 2 diabetes. Diabetes Obes Metab 14:944–950

Hirsch IB, Abelseth J, Bode BW et al (2008a) Sensor-augmented insulin pump therapy: results of the first randomized treat-to target study. Diabetes Technol Ther 10(5):377–383

Hirsch IB, Armstrong D, Bergenstal RM et al (2008b) Clinical application of emerging sensor technologies in diabetes management: consensus guidelines for continuous glucose monitoring (CGM). Diabetes Technol Ther 10(4):232–244

Hoevelmann U, Bysted BV, Mouritzen U, Macchi F, Lamers D, Kronshage B, Møller DV, Heise T (2018) Pharmacokinetic and pharmacodynamic characteristics of dasiglucagon, a novel soluble and stable glucagon analog. Diabetes Care 41(3):531–537

Home P (2019) Cardiovascular outcome trials of glucose-lowering medications: an update. Diabetologia 62 (3):357–369

Hovorka R, Allen JM, Elleri D et al (2010) Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. Lancet 375(9716):743–751

Htike ZZ, Zaccardi F, Papamargaritis D et al (2017) Efficacy and safety of glucagon-like peptide-1 receptor agonists in type 2 diabetes: a systematic review and mixed-treatment comparison analysis. Diabetes Obes Metab 19(4):524–536

Jacobs PG, El Youssef J, Castle J, Bakhtiani P, Branigan D, Breen M, Bauer D, Preiser N, Leonard G, Stonex T, Ward WK (2014) Automated control of an adaptive bihormonal, dual-sensor artificial pancreas and evaluation during inpatient studies. IEEE Trans Biomed Eng 61(10):2569–2581

Jacobs PG, El Youssef J, Reddy R et al (2016) Randomized trial of a dual-hormone artificial pancreas with dosing adjustment during exercise compared with no adjustment and sensor-augmented pump therapy. Diabetes Obes Metab 18(11):1110–1119

JDRF Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group (2009) The effect of continuous glucose monitoring in well-controlled type 1 diabetes. Diabetes Care 32(8):1378–1383

JDRF Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group (2010) Validation of measures of satisfaction with and impact of continuous and conventional glucose monitoring. Diabetes Technol Ther 12(9):679–684

Jensen L, Kupcova V, Arold G, Pettersson J, Hjerpsted JB (2018) Pharmacokinetics and tolerability of semaglutide in people with hepatic impairment. Diabetes Obes Metab 20(4):998–1005

Kaiserman K, Buckingham BA, Prakasam G, Gunville F, Slover RH, Wang Y, Nguyen X, Welsh JB (2013) Acceptability and utility of the mySentry remote glucose monitoring system. J Diabetes Sci Technol 7 (2):356–361

Kalra S, Gupta Y (2016) Endocrine and metabolic effects of glucagon like peptide 1 receptor agonists (GLP1RA). J Pak Med Assoc 66(3):357–359

Kalra S, Baruah MP, Sahay RK et al (2016) Glucagon-like peptide-1 receptor agonists in the treatment of type 2 diabetes: past, present, and future. Indian J Endocrinol Metab 20(2):254–267

Kapitza C, Dahl K, Jacobsen JB, Axelsen MB, Flint A (2017) Effects of semaglutide on beta cell function and glycaemic control in participants with type 2 diabetes: a randomised, double-blind, placebo-controlled trial. Diabetologia 60(8):1390–1399

Khunti K, Millar-Jones D (2017) Clinical inertia to insulin initiation and intensification in the UK: a focused literature review. Prim Care Diabetes 11:3–12

Khunti K, Wolden ML, Thorsted BL et al (2013) Clinical inertia in people with type 2 diabetes: a retrospective cohort study of more than 80,000 people. Diabetes Care 36:3411–3417

Kordonouri O, Pankowska E, Rami B et al (2010) Sensor-augmented pump therapy from the diagnosis of childhood type 1 diabetes: results of the Paediatric Onset Study (ONSET) after 12 months of treatment. Diabetologia 53(12):2487–2495

Koschinsky T, Heinemann L (2001) Sensors for glucose monitoring: technical and clinical aspects. Diabetes Metab Res Rev 17(2):113–123

Kropff J, Choudhary P, Neupane S et al (2017) Accuracy and longevity of an implantable continuous glucose sensor in the PRECISE study: a 180-day, prospective, multicenter, pivotal trial. Diabetes Care 40(1):63–68

Kumareswaran K, Thabit H, Leelarathna L et al (2014) Feasibility of closed-loop insulin delivery in type 2 diabetes: a randomized controlled study. Diabetes Care 37 (5):1198–1203

Landgraf W, Sandow J (2016) Equivalent recombinant human insulin preparations and their place in therapy. Eur Endocrinol 12(1):12–17

Leelarathna L et al (2014) Evaluating the performance of a novel embedded closed-loop system. J Diabetes Sci Technol 8(2):267–272

Lehrke M, Marx N (2017) Diabetes mellitus and heart failure. Am J Med 130(6S):S40–S50

Leite NC, Villela-Nogueira CA, Cardoso CR, Salles GF (2014) Non-alcoholic fatty liver disease and diabetes: from physiopathological interplay to diagnosis and treatment. World J Gastroenterol 20(26):8377–8392

Lipska KJ, Yao X, Herrin J et al (2017) Trends in drug utilization, glycemic control, and rates of severe hypoglycemia, 2006–2013. Diabetes Care 40:468–475

Luijf YM, DeVries JH, Zwinderman K et al (2013) Day and night closed-loop control in adults with type 1 diabetes: a comparison of two closed-loop algorithms driving continuous subcutaneous insulin infusion versus patient self-management. Diabetes Care 36 (12):3882–3887

Madsbad S (2016) Review of head-to-head comparisons of glucagon-like peptide-1 receptor agonists. Diabetes Obes Metab 18(4):317–332

Maniatis AK, Klingensmith GJ, Slover RH et al (2001) Continuous subcutaneous insulin infusion therapy for children and adolescents: an option for routine diabetes care. Pediatrics 107(2):351–356

Mann JFE, Ørsted DD, Buse JB (2017) Liraglutide and renal outcomes in type 2 diabetes. N Engl J Med 377 (22):2197–2198

Mannucci E, Monami M (2017) Cardiovascular safety of Incretin-based therapies in type 2 diabetes: systematic review of integrated analyses and randomized controlled trials. Adv Ther 34(1):1–40

Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, Lingvay I, Rosenstock J, Seufert J, Warren ML, Woo V, Hansen O, Holst AG, Pettersson J, Vilsbøll T, SUSTAIN-6 Investigators (2016) Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med 375(19): 1834–1844

Marso SP, McGuire DK, Zinman B et al (2017) Efficacy and safety of degludec versus glargine in type 2 diabetes. N Engl J Med 377:723–732

Mastrototaro J, Shin J, Marcus A et al (2008) The accuracy and efficacy of real-time continuous glucose monitoring sensor in patients with type 1 diabetes. Diabetes Technol Ther 10(5):385–390

Mbanya JC, Sandow J, Landgraf W, Owens DR (2017) Recombinant human insulin in global diabetes management – focus on clinical efficacy. Eur Endocrinol 13 (1):21–25

Merger SR, Kerner W, Stadler M, Zeyfang A, Jehle P, Müller-Korbsch M, Holl RW; DPV Initiative; German BMBF Competence Network Diabetes mellitus (2016) Prevalence and comorbidities of double diabetes. Diabetes Res Clin Pract 119:48–56

Monami M, Nreu B, Scatena A, Giannini S, Andreozzi F, Sesti G, Mannucci E (2017) Glucagon-like peptide-1 receptor agonists and atrial fibrillation: a systematic review and meta-analysis of randomised controlled trials. J Endocrinol Investig 40(11):1251–1258

Muskiet MHA, Tonneijck L, Huang Y et al (2018) Lixisenatide and renal outcomes in patients with type 2 diabetes and acute coronary syndrome: an exploratory analysis of the ELIXA randomised, placebo-controlled trial. Lancet Diabetes Endocrinol 6 (11):859–869

Neumiller JJ, Campbell RK (2009) Liraglutide: a once-daily incretin mimetic for the treatment of type 2 diabetes mellitus. Ann Pharmacother 43:1433–1444

Nuffer W, Guesnier A, Trujillo JM (2018) A review of the new GLP-1 receptor agonist/basal insulin fixed-ratio combination products. Ther Adv Endocrinol Metab 9 (3):69–79

O'Connell MA, Donath S, O'Neal DN et al (2009) Glycaemic impact of patient-led use of sensor-guided pump therapy in type 1 diabetes: a randomised controlled trial. Diabetologia 52(7):1250–1257

Okere AN, Montesdeoca J, Glasper A, Diaby V (2018) An evaluation of the clinical therapeutic effect of Lixisenatide in type 2 diabetes patients: a systematic literature review. Curr Diabetes Rev 14(4):363–375

Pankowska E, Skorka A, Szypowska A et al (2005) Memory of insulin pumps and their record as a source of information about insulin therapy in children and adolescents with type 1 diabetes. Diabetes Technol Ther 7 (2):308–314

Patterson CC, Dahlquist GG, Gyurus E et al (2009) Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–2020: a multicentre prospective registration study. Lancet 373(9680):2027–2033

Peterson SC, Barry AR (2018) Effect of glucagon-like peptide-1 receptor agonists on all-cause mortality and cardiovascular outcomes: a meta-analysis. Curr Diabetes Rev 14(3):273–279

Petrosyan Y, Bai YQ, Koné Pefoyo AJ, Gruneir A, Thavorn K, Maxwell CJ, Bronskill SE, Wodchis WP (2017) The relationship between diabetes care quality and diabetes-related hospitalizations and the modifying role of comorbidity. Can J Diabetes 41(1):17–25

Pfeffer MA, Claggett B, Diaz R et al (2015) Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. N Engl J Med 373(23):2247–2257

Plotnick LP, Clark LM, Brancati FL et al (2003) Safety and effectiveness of insulin pump therapy in children and adolescents with type 1 diabetes. Diabetes Care 26 (4):1142–1146. PMID: 29285508

Richter L, Freisinger E, Lüders F, Gebauer K, Meyborg M, Malyar NM (2018) Impact of diabetes type on treatment and outcome of patients with peripheral artery disease. Diab Vasc Dis Res 15(6):504–510

Rosenstock J, Aronson R, Grunberger G et al (2016a) Benefits of LixiLan, a titratable fixed-ratio combination of insulin glargine plus lixisenatide, versus insulin glargine and lixisenatide monocomponents in type 2 diabetes inadequately controlled on oral agents: the LixiLan-O randomized trial. Diabetes Care 39 (11):2026–2035

Rosenstock J, Diamant M, Aroda VR, Silvestre L et al (2016b) Efficacy and safety of LixiLan, a titratable fixed-ratio combination of lixisenatide and insulin glargine, versus insulin glargine in type 2 diabetes inadequately controlled on metformin monotherapy: the LixiLan proof-of-concept randomized trial. Diabetes Care 39(9):1579–1586

Sandow J, Landgraf W, Becker R, Seipke S (2015) Equivalent recombinant human insulin preparations and their place in therapy. Eur Endocrinol 11(1):10–16

Sharma D, Verma S, Vaidya S, Kalia K, Tiwari V (2018) Recent updates on GLP-1 agonists: current advancements & challenges. Biomed Pharmacother 108:952–962

Sorli C, Harashima SI, Tsoukas GM, Unger J, Karsbøl JD, Hansen T, Bain SC (2017) Efficacy and safety of once-weekly semaglutide monotherapy versus placebo in patients with type 2 diabetes (SUSTAIN 1): a double-blind, randomised, placebo-controlled, parallel-group, multinational, multicentre phase 3a trial. Lancet Diabetes Endocrinol 5(4):251–260

Stark Casagrande S, Fradkin JE, Saydah SH, Rust KF, Cowie CC (2013) The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988–2010. Diabetes Care 36(8):2271–2279

Sulli N, Shashaj B (2003) Continuous subcutaneous insulin infusion in children and adolescents with diabetes mellitus: decreased HbA1c with low risk of hypoglycemia. J Pediatr Endocrinol Metab 16(3):393–399

Swan KL, Weinzimer SA, Dziura JD et al (2008) Effect of puberty on the pharmacodynamic and pharmacokinetic properties of insulin pump therapy in youth with type 1 diabetes. Diabetes Care 31(1):44–46

Swan KL, Dziura JD, Steil GM et al (2009) Effect of age of infusion site and type of rapid-acting analog on pharmacodynamic parameters of insulin boluses in youth with type 1 diabetes receiving insulin pump therapy. Diabetes Care 32(2):240–244

Tamborlane WV, Beck RW, Bode BW et al (2008) Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med 359(14):1464–1476

Tauschmann M, Hovorka R (2017) Insulin delivery and nocturnal glucose control in children and adolescents with type 1 diabetes. Expert Opin Drug Deliv 14 (12):1367–1377

Thabit H, Leelarathna L, Wilinska ME et al (2015) Accuracy of continuous glucose monitoring during three closed-loop home studies under free-living conditions. Diabetes Technol Ther 17(11):801–807

The DCCT RESEARCH GROUP (1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 329(14): 977–986

Valentine V, Goldman J, Shubrook JH (2017) Rationale for, initiation and titration of the basal insulin/GLP-1RA fixed-ratio combination products, IDegLira and IGlarLixi, for the management of type 2 diabetes. Diabetes Ther 8(4):739–752

Vilsbøll T, Christensen M, Junker AE et al (2012) Effects of glucagon-like peptide-1 receptor agonists on weight loss: systematic review and meta-analyses of randomised controlled trials. BMJ 344:7771

Vilsbøll T, Vora J, Jarlov H, Kvist K, Blonde L (2016) Type 2 diabetes patients reach target glycemic control faster using IDegLira than either insulin degludec or liraglutide given alone. Clin Drug Investig 36 (4):293–303. PMID 26894800

Weinstein RL, Schwartz SL, Brazg RL et al (2007) Accuracy of the 5-day FreeStyle navigator continuous glucose monitoring system: comparison with frequent laboratory reference measurements. Diabetes Care 30 (5):1125–1130

Weinzimer SA, Steil GM, Swan KL et al (2008) Fully automated closed loop insulin delivery versus semi-automated hybrid control in pediatric patients with type 1 diabetes using an artificial pancreas. Diabetes Care 31(5):934–939

Werner U (2014) Effects of the GLP-1 receptor agonist lixisenatide on postprandial glucose and gastric emptying–preclinical evidence. J Diabetes Complicat 28:110–114

Wielgosz A, Dai S, Walsh P, McCrea-Logie J, Celebican E (2018) Comorbid conditions in Canadians hospitalized because of diabetes. Can J Diabetes 42(1):106–111

Willi SM, Planton J, Egede L et al (2003) Benefits of continuous subcutaneous insulin infusion in children with type 1 diabetes. J Pediatr 143(6):796–801

Wilson DM, Beck RW, Tamborlane WV et al (2007) The accuracy of the FreeStyle navigator continuous glucose monitoring system in children with type 1 diabetes. Diabetes Care 30(1):59–64

Wilson LM, Castle JR (2018) Stable liquid glucagon: beyond emergency hypoglycemia rescue. J Diabetes Sci Technol 12(4):847–885

Wolpert HA (2008) The nuts and bolts of achieving end points with real-time continuous glucose monitoring. Diabetes Care 31(Suppl. 2):S146–S149

Wysham C, Grimm M, Chen S (2013) Once weekly exenatide: efficacy, tolerability and place in therapy. Diabetes Obes Metab 15(10):871–881

Wysham CH, Campos C, Kruger D (2018) Safety and efficacy of insulin degludec/liraglutide (IDegLira) and insulin glargine U100/lixisenatide (iGlarLixi), two novel co-formulations of a basal insulin and a glucagon-like peptide-1 receptor agonist, in patients with diabetes not adequately controlled on oral anti-diabetic medications. Clin Diabetes 36(2):149–159

Zander M, Madsbad S, Madsen JL et al (2002) Effect of 6-week course of glucagon-like peptide 1 on glycaemic control, insulin sensitivity, and β-cell function in type 2 diabetes: a parallel-group study. Lancet 359:824–830

Pharmacodynamic Evaluation: Gastroenterology

11

Petar Nikolov, Georgi Banishki, and Milena Nikolova-Vlahova

Contents

Abstract

Pharmacodynamics aims to explain the complex relationship between the medication's dose, physiological or pathological response, and the chemical nature of the drug. The human gastrointestinal tract (GIT) is a strictly hierarchic body system with numerous functions and is often a therapeutic target, crosslink, or can even serve as a measurement for drug's physiologic and biochemical effects. The

P. Nikolov (✉)
Gastroenterology, Ardea Medical Center, Sofia, Bulgaria
e-mail: petar.k.nikolov@gmail.com

G. Banishki
University of Aberdeen, Aberdeen, UK
e-mail: banishki@abv.bg

M. Nikolova-Vlahova
Clinic of Nephrology, University Hospital Alexandrovska, Medical University, Sofia, Bulgaria
e-mail: milena_i_dani@abv.bg

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_50

pharmacological effects and the pharmacodynamics evaluation in the GIT would not have been possible without three distinct receptor families that have been known to have an enormous role in the modulation of the GIT functions: serotonergic, cannabinoid, and opioid receptors. In addition, the route of administration can be of great importance for the absorption and pharmacological response of the drug and there are several main routes of administration in gastroenterology: oral, parenteral, transmucosal, and local. Furthermore, a targeted and individualized approach for drug monitoring was developed that takes into account individual patient variability through careful gathering of pharmacokinetic and pharmacodynamic data: therapeutic drug monitoring, allowing to really individualize patient's dose. Various novel imaging methods are used in gastroenterology, e.g., PET scan, MRI, and molecular endoscopy, and they all use tracers and contrast agents. They allow for early an accurate detection of various lesions in the GIT. Current understanding of oral tolerance has allowed the development of two groups of medications with a unique pharmacodynamic profile: oral vaccines, inducing immune response and oral tolerogens, initiating immunomodulation with alteration of immune response directed at the development of local and/or systemic immune tolerance.

General Overview

Petar Nikolov

Pharmacodynamics stands for the study of the physiologic, biochemical and molecular effects of medications on the body and involves receptor binding, postreceptor signaling and biochemical interactions. Pharmacodynamics also aims to explain the complex and multifactorial relationship between the medication's dose, physiological or pathological response, and the chemical nature of the drug.

The human gastrointestinal tract (GIT) is a really complex and yet strictly hierarchic body system with numerous functions:

- Provides route and safe passage of food through the body.
- Plays a key role in the food's processing, degradation and utilization.
- Acts as an "opened door" to the outside world and thus has a major role in the life adaptation and preservation of the biological individuality of humans – GIT is extremely important for the proper functioning of the innate and acquired immune response.
- At the same time plays a pivotal role in the maintenance of oral tolerance.
- GIT is a crossroad for all other body systems from a metabolic, regulatory, and signaling standpoint, thus allowing for a therapeutic intervention at many levels.
- Small and large intestine host the intestinal flora, which is sometimes regarded as the "forgotten organ" in the human body as it is biochemical activity is only comparable with the one of the liver.
- Numerous endocrine and exocrine secreting cells adaptively interact with each other, thus keeping an equilibrium with the other body systems, the nutritional habits, and the outside world.
- Oral and rectal drug intake are often the preferred routes of drug administration for many therapies and would not have been possible without the unique and ubiquitous functions of the GIT.
- Liver is the most advanced and sophisticated biochemical laboratory in the human body and plays a primordial and diverse role in the drug's pharmacodynamics.

All the above features of the GIT may be a therapeutic target, crosslink or serve as measurement for drug's physiologic and biochemical effects. With this regard, the human GIT and the science that studies it – gastroenterology – are an important milestone in the pharmacodynamic

evaluation. Moreover, the human GIT hosts a galore of receptors that are of vital importance for the optimal functioning of the human body as it is a crossroad and often the pharmacological effects of the medications aiming at the GIT may biologically go in their effects beyond the GIT due to its ubiquitous function.

Receptors in the GIT

Georgi Banishki

Targeting specific receptors has always been one of the most preferred options in drug development and the GIT has been no exception in that area. There have been three different receptor families that have been identified that have been of the greatest interest to researchers due to their role in modulating GI functions and these are the serotonergic (5-HT), cannabinoid, and opioid receptors.

Serotonergic Receptors

There have been 12 different serotonin receptors identified to date, all of which are G-protein coupled receptors (GPCRs) and for most time they have been associated with the nervous system and their importance in mood control, depression, anxiety, sleep, etc. However, the greatest store of serotonin in the body is the gut where can be found about 95% of it (Gershon 2013). It was first in the 1950 that was proven that serotonin plays a major role in peristaltic activity as serotonin secreted from enterochromaffin cells in gut mucosa evokes peristaltic activity. Subsequent studies on rats using tryptophan-deficient diet (tryptophan is precursor of serotonin) showed that peristaltic activity was not impaired, but serotonin still has a modulating effect on the gut motility (Gershon 2013). Drugs that target serotonin receptors have been shown to be very effective in patients with IBS. In previously treatment resistant patients, Alosetron, a 5-HT_3 antagonist, was shown to be very effective against IBS with diarrhea, and Tegaserod, a 5-HT_4 agonist, was effective in patient who had IBS with constipation. Both drugs had some severe side effects which have led to their restricted use, but they were a successful proof of concept as it confirmed serotonergic bowel dysfunction to be an important factor in IBS (Gershon 2013).

Another major finding regarding serotonin signaling in the gut is its role in inducing inflammatory response. Animal studies using mice which lacked serotonin reuptake transporter (SERT), the activity of serotonin was enhanced and prolonged, and these mice were susceptible to developing trinitrobenzenesulfonic acid (TNBS)-induced or IL-10 KO-associated colitis (Gershon 2013). The exact pathway for this is yet to be understood although there has been evidence suggesting that by stimulating 5-HT_7 receptors on dendritic cells to launch the innate immune mechanisms, serotonin can cause inflammation of the bowel (Gershon 2013). Studies using KO mice which are missing the synthesizing enzymes for serotonin, tryptophan hydroxylase 1 and 2 (TPH1 and TPH2), have also confirmed the importance of serotonin in gut inflammation. Most interestingly, mice which have TPH1 gene knocked out had a reduction in inflammation while mice where lacked TPH2 inflammation was increased and thus led to the conclusion that serotonin can act as both the sword and shield of the gut (Gershon 2013). Serotonin has also been confirmed as an important factor in liver regeneration with 5-HT_2 receptors found on hepatocytes being involved in promoting DNA synthesis and hepatocytes proliferation (Gershon 2013). Overall, as we understand more about the importance of serotonin in gut function, it is an area that is going to attract even more attention by researchers looking for new therapeutics for different GI disorders.

Cannabinoid Receptors

Cannabis has been used to treat different GI ailments for centuries, but only recently with latest scientific discoveries researchers are beginning to

understand the pharmacologic pathway for this phenomenon. Cannabinoids elicit response through two main types of GPCRs called the cannabinoid 1 and cannabinoid 2 receptors (CB1 and CB2). CB1 receptors are found throughout the whole enteric nervous system (ENS) and the colon epithelium, while CB2 receptors are found primarily in the immune system, and thus they are both involved in multiple processes ranging from GI motility to regulating gastric secretion (Izzo and Sharkey 2010). The main active substance in cannabis, THC, has long been associated with craving of food, and there has been significant evidence accumulated which has directly linked CB1 receptors in the brain and gut with increasing food intake and body weight gain (Izzo and Sharkey 2010). Several antagonists have been developed that target specifically CB1s for promoting weight loss and treating obesity, but only one has reached the market (Rimonabant) and it was later withdrawn due to increased risk of depression (Izzo and Sharkey 2010). Nevertheless, this remains an area with very high potential as rat studies have shown that CB1 expression is upregulated in obesity-prone rats thus confirming the therapeutic potential for future CB1 antagonists that are unable to cross the blood-brain barrier (Izzo and Sharkey 2010). On the other hand, activation of CB1 receptors has also been shown to reduce gastric secretions and decrease gastric ulcers in rodents which is another area that presents exciting new opportunities for future drugs (Izzo and Sharkey 2010).

However, probably the greatest interest has been the involvement of CB receptors in controlling inflammation. Cannabis for many years has been used by patients suffering of autoimmune disease including those affecting the GIT such as Crohn's disease (CD) or ulcerative colitis (UC). Preclinical experiments in humans have shown increased expression of CB receptors and/or enhanced endocannabinoid levels in intestinal biopsies of patients suffering from CD, UC, diverticulitis, and celiac disease (Izzo and Sharkey 2010). Both CB1 and CB2 receptors are possibly involved as in vitro studies have shown them to modulate inflammatory responses and CB1s were also shown to promote gut healing (Izzo and Sharkey 2010). In rodent models, both CB1 and CB2 agonists have shown to be effective in reducing inflammation from trinitrobenzene sulfonic acid (TNBS) and oil of mustard induced colitis, and in rodents where CB1 antagonists were applied they were shown to be more susceptible to induced colitis (Izzo and Sharkey 2010). CB2 receptors by being found mainly in immune cells (B cells, killer T cells) has been shown to be involved in suppressing activated macrophages and the secretion of proinflammatory cytokines such as TNFα (Izzo and Sharkey 2010). In addition, as previously stated both CB1 and CB2 play an important role in regulating gut motility and secretion, their activation by an exogenic compound could cause an even greater reduction in gut inflammation through this process as well. All of this accumulated evidence from in vitro and in vivo studies is highly indicative of the huge future potential of CB receptors in the management of diseases such as CD or UC.

Opioid Receptors

Just like cannabis, opium and its many different derivatives have found therapeutic applications long before modern medicine. For centuries, it has been used for treatment of pain and diarrhea in instances such as cholera infections. It has been confirmed that opiates act by targeting specific opioid receptors, but unlike serotonin and cannabinoid receptors their function is much better understood and utilized. Opioid receptors are also all GPCRs and can be subdivided into three classes – μ-opioid receptors (MOR), κ-opioid receptors (KOR), and δ-opioid receptors (DOR). All three types are found in the myenteric and submucosal plexus of the ENS and MORs are also found in immune cells in the lamina propria of the gut. All opioid receptors are directly linked to controlling Cl^- secretions in the gut and thus water movement, delaying transit from the small intestine to the colon, elevating the resting anal sphincter pressure, and regulating intestinal inflammation (Holzer 2009). The most commonly

used drug acting on MORs is Loperamide which is the most commonly used drug to treat diarrhea caused by infections of IBS. Loperamide is a MOR agonist; it cannot pass the blood-brain barrier and targets the receptors in the ENS thus causing decreased propulsive motility and intestinal secretion (Holzer 2009). A common side effect as one would expect is constipation, but in overdoses it can also cause fatal arrhythmia and it should be avoided in patients with IBD where it can result in toxic mega colon. There is one other antidiarrheal drug which utilizes the MOR signaling pathway without crossing the blood-brain barrier, but indirectly. Racecadotril or acetorphan inhibits enkephalinases, the enzymes which degrade endogenous opioids, thus increasing their concentration which therefore leads to delayed bowel transit (Holzer 2009).

MOR mediated constipation is also quite a common problem for patient treated with opiate analgesics and which suffer from the so-called opioid-induced bowel dysfunction (OBD). Apart from constipation, OBD also included incomplete evacuation, abdominal distention bloating and discomfort, and gastroesophageal reflux, and it persists throughout the whole treatment of the patient and even though he can develop resistance to the analgesic effects of the opioid, the GI effects remain largely the same (Holzer 2009). This can be somehow managed using naloxone, an inverse MOR agonist which could counteract the undesirable effects without compromising the analgesia. This however is limited by its narrow therapeutic range and ability to cross the blood-brain barrier and at higher doses it can greatly reduce analgesia (Holzer 2009). As a result, one approach that has been attempted is by using peripherally restricted opioid receptor antagonists like n-methylnaltrexone, which has both low oral bioavailability and cannot cross the blood-brain barrier. This concept has subsequently been verified in rat, dog, and human studies using both oral and parenteral formulations, and it has now been approved for human use by both the FDA and EMA (Holzer 2009). Nevertheless, its long-term safety and tolerability are not yet known so the recommendation is not to use it for longer than 4 months. Other such antagonists have been developed (e.g., alvimopan), but they have found only limited clinical use. Nevertheless, peripheral MOR antagonists remain an area of great interest for the future.

Treatment Routes and Drug Delivery

Petar Nikolov

The ubiquitous physiological properties of the human GIT along with the various pathology to be found there predispose for a galore of treatment options to be considered in humans. The development of numerous acid suppressing drugs, antiviral agents for viral hepatitis, biologics for inflammatory bowel disease, live biotherapeutics for intestinal disease, etc. has changed the face of gastroenterology forever. These treatment options come with particular drug delivery techniques. In an attempt to improve the efficacy and safety profile of medications, researchers have developed different methods such as individualizing drug therapy, dose titration, therapeutic drug monitoring, delivering drug at controlled rate, targeted delivery, etc. (Tiwari et al. 2012).

The most commonly used routes of drug administration in gastroenterology are:

- Oral – delivering the drug into the stomach, small or large intestine
- Parenteral – subcutaneous, intravenous, intra-arterial, intramuscular, intralesional
- Transmucosal – transrectal, transnasal, and sublingual
- Local – mostly suppositories or enemas

Oral Drug Administration

The oral drug administration is probably the most commonly used method of drug administration. It is using oral formulations that could open into the stomach, small or large intestine.

There is a great and somewhat unmet need in oral delivery of protein and peptide drugs, suitable devices for delivering the therapeutic agents into

the systemic circulation. Numerous gelatin capsules, film tablets, sustained release capsules, etc. have been developed in the last decades so to boost the efficacy and tolerability of numerous medications aiming GIT pathology (Tiwari et al. 2012).

Oral administration in gastroenterology is only indicated in cases where patients can swallow properly, it is believed that this route of administration would be more beneficial as compared with the others, the drug is not likely to be destroyed or inactivated by stomach acid, pancreatic enzymes, bile acids or colonic bacteria and last but not least the drug would not be inactivated in the intestinal wall and/or the liver (first pass metabolism). Oral medications to be given with a glass of water in an upright position and washed down with a sufficient amount of water. Oral medications should not be given to a recumbent patient due to the risk of aspiration, chocking and also due to the risk of damages to the esophageal mucosa especially by some medications (e.g., tetracyclines, iron salts). To prevent gastric irritation and to achieve the desired concentration, some researchers have developed enteric coated tablets that resist the gastric acid and disintegrate in the intestine alkaline contents. This also helps to achieve the desired concentration of the drug in the small intestine (e.g., in Crohn's disease) and last but not least to retard the absorption of the drug. Furthermore preparations with colonic release have been developed. Oral formulations can be designed so to release the active substance over different period of time so there is a normal and controlled release oral formulations.

The constantly increasing number of peptide and protein drugs being investigated demands the development of novel dosage forms which exhibit also site-specific release. Delivery of drugs into systemic circulation through colonic absorption represents a novel mode of introducing peptide and protein drug molecules and drugs that are poorly absorbed from the upper GIT (Pinto-Alphandary et al. 2000).

Specific targeting of drugs to the colon is recognized to have numerous therapeutic advantages per se and drugs, which are destroyed by the stomach acid and/or metabolized by pancreatic enzymes or affected by bile acids, are slightly affected into the colon. Colon targeting is of value for the topical treatment colonic pathology such as Crohn's disease, ulcerative colitis, amebiasis, and colorectal cancer. Sustained colonic release of medications can be useful also in the treatment of non-GIT conditions. Peptides, proteins, oligonucleotides, colonic diagnostic agents, and even oral vaccines are potential candidates of interest for colon-specific drug delivery. The diverse microflora and numerous enzymes present in the human colon are being exploited to release drugs in the colon (Tiwari et al. 2012); however, some pharmacodynamic obstacles are also involved in the effective local delivery of drugs to the colon due to the artificial bypass of the stomach and small intestine: unpredictable effect of the gut flora that could vary in its composition from person to person (Lagier et al. 2012), differential pH conditions in the colon, differences in the dietary habits, long transit time during the passage from mouth to colon create difficulties in the safe delivery of drugs to the large intestine (Tiwari et al. 2012).

Recent technological achievements such as drug coating with pH-sensitive and bacterial degradable polymers, embedding in bacterial degradable matrices and designing into prodrugs are aiming to effectively target drugs to the colon. The use of pH changes is similar to the enteric coating and consists of employing a polymer with an appropriate pH solubility profile. The concept of using pH as a trigger to release the drug in the colon is based on the pH conditions that vary significantly down the GIT. Polysaccharide and azopolymer coating, which is refractory in the stomach and small intestine yet degraded by the colonic bacteria, have been used as carriers for colon-specific targeting. Last but not least, the availability of good preclinical models and clinical methods promoted the quick development and evaluation of colon-specific drug delivery systems for clinical practice (Tiwari et al. 2012).

Parenteral Route of Administration

Routes of administration other than the oral are called parenteral. These are used mostly when oral

therapy is not possible, not well tolerated (e.g., oral therapy triggers vomiting, diarrhea), patients cannot swallow, drug is not absorbed orally, to avoid drug modification by the GIT and when rapid systemic action and dose accuracy are to be ensured. GIT sometimes limits the bioavailability of certain medications because of its protease enzymes and bacteria-rich environment as well as general pH variability from pH 1–7. These extreme conditions make oral delivery particularly challenging for the some medications, e.g., biologics, insulin.

A common parenteral route of administration in gastroenterology is the subcutaneous injection. It is used for the application of nonirritant substances (e.g., somatostatin analogues). The drug absorption is slower but the action is sustained and uniform. It often comes with great efficacy and variable tolerability depending on the type of the active substance administered. Overall, the immunogenicity of subcutaneously administered proteins depends upon antigen presentation and processing by lymph nodes and migratory cutaneous dendritic cells in the subcutaneous space (Fathallah et al. 2013). Another parenteral route of administration is the intravenous one. In this case, drugs are given directly into a vein. Normally the drug produces a rapid effect and the target serum concentration can be achieved with lower doses administered. The drug may be given as a bolus, over 5–10 min, or as continuous infusion (e.g., rehydration), over prolonged periods of time. Some medications are considered to have irritant effect when administered intravenously (e.g., iron, cancer chemotherapy, potassium solutions, parental feeding, etc.). Use of intra-arterial administration in gastroenterology is very limited and is used mostly in cases of angiography and embolization therapy (e.g., hepatocellular carcinoma). The intramuscular route of administration allows for the administration of soluble substances, mild irritants, colloids, and suspensions. The volume of injection should not exceed 10 ml. The intramuscular administration of vaccines optimizes the immunogenicity of the vaccine and minimizes adverse reactions at the injection site, e.g., HBV vaccine (Zuckerman 2000). Intralesional injections have features of both parenteral and local drug administration. Intralesional injections in gastroenterology are often given under ultrasound control or via endoscope (e.g., endoscopic intralesional steroid injection in refractory esophageal strictures, endoscopic intralesional injection of diluted epinephrine (1:10,000) in the prevention of recurrent bleeding).

Transmucosal Route of Administration

The transmucosal route of administration is characterized by several main features: it is normally painless; offers greater flexibility in a variety of clinical situations, including patients who cannot swallow oral medications and/or in cases when it is not possible to establish intravenous access. Additionally it is characterized by a rapid onset of pharmacological effect, which is often preferred for drugs, especially in the treatment of the acute disorders. Human mucosa has rich blood and lymph supply and many drugs can cross the rectal mucosal membrane like any other lipid membrane, meaning that unionized and lipophilic substances are readily absorbed. Many drugs are using the so-called transrectal administration route. The rectum has rich blood and lymph supply: the portion of the drug absorbed from the upper rectal mucosa is carried by the superior hemorrhoidal vein into the portal circulation, whereas the portion absorbed by the lower rectum enters directly into the systemic circulation via the middle and inferior hemorrhoidal vein. Because of that absorption pattern approximately 50% of drug absorbed by the rectum bypasses the liver and additionally CYP3A4 is not present in the lower intestinal segments, meaning that the chances for first pass metabolism are significantly lower as compared with oral drug administration (e.g., indomethacin suppositories used after ERCP for the prevention of post ERCP-pancreatitis). The transnasal and sublingual routes are less commonly used in gastroenterology but again can be really efficient and with a really good safety profile overall (e.g., intranasal fentanyl in procedural and postprocedural pain in children and sublingual nitroglycerin again in the prevention of post ERCP-pancreatitis).

Local Drug Administration

The local drug application in gastroenterology is mostly given in the form of suppositories (e.g., Mesalazine in distal forms of ulcerative colitis) or enemas. The systemic absorption in local application is negligibly low. Enemas can be divided into retention and evacuant enema. In retention enema the fluid containing the drug (e.g., methylprednisolone in ulcerative colitis, mesalazine foam in ulcerative colitis) is usually 100–120 ml. The evacuant enema (e.g., soap water enema before abdominal surgery, X-ray of GIT) aims to remove the fecal matter and flatus. The liquid stimulates bowel movements by distending the bowel wall, whereas soap acts as a softener. The overall quantity of fluid administered is usually up to 600 mg.

Fecal microbiota transplantation holds a special place in gastroenterology and could be given as a retention enema (but also in the form of oral capsules) for the local treatment of recurrent *Clostridium difficile* infection and also ulcerative colitis (Rossen et al. 2015).

The development of micelles, liposomes, and even nanoparticles are being currently researched and integrated in oral and parenteral GIT medications. The aim of these sophisticated drug delivery systems is to provide enhanced efficacy for existing and novel medications and/or reduced toxicity for patients. These drug delivery systems may be subjected to even further changes such as PEGylation of liposomes and nanoparticles so to boost their efficacy, formation of nanogels, and solid lipid nanoparticles. These novel drug delivery systems are largely experimental but have also shown some efficacy in some gastrointestinal tumors.

Therapeutic Drug Monitoring

Georgi Banishki

Management of serious progressive diseases of the GIT offers countless challenges to gastroenterologist from lack of therapeutic response, genetic polymorphisms, drug interactions to adverse drug reactions. This requires a targeted approach that takes into account individual patient variability through careful gathering of pharmacokinetic and pharmacodynamic data and this is where therapeutic drug monitoring (TDM) comes into account. TDM has been defined as using laboratory measurements usually a biological matrix of a parameter (e.g., drug metabolites) which after analysis will directly influence patient therapy (Dasgupta 2012). TDM is not applied for all lines of treatment but only in situations when: (1) clinical evidence is deemed insufficient, (2) correlation between serum or whole blood drug concentration and dosage is poor, (3) there is a narrow therapeutic range, (4) drug toxicity may lead to serious adverse events, (5) there is a link between serum or whole blood concentration of the drug and its therapeutic response or toxicity, and (6) there are clinical indications (e.g., toxicity despite no dosage adjustment) which require it (Dasgupta 2012). Naturally, one would expect that chemotherapeutic treatments would be subject to TDM, but that is the case in only a select number of cases. Previously discussed PET imaging has been a huge improvement in monitoring chemotherapy efficacy, but classical pharmacological tests are still rarely used – 5-fluoruracil treatments are one notable exception (Dasgupta 2012). However, in terms of the GIT TDM has started to make a mark in helping choose the best treatment for patients with IBD. In the past, steroids were the preferred choice for IBD management, but nowadays they are primarily used for controlling diseases flares at relapse periods, but for maintaining remission other drugs are now preferred. The main drugs that have been universally accepted for use in IBD maintenance can be subdivided into three classes and these are the aminosalicylates, TNFα inhibitors, and thiopurines. Aminosalicylates (e.g., mesalazine, sulfasalazine) are a class of anti-inflammatories and are the first-line choice for UC, but not so much for CD. These agents have been used for decades without close blood monitoring of metabolites and dosing regimens have been adjusted according to the clinical response and manifestation of symptoms. With anti-TNFs and thiopurines this is not the case due to their narrow therapeutic window and high risk of adverse drug

reactions, and close TDM is deemed necessary for all patients. This has allowed for selecting the best therapy for the patient and if needs be altering it in order to account for evolving loss of response (LOR) or patient safety concerns.

Anti-TNFs are all monoclonal antibodies and are the latest line of drugs used to treat IBD that have come during the last decade. They are the primary recommended therapy for advanced CD or UC and the following drugs have been approved for human use – infliximab, adalimumab, certolizumab (only in the USA), and golimumab. The first concern when giving one of these drugs is the occurrence of primary LOR, or lack of effect after the first phase of treatment which can occur in up to one third of all patients (Kopylov et al. 2014). The second concern is secondary LOR which is also quite common and which is much harder to assess and can occur at any point of a treatment regimen. For some drugs like infliximab, it has been reported to appear in up to two-thirds of all patients in the first year, while in others like adalimumab it has reported as every fourth patient. The main factors that can cause primary LOR are disease progression, age of patient, genetic polymorphisms, smoking, and prior exposure to such drugs. For secondary LOR, the main cause has been immunogenicity and the development of antidrug antibodies (ATIs) (Kopylov et al. 2014). ATIs develop for both chimeric and fully humanized anti-TNFs and act by interfering with their binding to TNFα molecules. Consequently, the main tools that have been used for TDM in patients treated with anti-TNFs have been evaluation of serum levels of drug metabolites and ATIs as well disease activity measurements. The most common method used is double-antigen ELISA in which the drug molecule (e.g., infliximab) is both the capture antigen and the detection antibody. This has the drawback of being unable to detect ATIs in the presence of the drug in the serum which can be fixed by using antihuman λ antigen detection antibody (AHLC) ELISA which has this capacity (Kopylov et al. 2014). Disease activity is measured by regular monitoring of an inflammatory marker (e.g., CRP/FCP) and endoscopy (Fig. 1). In most cases of primary LOR, this is not due to low levels of drug metabolites but due to increased clearance (e.g., fecal loss in UC) and ATI formation which result in low serum levels (Kopylov et al. 2014). Generally, in such cases increasing the dosing

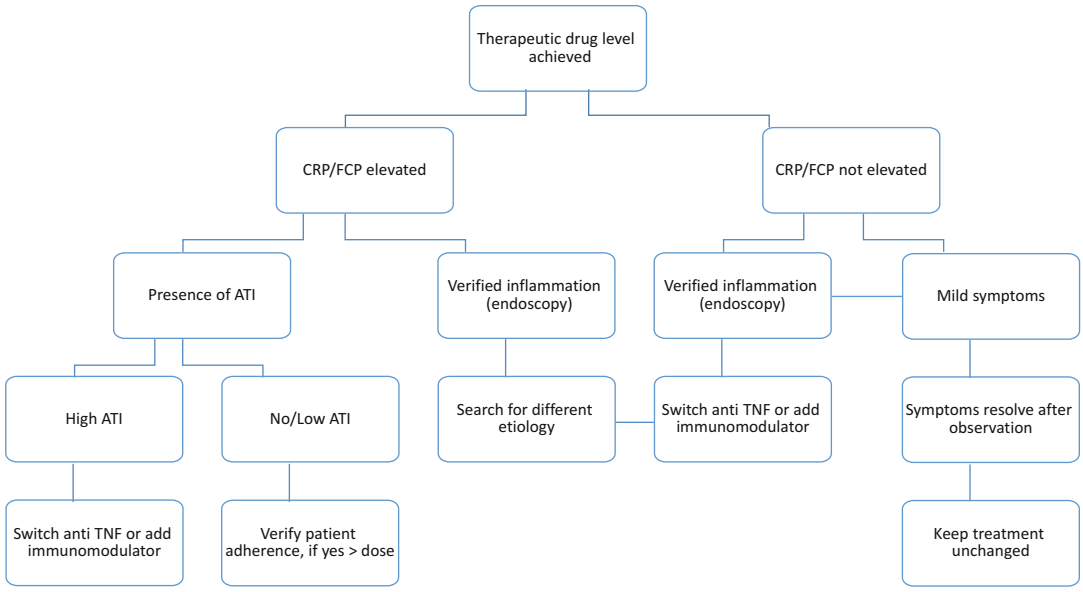


Fig. 1 TDM based algorithm for management of loss of response to TNFα inhibitors (Adapted for use from Kopylov et al. 2014)

regimen has been ineffective, but switching from one medication to another is something that could have positive results (e.g., infliximab replaced with adalimumab). For cases when the patient experiences secondary LOR, a careful assessment of the level of inflammation and drug metabolites is needed. In case that there is active inflammation and high drug levels, the best option would be to switch to another medication. However, if the inflammation is seen, but drug levels are low there would be two options – first would be to increase the dose and the other, but much less common would be to add an immunomodulator in order to suppress ATI formation. Clinical studies where azathioprine or methotrexate was given in addition to infliximab showed reduction of ATIs and the return of clinical response and improvement in disease management (Kopylov et al. 2014).

Thiopurines are immunomodulators or immunosuppressants that act by blocking purine synthesis and thus inhibiting T cell production. The main drug of this class azathioprine was initially developed as an anticancer medication, but was then found out to be extremely effective in transplantations and later became the mainstay in IBD management. In recent years, azathioprine has been gradually replaced as first choice by the monoclonal antibodies as they have been deemed to be more target specific and with less side effects (Kopylov et al. 2014). The most common cause for discontinuation of treatment with thiopurines have been its adverse effects such as myelosuppression due to their interference with DNA synthesis, but there are also about 9% of all patients who do not respond to this line of treatment (Kopylov et al. 2014). Therefore, therapeutic drug monitoring is done by monitoring thiopurine metabolites, but also by monitoring blood counts and doing routine checks on pancreatic enzymes. One other factor that has also been recommended to be taken into account is genetic polymorphisms. Once administered azathioprine is rapidly converted to active metabolite 6-mercaptopurine (6-MP) by a nonenzymatic reaction and after that there are two competing pathways (Kopylov et al. 2014). The first is mediated by the enzyme thiopurine methyltransferase (TPMT) which converts 6-MP to the inactive 6-MMP which gets excreted or the other pathway which converts it to 6-thianoguanine (6-TGN). 6-TGN is an active metabolite which causes myelosuppression, and in patients who were shown to have low or intermediate acting, TPMT was at much greater risk of myelosuppression due to 6-TGN accumulation (Kopylov et al. 2014). Three such alleles have been confirmed in Caucasian (TPMT*2, TPMT*3A, or TPMT*3C) and one in African-American (TPMT*3C) populations, and the FDA now recommends genotype or phenotype assessment of TPMT prior to initiating azathioprine (Kopylov et al. 2014). In case when patients are found to be homozygous of any of these, azathioprine should be avoided, and in case when they are heterozygous a decrease in recommended dose by 30–70% has been suggested (Kopylov et al. 2014). Nevertheless, even if the patient has been confirmed not to be heterozygous for any of these alleles regular CBCs should be performed as myelosuppression has been seen in patients with normal TPMT after long-term treatment with azathioprine (Kopylov et al. 2014).

Imaging in Gastroenterology

Georgi Banishki

As in other functional areas, imaging plays a major role in diagnosing GI diseases and even though there were times when it relied on radiological studies it has expanded to include novel technologies which utilize tracers and contrast agents. The imaging techniques used play a great role both in diagnosing and assessing treatment efficacy and can be subdivided into two categories: anatomical and functional. The anatomical techniques such as computer tomographic (CT) scans as their name signifies focus on observation of structural changes and identification of anatomical landmarks induced by GI ailments (e.g., solid tumors), while the functional ones rely on detecting functional and metabolic changes. These techniques can include functional magnetic resonance imaging (fMRI), but also the recently

developed molecular imaging techniques such as positron emission tomography (PET). It is important to note that no single technique is superior to another which is why they are always used in connection, but this section will focus on the ones that rely on the use of pharmacological agents and their molecular interaction with biological targets.

Positron Emission Tomography (PET) Imaging

Despite of research in the field dating as far as the 1950s, PET scanning has been approved for use in colorectal and esophageal cancer diagnostics and management since 2001 and since then it has revolutionized the field. Several studies have shown that PET results have caused a significant change in cancer management in 25% and in some cases up to 40% of patients (Bailey et al. 2005). PET scanning is a branch of nuclear imaging which uses radioactive isotopes called nucleotides or tracers aiming to particular organs and structures. The radioactive isotope can be swallowed, injected or inhaled depending on the target of interest. Each radiotracer is specifically designed to be taken by a metabolic pathway at the targeted tissues where it will accumulate. All radionucleotides are designed to have short half-lives and as they are decaying they are emitting positrons (e+) which travel in the tissue until they meet an electron (e−) which annihilates both and result in the production of photons which then get detected by the scanning device. Once all of these emissions undergo computer reconstruction a 3D image is created and in which the targeted areas will be highlighted and thus may indicate that there are metabolic processes associated with a particular disease in this area (Bailey et al. 2005). PET imaging has found extensive application mainly in neurology and oncology, and it is in the diagnosis and treatment monitoring of different cancers that it has found its greatest use in the GIT (Table 1).

Table 1 UK intercollegiate Committee recommended indications for clinical PET studies in the GIT: (A) supported by randomized controlled clinical trials, meta-analyses, and systematic reviews, (B) by experimental or observational studies, and (C) other evidence (Adapted from Bailey et al. 2005)

Oncology applications	Indication	Not indicated routinely	Not indicated
Esophagus	Staging of primary cancer (B) Assessment of disease recurrence in previously treated cancers (C)	Assessment of neoadjuvant chemotherapy (C)	
Stomach	No routine indication (C)	Assessment of gastroesophageal malignancies and local metastases (C)	
Small bowel	No routine indication (C)	Proven small bowel lymphoma to assess extent of disease (C)	
Liver	Equivocal diagnostic imaging (CT, MRI, ultrasound) (C) Assessment pre- and posttherapy intervention (C) Exclude other metastatic disease prior to metastectomy (C)		Routine assessment of hepatoma (C)
Pancreas		Staging a known primary (C) Differentiation of chronic pancreatitis from pancreatic carcinoma (C) Assessment of pancreatic masses to determine benign or malignant status (C)	
Colon and rectum	Assessment of recurrent disease (A) Prior to metastectomy of colorectal cancer (C)	Assessment of tumor response (C) Assessment of a mass that is difficult to biopsy (C)	Assessment of polyps (C) Staging a known primary (C)

Anatomical imaging techniques remain the mainstay in early diagnosing of cancer, but in other aspects such as tumor staging and monitoring treatment outcome and disease progression is where PET being applied to the best benefit. Accurate evaluation of treatment response is critical for optimal treatment decisions in different types of cancer, and Response Evaluation Criteria in Solid Tumors (RECIST) group has come up with criteria on how to evaluate tumor response to treatment (Table 2).

RECIST assesses tumor response by the extent of tumor size reduction and in the past this assessment was done primarily using CT scans or MRI, but as these can be sometimes insufficient in visualizing all tumor infiltration points. This has led to the fusion of PET and CT scan technologies and the development of fluorodeoxyglucose-positron emission tomography (FDG-PET/CT). FDG is a glucose analogue radiotracer which is differentially taken up by malignant cells due to their higher glucose metabolism and this can be used to monitor both short- and long-term metabolic response of tumor after chemotherapy (Van Cutsem et al. 2016). This method utilizes the affinity of tumor cells for FDG, which is strongly linked to tumor grade (aggressiveness) and cellularity. Changes in FDG uptake can be detected after a single course of chemotherapy and as early as 24 h after treatment and thus can help discriminate malignant CRC tumors that are unlikely to respond to treatment. Metastatic CRC is usually highly responsive to FDG, with the exception of mucinous tumors, which may not be detected by a FDG-PET/CT scan (Van Cutsem et al. 2016). Other digestive tumors that can result in false-negative results are low grade neuroendocrine tumors, well-differentiated hepatocellular carcinoma, nonmass forming gastric tumors, and mucinous or cystic pancreatic tumors. Nevertheless, FDG-PET/CT has shown to be much more sensitive in tumor detection and especially for evaluating responses to chemotherapy and especially for targeted treatments than CT scans or FDG-PET imaging on their own. For example, in the multicenter SoMore trial, FDG-PET/CT was used to assess early metabolic response following treatment with combination of sorafenib and capecitabine in chemorefractory malignant CRC patients (Van Cutsem et al. 2016). Just after one single treatment cycle (week 3), FDG-PET/CT was able to differentiate responsive from unresponsive lesions (Van Cutsem et al. 2016). These and other such findings support the notion that FDG-PET/CT will be able to greatly improve the management of chemotherapies in the future, but more such studies will be needed in order to verify this. Furthermore, as the RECIST scale was designed for CT scanning assessment a much less ambiguous response scale called the PET response criteria in solid tumors (PERCIST) has been developed which takes metabolic criteria into much greater account (Table 3). Nevertheless, due to its lack of focus on treatment resistance

Table 2 Response Evaluation Criteria in Solid Tumors (RECIST) (Adapted from Van Cutsem et al. 2016)

Grade	Response criteria
Complete response	Disappearance of all target lesions. Any pathological lymph nodes(whether target or nontarget) must have reduction in short axis to <10 mm
Partial response	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters
Progressive disease	At least a 20% increase in the sum of diameters of target lesions, the appearance of one or more new lesions is also considered progression
Stable disease	Neither sufficient shrinkage to qualify for partial response nor sufficient increase to qualify for progressive disease

Table 3 PET response criteria in solid tumors (PERCIST) (Adapted from Poeppel et al. 2002)

Grade	Response criteria
Complete metabolic response	Complete resolution of FDG uptake
Partial metabolic response	≥30% decrease in FDG uptake with absolute drop in standardized uptake value (SUD)
Progressive metabolic disease	≥30% increase in FDG uptake or appearance of new FDG-avid lesions
Stable metabolic disease	No partial metabolic response or disease progression

improved FDG-PET response criteria will need to be developed in the future (Van Cutsem et al. 2016).

One other benefit of FDG-PET/CT compared to normal CT scan is that can provide complementary metabolic information that can enable the detection of tumors at unexpected sites or in appearing morphologically normal structures. A common problem associated with colorectal cancer (CRC) is its ability to metastases and most specifically to form metastases in the liver and you cannot achieve disease control without limiting this spread. Once again CT scans and MRI are the preferred method, but FDG-PET/CT is starting to make an impact here as well. FDG-PET/CT can effectively detect extrahepatic disease and has higher sensitivity (64% vs. 89%) and specificity (70% vs. 90%) in this setting (Van Cutsem et al. 2016). Performing FDG-PET/CT in addition to conventional imaging can further support decision-making as one study of 150 patients with metabolic CRC found the addition of FDG-PET to CT resulted in the avoidance of an unnecessary laparotomy in a significant proportion of patients (38%) (Van Cutsem et al. 2016). The expansion of FDG-PET/CT to evaluate the response of metastatic disease remains a growing area of research. FDG remains the golden standard in PET imaging in 90% of all applications, but there are other PET radiotracers that have been tested in oncology and are likely to have greater importance in the future. One such example is ^{15}O-water which can indicate perfusion and hypoxia which are both markers of tumor angiogenesis and increased metastatic activity (Van Cutsem et al. 2016).

There have been attempts to expand PET imaging into other disease areas and one such example has been lymphocytic gastritis which accounts for about 5% of cases of chronic gastritis and its symptoms are often nonspecific and include abdominal pain, nausea and vomiting, or weight loss (Murphy et al. 2017). Therefore, using endoscopic appearance of the mucosa for diagnosis can be unreliable as it can appear as nodules, erosions, enlarged and prominent rugae, or normal. In a case study by Murphy and colleagues, they were able to diagnose for the first time lymphocytic gastritis using PET scan. A 20-year-old man presented for evaluation of diffuse abdominal pain after he had completed chemotherapy 2 months earlier for stage IV diffuse large B-cell lymphoma (DLBCL). During the examination, abnormal PET findings showed him to have a new area of intense activity within the proximal stomach (Fig. 2a). Subsequent upper endoscopy (Fig. 2b) showed scattered shallow ulcers throughout the stomach and gastric biopsies confirmed the patient to be suffering from lymphocytic gastritis-type pattern with acute inflammation and reactive gastropathy. Further immunostaining showed that these were CD3+ T-cells rather than the B-cells expected in recurrent DLBCL and along with the negative *H. pylori* and Celiac disease serologies confirmed the final diagnosis of lymphocytic gastritis. The patient underwent treatment with several days of daily proton pump inhibitors (PPI) and the repeat upper endoscopy and PET scan demonstrated resolution (Fig. 2c and d). Despite the fact this was a single event, this has led to the speculation that the increase in intraepithelial lymphocytes and associated inflammation is the etiology of hypermetabolic activity seen on the PET scan and could account for other abnormal observations from previous studies (Murphy et al. 2017).

Contrast Enhanced MRI (CE-MRI)

Chronic inflammatory conditions such as Crohn's disease (CD) have always been diagnosed using invasive techniques such as endoscopy and biopsies. However, as like with any other autoimmune disease CD is characterized by alternating remission and relapsing phases which lead to progressive intestinal damage and loss of function this requires a closer examination and a larger view of the gut and endoscopy is insufficient in these aspects. Therefore, cross-sectional imaging techniques like computed tomography enterography (CT) or MRI have been recommended as suitable techniques for assessing both mucosal healing and long-term disease progression and thus can help to establish the best treatment strategy (Savarino et al. 2017). As CD patients have a high risk of developing anal abscesses and fistulas

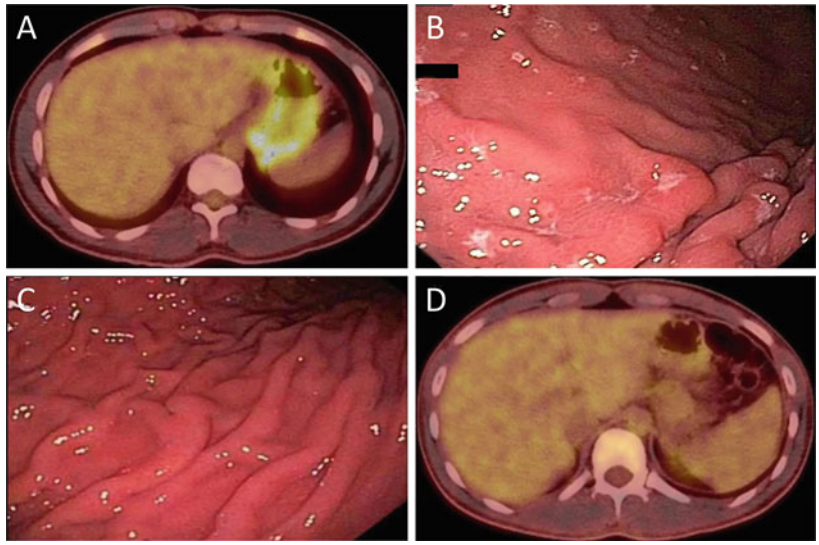


Fig. 2 (**a**) Intense activity in proximal stomach detected by PET; (**b**) shallow gastric ulcers seen by endoscopy; (**c**) full mucosal healing after repeat endoscopy; (**d**) resolution of hypermetabolic activity after repeat PET and PPI therapy (By permission of Murphy et al. 2017)

(30–50%), and CT has high radiation hazard, MRI with its high contrast resolution has become the preferred choice for assessing treatment options for CD. MRI has demonstrated high accuracy for the assessment of mucosal lesions and has shown to be a reliable alternative to ileocolonoscopy as it reveals not only the gut mucosa, but all other bowel layers and is thus preferred by the patients and physicians for frequent disease examinations (Savarino et al. 2017).

In order to achieve better contrast of the image, IV injection of Gadolinium-based contrast agents (GBCA) (contrast-enhanced MRI, CE-MRI) has been demonstrated to be of crucial importance for evaluating mucosal inflammation, transmural involvement, and extraintestinal disease (Savarino et al. 2017). GBCAs are small molecular chelates with high stability containing a Gadolinium ion (Gd3+) which causes reduction of the T1 (rate of longitudinal relaxation) and T2 (rate of transverse relaxation) by modifying the relaxation of closer water protons which in turn leads to positive enhancement on T1-weighted images (Savarino et al. 2017). This in turn helps to better distinguish wall thickness and inflammation that are associated with CD. Nine different GBCAs have been approved for human use since 1986 and until recently they were all thought to be quite safe (Savarino et al. 2017). However, there has been accumulating evidence of Gadolinium accumulation in tissues. There have been animal models and clinical findings which have demonstrated Gadolinium accumulation in brain tissue (Savarino et al. 2017). This has been confirmed only for patients who had CE-MRI for neurological conditions, but there is mounting evidence of this also occurring in patients with no brain disorders. GBCAs are also suspected as a possible cause of the life-threatening condition nephrogenic systemic fibrosis (NSF) which could result in potentially fatal renal failure (Savarino et al. 2017). This has led to warning being given to GBCAs being used in patients with estimated glomerular filtration rate <30 mL/min/1.73 m2 and the European Medicines Agency has recommended the suspension of the market authorizations for four linear GBCAs "because of evidence that small amounts of the Gadolinium they contain are deposited in the brain" (gadobenate, gadodiamide, gadopentetate, and gadoversetamide) (Savarino et al. 2017). Nevertheless, CE-MRI remains the most preferred choice for assessing CD treatment response and progression and will likely remain so in the future.

Molecular Endoscopy

Nuclear medicine has its advantages, but in terms of early cancer detection optical imaging such as endoscopy remains the preferred choice due to neoplasms usually occurring in the epithelial and mucosal surfaces of the GIT. However, observation of structural changes and identification of anatomical landmarks can sometimes be insufficient as there are flat genetically heterogeneous precancerous lesions that are not visible to the naked eye and are virtually undetectable when using standard white light endoscopy. This has led to the development of a new emerging technique, molecular endoscopy which has been designed to better visualize genetic and molecular changes typical of cancerous disease by utilizing autofluorescence or fluorescent molecular compounds targeted at disease specific markers. Molecular endoscopy offers the opportunity to substantially improve specificity through detection of targets that are unique to disease. This technique has been demonstrated in preclinical models using genetically engineered animals that replicate the molecular pathogenesis of human disease. Molecular endoscopy has been demonstrated clinically in the colon and esophagus and is being developed for use in the stomach, biliary tract, and pancreatic duct (Table 4) (Lee and Wang 2016).

Molecular endoscopy has been specifically designed to exploit the uniqueness and high expression levels of targets that are specific of GI diseases. Consequently, for molecular imaging have been recognized agents with high affinity and target specificity as well as rapid clearance and preferably low cost. Fluorophores have been used for quite a while in contract imaging techniques such as flow cytometry and several such as fluorescein isothiocyanate (FITC) have been approved for human use (Lee and Wang 2016). The way that fluorophores work is by emitting light after light excitation and when attached to an antibody with high target specificity they are the perfect for staining and thus identifying objects of interest. Other than antibodies a broad range of platforms for fluorescent dyes have been tried out in both preclinical and clinical studies, including enzyme-activatable probes, peptides, and lectins (Lee and Wang 2016). FITC and other such fluorophore are generally inexpensive, emit light in the visible spectrum and thus provide images with high resolution and reduced depth.

In preclinical studies molecular endoscopies has been demonstrated to detect overexpression of cathepsin B in colonic adenomas and also to monitor tumor shrinkage after treatment with sirolimus in mice animal models (Lee and Wang 2016). In clinical studies, fluorescent antibodies and peptides have been used to demonstrate nonpolypoid lesions in the proximal colon. It had been previously demonstrated that in the proximal colon up to 27% of all lesions can be invisible to white light endoscopy. In a study where 15 patients were injected with Cy5-labeled GE-137 peptide, a peptide that binds to c-Methionine overexpressed in dysplastic crypts, Burggraaf and colleagues successfully demonstrated that flat lesions can be detected with this method (Lee and Wang 2016). Using a modified fiber-optic colonoscope, they showed that all of the 47 tubular adenomas detected had increased uptake of this fluorescent peptide. Furthermore, nine additional adenomas that remained unseen under white light alone were detected (Lee and Wang 2016). These were all small in size and of nonpolypoid nature thus confirming the potential of this imaging technique. Last but not least, molecular endoscopy has been demonstrated also into monitoring therapy effectiveness. Adalimumab is a drug used to treat CD by counteracting Tumor Necrosis Factor α (TNFα) and in up to 50% of patients there is no response to treatment. Atreya and colleagues gave topically 25 patients FITC-labeled anti-TNF antibody in 25 patients and then treated with adalimumab (Lee and Wang 2016). Using confocal laser endomicroscope it was demonstrated that patients with high number of cell with membrane-bound TNF responded much better to the therapy and this was demonstrated over a course of a whole year (Lee and Wang 2016). Molecular endoscopy has also been effectively demonstrated in Barrett's esophagus imaging. In two separate clinical experiments using topically administered FITC-labeled peptide specific for CypA, the fluorescence images collected demonstrated up to

Table 4 Summary of key ex vivo and in vivo GI studies using targeted fluorescent labels and the different detection instruments used (Adapted for use from (Lee and Wang 2016)

Indication	Target	Carrier molecule	Fluorescent label	Administration method	Species	Instrument	Year of experiment
Colitis	GGT	Enzyme	Rhodamine green	Topical	Mouse	Wide-field fluorescence endoscopy	2013
Colonic adenoma	Cathepsin B	Enzyme	Cy5.5	Injection	Mouse	Wide-field fluorescence endoscopy	2010
Colonic adenoma	EGFR	Peptide	Cy5.5	Topical	Mouse	Wide-field fluorescence endoscopy	2015
Colonic adenoma	HER2	Peptide	Cy5.5	Topical	Mouse	Wide-field fluorescence endoscopy	2016
Colonic adenoma	Claudin-1	Peptide	Cy5.5	Topical	Mouse	Wide-field fluorescence endoscopy	2016
Colorectal cancer	EGFR	Antibody	FITC	Injection	Mouse	Confocal laser endomicroscopy	2010
Colorectal cancer	VEGF	Antibody	AF488	Injection	Mouse	Confocal laser endomicroscopy	2010
Sessile serrated adenoma	Unknown	Peptide	FITC	Topical	Human	Wide-field fluorescence endoscopy	2016
Colonic adenoma	c-met	Peptide	Cy5	Injection	Human	Wide-field fluorescence endoscopy	2015
Crohn's disease	Membrane-bound TNF	Antibody	FITC	Topical	Human	Confocal laser endomicroscopy	2014
Colonic adenoma	EGFR	Antibody	AF488	Topical	Human	Confocal laser endomicroscopy	2013
Colonic adenoma	Caspr-1	Peptide	FITC	Topical	Human	Confocal laser endomicroscopy	2008
Esophageal squamous cell cancer	Periostin	Antibody	Cy5.5	Injection	Mouse	Wide-field fluorescence endoscopy	2013
Barrett's neoplasia	HER2	Antibody	AF488	Injection	Rat	Confocal laser endomicroscopy	2015
Barrett's neoplasia	Glycans	Lectin	AF488	Topical	Human (ex vivo)	Wide-field fluorescence endoscopy	2012
Barrett's neoplasia	CypA	Peptide	FITC	Topical	Human	Confocal laser endomicroscopy	2016
Barrett's neoplasia	CypA	Peptide	FITC	Topical	Human	Confocal laser endomicroscopy	2013
Gastric cancer	MG7	Antibody	AF488	Injection	Mouse (xenograft)	Confocal laser endomicroscopy	2013

76% specificity for detection of high-grade dysplasia and 97% specificity for esophageal adenocarcinoma and most importantly to distinguish previously unidentifiable flat lesions (Lee and Wang 2016). Overall, this imaging technique has demonstrated huge promise and is currently investigated in other GI areas and overtime with improvements to both endoscopic tools and molecular probes it has the potential one day to become a universally used by gastroenterologists method.

Oral Vaccines and Oral Tolerogens

Milena Nikolova-Vlahova

Oral vaccines and oral tolerogens are only possible because of the unique functions of the human GIT; however, their biological effect goes far beyond the GIT thanks to its ubiquity.

Vaccines are undoubtedly one of the major successes of modern medicine. They are generally composed of killed or attenuated causative organisms, their subunits or nucleic acid. The subunit vaccines are considered to be very selective and specific but are poorly immunogenic. Therefore, in cases when the vaccine has lower immunogenicity, its effect could be promoted with the use of adjuvants that form complexes with the immunogenic parts and ensure slower release and more prolonged exposure to the antigen. These adjuvants are divided in two types: vaccine delivery systems (emulsions, micro- and nanoparticles, immune-stimulating complexes, and liposomes) and immunostimulatory adjuvants that are nonimmunogenic and nontoxic per se but potentiate the immune response to the antigen (aluminum and its salts, oil emulsions, synthetic polynucleotides, ISCOMs, etc.) (Wang and Coppel 2008; Saroja et al. 2011; Zhu and Berzofsky 2013).

Oral vaccines represent an easy, patient-friendly and needle-free method of protection against mucosal and nonmucosal pathogens. On the other hand, oral tolerogens are antigens that reduce systemic allergic immune response to foreign antigens (Chehade and Mayer 2005). Tolerance is immune phenomenon in which the immune response decreases following repetitive stimulation with high or low grade concentrations of antigen due to alteration of the mucous and systemic immune response.

Oral Mucosal Immune Tolerance, Suppression, and Silencing

When discussing the oral immune system modulation, of both local and systemic immune response, we should elucidate the concept of immune tolerance, suppression, and silencing. Immune tolerance is immune phenomenon associated with decrease or abolishment of immune response during or after administration of low or high doses of antigen (including allergen) due to repetitive contact with the immune system cells and decrease in immune response (Chehade and Mayer 2005; Shnawa 2015). Immune tolerance can be local/mucosal or systemic. Both types are associated with clonal cell deletion and anergy, T-regulatory activity, and clonal negative selection. Mucosal immune tolerance also involves T-cell apoptosis, changes in Th3 and dendritic cell activity and secretory IgA clonal deletion.

Immune suppression is inhibition of immune response due to decreased levels and/or activation of immunocompetent cells, humoral factors, and antibodies. It can be due to inborn defects, contact with myelosuppressive substances (including cytotoxics), radiation, etc., or antagonists to natural immune mechanisms (i.e., monoclonal antibodies against T/B cells, antibodies, cytokines and their receptors) (Shnawa 2015).

Oral mucosal immune silencing (also called oral tolerance) is a mechanism of immune tolerance, related to changes in membrane expression of molecules (i.e., B7-H) on the surface of Langerhans'cells due to contact with microbial antigens and by-products (Shnawa 2015).

Oral vaccines induce immune response and oral tolerogens cause immunomodulation with alteration of immune response directed at the development of local and/or systemic immune tolerance.

Oral Vaccines

Vaccines are an effective strategy aimed of prevention of many infectious and noninfectious disease. Different routes of administration have been developed, including oral, intranasal, intracolorectal, buccal, intravaginal, subcutaneous, intradermal, intramuscular, etc. Oral vaccines ensure durable protection against mucous and nonmucous infections, including polio, rotavirus and adenovirus infection, typhoid fever. Oral vaccines against several nonmucous diseases have been developed, including malaria, Japanese encephalitis, hepatitisB, etc. but their safety, tolerability, and efficacy in humans are to be elucidated.

The immunogenicity of all oral vaccines depends on the stimulation of mucosa-associated lymphatic tissue (MALT), gut-associated lymphatic tissue (GALT), dendridic cells, secretory IgA response, and T-cells. All oral vaccines are composed of killed or attenuated microorganisms or of their immunogenic substances. The *delivery system* is the crucial link, determining the site of liberation of the immunogenic substance and the site of stimulation of mucous immune response (Wang and Coppel 2008; Saroja et al. 2011; Kraan et al. 2014). Moreover, orally administered vaccines can trigger nongastrointestinal response (urogenital or systemic) depending on the site of liberation of antigens and the delivery system.

Following delivery systems have been developed (Wang and Coppel 2008; Saroja et al. 2011; Kraan et al. 2014):

- Live bacterial vectors expressing recombinant antigens and colonizing the gut (mainly nonpathogenic *Salmonella*).
- Particulate formulations – vaccine particles are covered with substances that protect them from degradation until the antigens reach their target sites (small or large intestine) – microparticles of biodegradable polymers, cochleates, liposomes (in lipopolysaccharides and/or phospholipids), ICOMs (covered with cholesterol + QuilA), virus-like particles (VLPs). Moreover, the antigens can be incapsulated in a structure along with targeting substances. The encapsulated antigens induce stronger and site-specific immune response. Some of these techniques have been proven to be effective in animal models but their efficacy, safety, and tolerability on humans need to be further elucidated.
- Nanoparticles – these small particles of encapsulated antigens ensure safe passage of the vaccine to the site of action (usually the small intestine) and effective stimulation of the immune response – both humoral and cellular. As in microparticles, the vaccine antigens are encapsulated in polymers, starch, lipid-containing vesicles, VLPs and ISCOMs, that are dissolved and subsequently absorbed in the small intestine and exercise both their local and systemic immunogenic effect with subsequent immune response.
- Transgenic plants – genetically modified plants that express antigenic proteins expressed on the surface or within the plant cell – a cost-effective alternative of particulate formulations.
- Mucosal adjuvants – cholera toxin and heat-labile *E. coli* toxin – toxic and unsafe in humans.
- Several oral malaria vaccines have been developed. The main principles of their action are: naked bacterial antigens, oral vectors expressing plasmodium antigens (*Salmonella, Lactococcus*), encapsulated plasmodium antigens (i.e., PLGA-encapsulated SPf66).
- Oral vaccines against hepatitis B – live vectors of HBs antigen, encapsulated HBs antigen (in biodegradable capsula in the form of microparticles), expression of HBs antigen by genetically modified potatoes and lettuce.
- Oral vaccines against Japanese viral encephalitis – viral protein with CpG adjuvant and/or expressed by genetically modified *E. coli*.
- Several new vaccine delivery systems that ensure needle-free administrations of vaccines have been developed, including the MucoJet system (Miller and Greenberg 2017). The MucoJet has been tested in laboratory animals (rabbits). It uses a gas-generating system that produces a high-pressure jet of vaccine that penetrates the buccal mucosa and is capable of inducing both local and systemic immune response (Miller and Greenberg 2017).

The immune potentiators (adjuvants) that are added to orally administered vaccines in order to increase their immunogenicity are two biologically similar enterotoxins – cholera toxin and heat-labile *E. coli* toxin, and the Toll-like receptor (TLR) ligands – Poly(I:C), FSL-1, MPLA, Pam3CSK4, cytosine-phosphate guanosine, etc. (Wang and Coppel 2008; Saroja et al. 2011; Zhu and Berzofsky 2013). These immunogens activate the NK cells, NKT, B cells, CD4+ T cells in the spleen and induce stronger Th1 response after sublingual vaccine delivery. Aluminum and its salts, oil emulsions, and synthetic polynucleotides are nontoxic and nonimmunogenic but nonspecifically enhance the vaccine-triggered immune response.

Oral Tolerogens

Oral tolerogens are antigens that after repeated oral administration lead to a decrease in allergic (IgE-mediated), secretory (IgA-mediated), or systemic (IgG-mediated) immune response (Chehade and Mayer 2005).

The oral administration of antigens change both Th2 (IL-4/IL-10) and Th3 (TGF-beta) response, dendritic cells and antigen-presenting cells in general, along with CD4 + CD25+ and LAP + T cells, cytokine levels and affect other metabolic and immune pathways (retinoic acid, Tres expression transcription factor FoxP3) (Chehade and Mayer 2005; Wang et al. 2013). Moreover, the development of oral tolerance is augmented by multiple humoral factors, i.e., IL-4, IL-10, anti-IL-12, TGF-beta, cholera toxin B subunit, Flt-3 ligand, anti-CD40 ligand, and repetitive oral contact with the allergen/antigen.

Oral tolerogens are used both in allergic (i.e., allergy to cow's milk, contact allergy) and in autoimmune diseases (systemic lupus, encephalomyelitis, myasthenia, arthritis, uveitis (Chehade and Mayer 2005; Wang et al. 2013). Several researches have been performed in chronic diseases, such as atherosclerosis, asthma, stroke and colitis. In immune disease all oral tolerogens have been tested mainly in experimental animals and in humans several studies in lupus, multiple sclerosis, uveitis, myasthenia, and arthritis have been conducted.

Effect of oral tolerogens depends mainly on three factors: antigen dose and form, host factors (age, genetics, and normal flora of the host) (Chehade and Mayer 2005).

- The effect is dose-dependent – frequent administration of low doses leads mainly to suppression of the immune response (allergic or inflammatory), whereas high doses lead mainly to clonal anergy and deletion.
- The form of the antigen – soluble antigens have more pronounced tolerogenic effect than particulate. Moreover, the possibility for cross-reaction and antigen mimicry should not be forgotten.
- Host genetics – in mice, the development of tolerance to ovalbumin is dependent on MHC genes, IL-4 and IL-10 secretion, interferon-gamma secretion, etc. In humans, the presence of DRB1*08, DRB1*08/12tyr16, and DQB1*04 is associated with higher prevalence of peanut allergy.
- Host age – the allergic and immune reactions are known to be more severe in younger subjects and especially in neonates, probably due to increased intestinal permeability and lymphocytic response in this age. Therefore, the tolerogenesis is expected to have different efficacy in different age groups.
- Normal flora of the host – The development of oral tolerance is highly dependent on the bacterial colonization of the host's GIT. The concomitant administration of *Lactobacillus spp.* along with oral desensitization to cow's milk leads to more effective tolerogenesis. Therefore, the normal bacterial flora promotes the development of normal oral tolerance.

Oral tolerance has been tested in humans in autoimmune diseases, such as multiple sclerosis, lupus, arthritis, uveitis, contact sensitivity. The studies in systemic lupus failed to demonstrate long-term efficacy of abetimus (LJP 394) for the activity of systemic lupus and lupus nephritis. Studies in humans are ongoing for other immune disease.

References

Bailey DL, Townsend DW, Valk PE, Maisey MN (2005) Positron emission tomography in clinical medicine. Positron emission tomography, 1st edn. Springer, New York, pp 1–12

Chehade M, Mayer L (2005) Oral tolerance and its relation to food hypersensitivity. J Allergy Clin Immunol 115(1):3–12

Dasgupta A (2012) Chapter 1 – introduction to therapeutic drug monitoring: frequently and less frequently monitored drugs. In: Dasgupta A (ed) Therapeutic drug monitoring, 1st edn. Elsevier, London, pp 1–29

Fathallah AM, Bankert RB, Balu-Iyer SV (2013) Immunogenicity of subcutaneously administered therapeutic proteins - a mechanistic perspective. AAPS J 15(4):897–900

Gershon MD (2013) 5-Hydroxytryptamine (serotonin) in the gastrointestinal tract. Curr Opin Endocrinol Diabetes Obesity 20(1):14–21

Holzer P (2009) Opioid receptors in the gastrointestinal tract. Regul Pept 155:11–17

Izzo AA, Sharkey KA (2010) Cannabinoids and the gut: new developments and emerging concepts. Pharmacol Ther 126:21–38

Kopylov U, Ben-Hurin S, Seidman E (2014) Therapeutic drug monitoring in inflammatory bowel disease. Ann Gastroenterol 27:304–312

Kraan H, Vrieling H, Czerkinsky C et al (2014) Buccal and sublingual vaccine delivery. J Control Release 190:580–592

Lagier J-C, Million M, Hugon P et al (2012) Human gut microbiota: repertoire and variations. Front Cell Infect Microbiol 2(136):1–19

Lee JH, Wang TD (2016) Molecular endoscopy for targeted imaging in the digestive tract. Lancet Gastroenterol Hepatol 1:147–155

Miller CS, Greenberg RN (2017) MucoJet: a novel oral microjet vaccination system. Oral Dis. https://doi.org/10.1111/odi.12697

Murphy CJ, Swanson B, Markow M et al (2017) Lymphocytic gastritis identified by abnormal PET scan. ACG Case Rep J 4:e62

Pinto-Alphandary H, Andremont A, Couvreur P (2000) Targeted delivery of antibiotics using liposomes and nanoparticles: research and applications. Int J Antimicrob Agents 13:155–168

Poeppel TD, Krause BJ, Heusner TA et al (2002) PET/CT for the staging and follow-up of patients with malignancies. Eur J Radiol 70:382–392

Rossen NG, MacDonald JK, de Vries EM et al (2015) Fecal microbiota transplantation as novel therapy in gastroenterology: a systematic review. World J Gastroenterol 21(17):5359–5371

Saroja CH, Lakshmi PK, Bhaskaran S (2011) Recent trends in vaccine delivery systems: a review. Int J Pharm Investig 1(2):64–74

Savarino E, Chianca V, Bodini G et al (2017) Gadolinium accumulation after contrast-enhanced magnetic resonance imaging: which implications in patients with Crohn's disease? Dig Liver Dis 49:728–730

Shnawa IMS (2015) Oral mucosal immune tolerance versus oral immune silencing: mini review. Am J Biomed. Life Sci 3(4–1):7–9

Tiwari G, Tiwari R, Sriwastawa B et al (2012) Drug delivery systems: an updated review. Int J Pharm Investig 2(1):2–11

Van Cutsem E, Verheul HMW, Flamen P et al (2016) Imaging in colorectal cancer: progress and challenges for clinicians. Cancers 8:1–14

Wang L, Coppel R (2008) Oral vaccines delivery: can it protect against non-mucosal pathogens? Expert Rev Vaccines 7(6):729–738

Wang X, Sherman A, Liao G (2013) Mechanisms of oral tolerance induction to therapeutic proteins. Adv Drug Deliv Rev 65(6):759–773

Zhu Q, Berzofsky JA (2013) Oral vaccines. Gut Microbes 4(3):246–252

Zuckerman JN (2000) The importance of injecting vaccines into muscle: Different patients need different needle sizes. Br Med J 321(7271):1237–1238

Pharmacodynamic Evaluation: Endocrinology

12

Michael A. B. Naafs

Contents

M. A. B. Naafs (✉)
Naafs International Health Consultancy, Oldenzaal, Netherlands
e-mail: naafs.healthconsultancy@gmail.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_35

Abstract

In this chapter, pharmacodynamic evaluations for the most common clinical endocrine disorders are discussed.

Special attention is given to new classes of oral antidiabetics and new sophisticated insulin analogues.

Clinical pharmacology in endocrine disorders is reviewed along the classical hypothalamus-pituitary-adrenal-gonadal axis, while thyroid disorders are viewed separately.

There is a focus on the hormone receptor interaction similarities and the drug-ligand-receptor binding.

The concept of hormone agonists and antagonists has a prominent place.

Neuroendocrinology pharmacodynamic evaluation is not included in this chapter.

Introduction

Endocrinology is the communication science in internal medicine.

Communication takes place by "classical" hormones and paracrine and intracrine mechanisms and by neuroendocrine signals. Hormones are classified traditionally as amines, peptides, proteins, and steroids.

Endocrinology focuses primarily on the endocrine organs that secrete hormones as the pituitary, thyroid, adrenals, ovaries, testes, and pancreas.

Endocrine diseases are disorders of deficiency, excess, or end-organ resistance to one or more hormones. Most endocrine diseases are chronic diseases that need lifelong treatment. Diabetes mellitus is one of the most common as is hypothyroidism and metabolic syndrome, including obesities and dyslipidemia.

Hormones exert their effect in the body by binding to a hormone-receptor complex. These are a wide family of proteins made up of receptors located on the cell surface as insulin and steroid hormone receptors or in the cytoplasma, so-called intracellular or nuclear receptors, used, for example, by testosterone.

After hormone binding to the receptor complex, several pathways can be signaled and activate the target cells. This process frequently shows similarities with the process of ligand (drug), receptor, and ligand-receptor interaction in pharmacodynamics (Salahudeen and Nishtala 2017).

In this chapter, the latest pharmacodynamic developments for the most common endocrine disorders will be discussed.

Diabetes Mellitus

Pharmacodynamics in DM can be studied at a population level, a mechanism-based mode, a metric model, or a drug-targeted-mediated pharmacodynamic model (Salahudeen and Nishtala 2017). In this section, the clinical pharmacodynamics of the most used oral antidiabetics, insulins, and its analogues will be viewed shortly.

Metformin-Biguanides

Metformin improves glucose tolerance in patients with type 2 DM lowering both basal and postprandial plasma glucose by decreasing hepatic glucose production, intestinal absorption of glucose, and insulin sensitivity by increasing peripheral glucose uptake and utilization (Thonos and Gregg 2017).

Metformin does not produce hypoglycemia under normal circumstances. Absolute bioavailability is around 50–60%. Increasing doses results in decreased absorption. Food delays the absorption of metformin. Peak plasma levels are achieved at a median 7 h. Metformin has a negligible protein bound and is excreted unchanged by tubular secretion in the urine. No hepatic metabolites have been found. The elimination half-life is approximately 17 h. Plasma half-life is 6 h. Metformin uses the erythrocyte mass as a compartment of distribution.

In elderly the total plasma clearance of metformin is delayed due to an age-related decline in renal function. No pharmacokinetic studies of metformin have been done in patients with hepatic insufficiency.

Metformin is contraindicated in severe renal impairment (eGFR below 30 ml/min.), known hypersensitivity to metformin, and acute or chronic metabolic acidosis, including diabetic ketoacidosis. Elderly, debilitated, or malnourished patients and those with adrenal or pituitary insufficiency are susceptible to the hypoglycemic effects. Beta-adrenergic blockers can mask this effect.

Lactic acidosis can be an infrequent complication in metformin users. Metformin decreases liver uptake of lactate. Risk factors are renal impairment; use of carbonic anhydrase inhibitors, topiramate, and radiologic contrast agents; hypoxia; and alcohol intoxication.

In a minority of patients, metformin can lead to vitamin B12 depletion, mostly without clinical symptoms. In these cases, metformin interferes with vitamin B12 absorption from the B12-intrinsic factor complex.

Sulfonylureas

Sulfonylureas are insulin secretagogues which means they work by causing the body to secrete insulin (Riddle 2017). Sulfonylureas bind to a channel of proteins in the pancreas. This is an ATP-potassium channel. Glucose-generated ATP is the ligand in the pancreatic beta cell to produce insulin.

Members of this drug class are glimepiride (Amaryl), glibenclamide (Daonil), gliclazide (Diamicron), glipizide (Glibenese), and tolbutamide (Rastinon). Another class of diabetes drugs which work in this way is the prandial glucose regulators, the meglitinides.

All sulfonylureas are absorbed by the intestine, each one with its specific absorption and bioavailability. After absorption, sulfonylureas bind almost completely to plasma proteins on an average of 95%. The volume of distribution is about 0.2 l/kg.

The biological effect of sulfonylureas lasts much longer than their plasma half-life because of receptor interaction and formation of active metabolites persisting 24 h or more. Moreover their half-life is prolonged in renal failure.

Genetic differences can also change the response to sulfonylureas. Some of these gene polymorphisms were identified in the genes encoding the potassium-ATP channel (KNCJ11 andABCC8). These mutations cause a change in insulin secretion and insulin response to treatment.

Most sulfonylureas are characterized by renal excretion. Gliclazide and above all cliquidone

show a predominant biliary clearance. Sulfonylureas lower blood glucose by 20% and HbA1C by 1–2%.

The most common side effect is hypoglycemia, sometimes lasting for hours and requiring hospital admission. Undesired weight gain due to increasing insulin secretion is around 2 kg. Sulfonylureas act directly on beta cells leading to progressive worsening of DM at the end. This phenomenon is called "secondary failure." Since ATP-potassium-dependent channels are present in cardiac cells and coronary vessels, sulfonylureas if present at the time of a myocardial infarction may impair adequate vasodilatation, resulting in a greater area of myocardial damage (Riddle 2017).

Alpha-Glucosidase Inhibitors

Alpha-glucosidase inhibitors (AGIs) sometimes referred as starch blockers are antidiabetic medicines that help to reduce post-meal blood glucose levels (Laar van de 2008).They do not have a direct effect on insulin secretion or sensitivity. They work by slowing down the digestion of carbohydrates found in starchy foods. Examples of AGIs include acarbose (Glucobay) and miglitol (Glyset). AGIs are normally used as a single treatment but can be taken in combination with sulfonylureas. AGIs slow down digestion by blocking enzymes in the small intestine that break down carbohydrates.

Pharmacokinetic evaluation of acarbose is difficult because only 2% is absorbed. Acarbose reversely binds to pancreatic-alpha-glucoside hydrolases. Those enzymes inhibit hydrolysis of complex starches to oligosaccharides. Acarbose raises the glucagon-like peptide (GLP-1) response due to inhibiting gastric emptying. The use of AGIs is limited by gastrointestinal side effects as meteorism, flatulence, and diarrhea. Unexplained severe liver function test disturbances have been reported.

Thiazolidinediones

The thiazolidinediones are also called TZDs or glitazones. TZDs work by targeting the PPAR-gamma receptor which activates a number of genes in the body and plays an important role how the body metabolizes glucose and fat (Marathur et al. 2016). TZDS can therefore help boost insulin sensitivity. Pioglitazone (Actos) is the most known member of this group. Peroxisome proliferate-activated receptors (PPARs) are structural similar to steroid or thyroid hormone receptors and are subcellular organelles found in most plant and animal cells.

TZDs are high-affinity ligands for PPAR-gamma, a nuclear receptor. PPAR-gamma has been known to regulate adipocyte differentiation, fatty acid storage, and glucose metabolism and is a target of antidiabetic drugs. PPAR-gamma agonists improve insulin resistance by opposing the effect of tumor necrosis factor-alpha (TNF-alpha) in adipocytes.

Pioglitazone can be used in type 2 DM alone or in combination with metformin, a sulfonylurea or in combination with insulin. Pioglitazone has been linked to a significant risk of bladder cancer in long-time users (Marathur et al. 2016).

DPP-4 Inhibitors

DPP-4 inhibitors are also called gliptins. Gliptins work by blocking the action of dipeptidyl peptidase-4 (DPP-4), an enzyme which destroys a group of gastrointestinal hormones called the incretins (Hippisley-Cox and Compland 2016; Karagiannis et al. 2012). Incretins stimulate the production of insulin after eating and reduce the production of glucose by the liver during digestion by its effect on glucagon.

Oral glucose stimulates the release of the endogenous incretins glucagon-like peptide (GLP-1) and glucose-dependent insulin-releasing peptide (GIP).The incretin effect is diminished in type 2 DM. Drugs in this class include sitagliptin (Januvia), vildagliptin (Galvas), and saxagliptin (Onglyza). DPP-4 inhibitors have been linked with an increased risk of pancreatitis.

GLP-1 Agonists

The GLP-1 agonists are also known as the incretin mimetics. These drugs work by mimicking the functions of the natural incretin hormones that

help lower post-meal blood sugars (Lovshin 2017).

GLP-1 agonists stimulate the release of insulin and inhibit the release of glucagon and slow glucose absorption by slowing gastric emptying. The GLP-1 receptor is a cell surface receptor which internalizes after stimulation and exerts its effect by the second messenger adenylyl cyclase. The GLP-1 receptor is expressed in pancreatic beta cells and also in the brain where it is involved in the control of appetite. The GLP-1 receptor binds glucagon-like peptide (GLP-1) and glucagon as its natural agonist.

Members of this drug class are exenatide (Byetta), lixisenatide (Lyxumia), dulaglutide (Trulicity), and liraglutide (Victoza).

SGLT-2 Inhibitors

SGLT-2 inhibitors are a new class of type 2 DM medications They are also called the glifozins. They block the reabsorption of glucose in the kidney, increase glucose excretion, and lower blood glucose levels (Zou et al. 2017).

SGLT-2 is a low-affinity, high-capacity glucose transporter located in the proximal tubule in the kidneys. The pharmacodynamic response to SGLT-2 inhibitors as assessed by urinary glucose excretion declines with increasing severity of renal impairment. SGLT-2 inhibitors have a rapid oral absorption, a long elimination half-life, and an extensive hepatic metabolism mainly via glucuronidation to inactive metabolites and a low renal excretion as a parent drug.

Drugs in this class are dapagliflozin (Forxiga) and empagliflozin (Jardiance). Adverse effects are yeast infections, urinary tract infections, and diabetic ketoacidosis.

Prandial Glucose Regulators

Prandial glucose regulators or "glinides" are insulin secretagogues working similar like sulfonylureas by the ATP-potassium channel but at a different site. Unlike sulfonylureas, they have a rapid onset but relatively short effect. Drugs in this class are repaglinide (Prandin) and nateglinide (Starlix) (Mondoza et al. 2013).

Amylin Analogues

Amylin analogues or agonists are injectable drugs that work similar to the hormone amylin and can be used in both type 1 and type 2 DM (Marathur et al. 2016). Amylin is released by the pancreas at the same time as insulin but in much smaller quantities, about 1% compared to insulin.

It inhibits the release of glucagon, slows food emptying from the stomach, and curbs appetite. Pramlintide acetate (Symlin) is the best known member of this drug class. Symlin is unbound in plasma. The half-life is around 50 min.

Insulin Types and Forms

Until the 1980s, animal insulin was the only treatment for insulin-dependent DM. Nowadays, largely human insulins and human insulin analogues are used. These are rapid, short-acting, intermediate, and long-acting insulins (Lispkak et al. 2017).

Examples of rapid-acting insulins are insulin lispro (Humalog) and insulin aspart (Novorapid). Action starts at 15–20 min after s.c. injection and lasts 2–5 h.

Short-acting insulins are Actrapid, Humulin S, and Velosulin. Action starts at 20 min and lasts 6–8 h.

Intermediate insulins are represented by the premixed insulin Humulin I, a human insulin made up of 30% short-acting (neutral) insulin and 70% intermediate-acting (isophane) insulin. It has a peak activity between 1 and 8 h with a duration of action lasting to 22 h.

Long-acting insulins have no peak activity which allows for a basal delivery through the day. Examples of long-acting insulins are insulin glargine (Lantus), insulin detemir (Levemir) or insulin degludec (Tresiba). Lantus has a consistent activity of 24 h. The duration of Levemir is slightly shorter than Lantus and therefore is often injected twice daily. Tresiba has an action duration of more than 42 h.

In addition, there are also premixed human insulin analogues which combine a rapid and a long-acting insulin. Examples are Humalog Mix 25, Humalog Mix 50, and Novomix 30.

All insulins use the insulin receptor, a transmembrane tyrosine kinase receptor, acting by the phosphorylation pathway. Ligands for this receptor are insulin; IGF-1 (insulin growth factor); IGF-2; the relaxin peptides 1, 2, and 3; and the insulin-like peptides 3–6.

There are important pharmacokinetic and pharmacodynamic differences between the long-acting and rapid-acting insulin analogues (Mondoza et al. 2013). These depend on the site of injection, concentration of the insulin formula, volume of the injected dose, depth of injection, thickness of the subcutaneous fat layer, exercise, local massage, heat exposure, and finally intrinsic properties of the insulins (Lispkak et al. 2017).

Insulin is usually cleared by receptor-mediated uptake and intracellular degradation. The main site of plasma extraction is the liver with smaller contributions by adipose tissue and muscle.

Glucagon

Glucagon is produced by the alpha cells and raises the blood glucose. Glucagon and insulin are part of a feedback system that keeps blood glucose levels stable.

Glucagon belongs to the secretin families of hormones. Blood glucose is elevated by promoting gluconeogenesis and glycogenolysis. Glucagon also regulates the role of glucose production through lipolysis. Glucagon induces lipolysis under conditions of insulin suppression such as type 1 DM.

The glucagon receptor is a membrane G-protein-coupled receptor using the adenylate cyclase system as a second messenger. Secretion of glucagon is merely stimulated by hypoglycemia and epinephrine and inhibited by insulin and somatostatin.

Glucagon (GlucaGen) is used as an emergency medicine to treat severe hypoglycemia in diabetic patients treated with insulin, who have passed out or cannot take some form of sugar by the mouth. Alternatively, an epinephrine emergency kit (Epipen) can be used in these circumstances (Posner and Camarga 2017).

Abnormally elevated levels of glucagon may be caused by pancreatic tumors, such as glucagonoma, which include necrolytic erythema migrans. It may occur alone or in the context of the genetic multiple endocrine neoplasia (MEN) type 1 syndrome.

Pituitary-Hypothalamus

The pituitary gland produces various hormones. In the anterior pituitary gland, ACTH, TSH, LH, FSH, PRL, GH, and MSH are produced, which act on different target glands or cells. The posterior pituitary produces ADH and oxytocin. The hypothalamus releases ADH, CRH, GnRH, GHRH, GHIH (somatostatin), oxytocin, PRH or PIH (dopamine), and TRH.

ACTH

ACTH (Synachten) is used as a diagnostic aid in the assessment of suspected adrenocortical hypofunction, Addison's disease. The binding sites of ACTH are located in the adrenal cortex where it becomes bound to a specific receptor. By activating the adenylate cyclase pathway, the pregnenolone is synthesized from cholesterol. From pregnenolone various corticosteroids are formed.

Pharmacodynamics are measured as the cortisol response at 0, 30, and 60 min after intravenous administration of 250 microgram synachten.

Growth Hormone

Human growth hormone (GH) is secreted by somatotrope cells in the anterior pituitary in a pulsatile fashion. The secretion is regulated by two hypothalamic peptides, growth hormone-releasing hormone (GHRH) which stimulates GH secretion and somatostatin which inhibits GH secretion by backregulation.

The growth hormone receptor belongs to the family of transmembrane proteins that includes the prolactin receptor. Signal transduction is by tyrosine phosphorylation.

Owing to the fact that GH has a short half-life, several approaches have been taken to create long-term agonists. These include the pegylation-sustained release formulations and ligand-receptor fusion proteins. Pegylation of a GH analogue (pegvisomant, Somavert) forms the basis of a successful treatment of acromegaly (Freda et al. 2015).

GH receptor expression can be modified by insulin, thyroid hormones, and sex hormones. Pharmacodynamic response to GH analogue injection in GH-deficient children is measured by plasma GH levels, IGF-1 (insulin-like growth factor), glucose, and free fatty acid (FFA) levels.

Synthetic human growth hormone is used in children with HGH deficiency or insufficiency, children born small for gestational age, and girls with Turner syndrome, Prader-Willi syndrome, and chronic kidney disease. In adults, the use of HGH includes HGH deficiency to rare pituitary tumors or their treatment, short bowel syndrome, or muscle wasting associated with HIV/AIDS.

However, the most common use is doping in combination with other performance drugs as anabolic steroids in an attempt to build muscle and improve athletic performance. Yet HGH's effect on athletic performance is unknown.

Members of the somatotropin drug class are Norditropin, Nutropin, Humatrope, Genotropin, and Saizen.

Prolactin

Prolactin (PRL) is a hormone that promotes lactation in mammals. Prolactin is produced in the front portion of the pituitary. Production of PRL is controlled by two hormones, dopamine and estrogen. Dopamine inhibits and estrogen increases PRL production.

Hyperprolactinemia leads to menstrual disturbances, estrogen and testosterone deficiency reproduction, and disease. Naturally occurring GnRH has a half-life of 2–4 min. GnRH agonists have substitutions for glycine which significantly increases the plasma half-life (Niamh 2013).

GnRH agonists can be used in pulsatile or continuous regimen to treat estrogen-dependent conditions as endometriosis, uterine leiomyomas, precocious puberty, and menorrhagia. GnRH analogues are used extensively during in vitro fertilization cycles to prevent an LH surge and allow for retrieval of the mature oocytes (Singh et al. 2014).

The GnRH receptor belongs to the G-protein-coupled transmembrane intracellular receptor family. They are also present in the gonads.

In men GnRH agonists are used in the treatment of prostate carcinoma by reducing the levels of testosterone (Shipley et al. 2017).

Members of this drug class are goserelin (Zoladex) and leuprorelin (Lupron). Pharmacodynamic response is measured by suppressed testosterone and PSA (prostate-specific antigen) levels.

In women GnRH antagonists are used in infertility treatment. Members of this drug class are ganirelix acetate (Ganirelix) and cetrorelix (Cetrotide) (37).

Pharmacodynamic response is measured by LH, FSH, and E2 concentrations.

Somatostatin

Somatostatin is also known as growth hormone-inhibiting hormone (GHIH). Somatostatin inhibits insulin and glucagon secretion. Somatostatin is produced by the hypothalamus and uses the membrane G-protein-coupled receptor and the adenylate cyclase system as a second messenger.

Somatostatin is also produced by the delta cells in the pyloric antrum, the duodenum, and the pancreatic islets. Somatostatin inhibits the release of growth hormone by opposing growth hormone-releasing hormone (GHRH), inhibits TSH, and inhibits the release of prolactin (PRL). It further inhibits the release of gastrin, cholecystokinin, motilin, vasoactive intestinal polypeptide (VIP), gastric inhibitory polypeptide (GIP), and enteroglucagon. It also suppresses the exocrine action of the pancreas.

Octreotide (Sandostatin) is a synthetic somatostatin analogue used in the treatment of acromegaly, carcinoid syndrome, insulinomas, glucagonomas,

and the VIPomas (Freda et al. 2015; Der-Nigoghossian et al. 2017).

Melanocyte-Stimulating Hormones

The melanocyte-stimulating hormones, alpha, beta, and gamma MSH, are collectively known as the melanotropins. MSH is produced in the hypothalamus. Acting in the hypothalamus, alpha MSH suppresses appetite and contributes to sexual arousal. Keratinocytes in the skin use the melanocortin receptor to produce melanin.

An increase in MSH will cause darker skin. MSH increases in humans during pregnancy. This causes increased pigmentation in pregnancy. MSH and ACTH share the same precursor molecule proopiomelanocortin (POMC).

For these reasons, patients with Cushing's disease, due to excess ACTH, can have hyperpigmentation as acanthosis nigricans. Patients with primary Addison's disease can have hyperpigmentation too.

Different levels of MSH are not the major cause of racial variation in skin color. However, in many red-headed people and other people who do not tan well, there are variations in their hormone receptors, causing them not to respond to MSH in the blood.

Alpha-MSH analogues as Melanolux are primarily used for their tan-stimulating effect and in erythropoietic porphyria. Melatonin is an alpha-MSH antagonist and is produced in the hypofyse from tryptophan precursors. Synthetic melatonin (Bio-Melatonin) is used in insomnia and jet lag (De Leo et al. 2016).

Oxytocin

Oxytocin is a hormone that causes contractions during labor and helps shrink the uterus after delivery. Oxytocin orders the body to let down milk when the baby suckles. It is also known as the "cuddle hormone" because it is released when people snuggle up or bond socially (Zanos et al. 2017).

The oxytocin receptor is a member of the G-protein-coupled receptor family. Oxytocin receptors are expressed by the myoepithelial cells of the mammary and in both the myometrium and endometrium at the end of pregnancy. Oxytocin receptors are also present in the central nervous system.

Oxytocin is used in inducing labor in problematic pregnancies or in helping to abort the fetus in cases of incomplete abortion or miscarriage.

Oxytocin (Pitocin) is usually used as an intravenous infusion. Pitocin is also used as a nasal spray in the treatment of autism.

Pharmacodynamics of oxytocin are measured by clinical outcomes. After prolonged exposure, desensitization of oxytocin receptors can occur.

Antidiuretic Hormone

Antidiuretic hormone (ADH) or vasopressin acts on the kidney and blood vessels. Vasopressin helps prevent loss of water from the body by reducing urine output and helping the kidneys reabsorb water. Vasopressin is used to treat diabetes insipidus which is caused by a lack of naturally occurring pituitary hormone (Christ-Crain and Fenske 2016).

Vasopressin receptors belong to the G-protein-coupled receptor family. They are located in the basolateral membrane of the kidney collecting ducts, pituitary gland, and vascular smooth muscle.

Vasopressin is used as an intravenous infusion (Pitressin) or as a nasal spray (desmopressin).

Vasopressin antagonists (VRAs) are drugs that block vasopressin receptors. Most commonly VRAs are used to treat hyponatremia caused by the syndrome of inappropriate secretion of antidiuretic hormone (SIADH), congestive heart failure (CHF), and cirrhosis (Alagiakrishnan 2016).

Members of this drug class are the "vaptans" as conivaptan (Vaprisol) and tolvaptan (Jinarc). Pharmacodynamic response can be measured by serum electrolytes and serum and urine osmolarities (Streeten et al. 2017).

The Thyroid Gland: Hyper- and Hypothyroidism

The thyroid gland controls metabolism, the way your body uses energy. The thyroid uses iodine in your food to make two hormones,

triiodothyronine (T3) and thyroxine (T4). The hypothalamus and the pituitary communicate to maintain T3 and T4 balance. The hypothalamus produces TSH-releasing hormone (TRH) to tell the pituitary to release thyroid-stimulating hormone (TSH).

When T3 and T4 levels are low in the blood as in hypothyroidism, the pituitary releases more TSH to order the thyroid gland to produce more T3 and T4. If T3 and T4 levels are high as in hyperthyroidism, the pituitary gland releases less TSH to the thyroid gland to slow the production of these hormones (Campbell et al. 2015).

The thyroid hormone receptor is a nuclear retinoid X receptor.

Symptoms of hyperthyroidism are anxiety, irritability, moodiness, palpitations, sweating, eye symptoms as exophthalmus, hand trembling, hair loss, weight loss, and missed or light menstrual periods.

Symptoms of hypothyroidism are trouble in sleeping, tiredness and fatigue, difficulty in concentrating, dry coarse skin, weight gain, depression, sensitivity to cold temperature, frequent heavy periods, and joint and muscle pain.

Hyperthyroidism can be treated with antithyroid medication, thyreostatics, primarily methimazole (Strumazol) and propylthiouracil (PTU). Another option is radioactive iodine or surgery.

Methimazole

Thiamazole inhibits the enzyme thyroperoxidase which normally acts in thyroid hormone synthesis by oxidizing the anion iodide (I-) to iodine I2, hypoiodous acid (HOI), and enzyme-linked hypoiodite (EOI) facilitating iodine's addition to tyrosine residues on the hormone precursor thyroglobulin (TBG) to synthesize T3 and T4. It does not inhibit the sodium-dependent iodide transporter in the thyroid follicular cells. Inhibition of this step requires competitive inhibitors as perchlorate or thiocyanate.

Potassium perchlorate is used in the treatment of amiodarone (Cordarone), an iodide-containing cardiac arrhythmic-induced thyrotoxicosis.

Pharmacodynamic response of the thyreostatics are measured by T4 and TSH levels.

Propylthiouracil

Propylthiouracil (PTU) is mainly used in patients with hypersensitivity or allergic reactions to methimazole. PTU inhibits the conversion of T4 to T3 in peripheral tissues and may therefore be an effective treatment for thyroid storm.

Hypothyroidism is treated with levothyroxine (Thyrax) or triiodothyronine (Cytomel). Cytomel is used in patients allergic to L-thyroxine.

Pharmacodynamic responses are measured by serum T4, T3, and TSH levels.

The Parathyroids

Parathyroid hormone (PTH) is secreted by the chief cells of the parathyroid glands and acts to increase the concentration of ionic calcium. In the blood calcitonin, a hormone produced by the C cells of the thyroid decreases ionic calcium concentration. The PTH receptor is located in the bone and the kidney.

PTH reduces the reabsorption of phosphate in the proximal tubulus of the kidney and stimulates 1,25 dihydroxyvitamin D, the active vitamin D metabolite, from 25-hydroxyvitamin D by its effect on renal 1 alpha-hydroxylase. PTH increases calcium absorption in the intestine in conjunction with vitamin D3.

PTH is important in bone remodeling and has a direct stimulating effect on the osteoblast and an indirect inhibiting effect on the osteoclast.

PTH uses the G-protein-coupled protein receptor with the adenylyl cyclase system as the second messenger.

Hyperparathyroidism, the presence of excess PTH in the blood, occurs in two distinct clinical circumstances. Primary hyperparathyroidism is due to abnormal levels of PTH from the parathyroid glands by either hyperplasia or an adenoma. Secondary hyperparathyroidism is due to an inappropriate high PTH level seen as a physiological response to hypocalcemia, as seen in renal insufficiency and vitamin D deficiency.

Hypoparathyroidism is a decreased function of the parathyroids with low PTH levels leading to hypocalcemia, causing cramping and twitching of muscles or even tetany. The

condition can be inherited, but it is also encountered after thyroid or parathyroid surgery. Teriparatide injection can be used as a treatment. Calcium replacement and vitamin D can ameliorate the symptoms but can increase the risk of kidney stones and chronic kidney disease (Black and Rosen 2016).

There are a number of rare but well-described genetic conditions affecting PTH metabolism, including pseudohypoparathyroidism, familial hypocalciuric hypercalcemia (FHH), and autosomal dominant hypocalciuric hypocalcemia.

In osteoporotic women, administration of the exogenous 1-34 PTH analogue teriparatide (Forsteo) by daily s.c. injection in conjunction with estrogens produced increases in bone mass and reduced vertebral and non-vertebral fractures. The intact1-84 PTH analogue Natpara is also used in the treatment of osteoporosis. These analogues can be used in sequential courses with anti-resorptive bone agents as the bisphosphonates (Loriaux 2017; Rossi et al. 2017). The bisphosphonate drug class includes APD (Pamidronate), alendronate (Fosamax), risedronate (Actonel), and ibandronate (Boniva).

Bisphosphonates inhibit the digestion of bone by stimulating the osteoclast to undergo apoptosis or cell death. While bone formation and bone resorption are normally coupled and are in balance, the net effect will be increased bone formation.

Calcitonin (Miacalcin Nasal) has a modest place in the treatment of osteoporosis, because more effective drugs as bisphosphonates are available in the prevention of bone loss.

The Adrenals

The adrenal cortex produces glucocorticoids, mineralocorticoids, androgens, estrogens, and progestins.

Glucocorticoids have a broad physiological role in the regulation of glucose metabolic pathways, stress, and modulation of the immune system. Mineralocorticoids are key regulators of mineral and water balance.

Cholesterol is the precursor to all steroid biosynthesis and is converted to a variety of steroid molecules in a series of reactions catalyzed by several cytochrome P450 enzymes. The vast majority of cholesterol is taken up from the LDL cholesterol pool.

Cortisol levels are highest in the morning and are increased by stress or severe infection. Too much cortisol from any cause lead to Cushing's syndrome, the symptoms and signs of which include fat redistribution to the face, upper back, and abdomen, weight gain, stretch marks, bruising, extra hair growth, irregular periods in women, loss of muscle, and trouble in sleeping (Posner and Camarga 2017).

Too little cortisol is part of the syndrome called Addison's disease marked by low energy, joint and abdominal pain, weight loss, diarrhea, fever, and electrolyte disturbances. If the adrenals make too little cortisol, ACTH levels rise. If the pituitary is not working, both ACTH and cortisol levels are low.

The glucocorticoid receptor is a G-protein-coupled nuclear receptor. The receptor is expressed in almost every cell of the body.

Glucocorticosteroids are used as classical glucosteroids or as glucocorticosteroid (GR) agonists that can be divided in selective and nonselective agonists. Classical glucocorticoids are hydrocortison, prednisone, prednisolone, and methylprednisolone.

Beclomethasone is a potent nonselective glucocorticoid agonist. Fluticasone is a highly selective GR agonist.

Mifepristone and ketoconazole, an antimycoticum, are GR antagonists used in the treatment of Cushing's syndrome. Mifepristone has a fourfold higher affinity for the glucocorticoid receptor than dexamethasone being in essential an antiprogestin. For this reason, it is used to abort early pregnancies (Singh et al. 2014).

The primary mineralocorticoid is aldosterone. Aldosterone is involved in the retention by sodium by active reabsorption in the collecting tubule of the kidney, while potassium is actively secreted by the collecting tubule, and water is passively reabsorbed. Aldosterone production is under the influence of ACTH. The mineralocorticoid receptor is located in the cytosol in the cell. Aldosterone and cortisol have similar affinity to the mineralocorticoid receptor.

Primary hyperaldosteronism, Conn's syndrome, is caused by either adrenal hyperplasia or by an adrenal adenoma. This results in hypertension and edema due to excessive sodium and water retention and accelerated excretion of potassium ions (Steinman et al. 2017).

Secondary hyperaldosteronism is caused by extra adrenal stimuli, such as renal hypoperfusion, which stimulates the renin-angiotensin-aldosterone system (RAAS) with resultant hypersecretion of aldosterone and edematous disorders as congestive heart failure cirrhosis with ascites and nephrotic syndrome. Causes of reduced renal blood flow include obstructive renal artery disease (atheroma, fibromuscular dysplasia of the renal artery).

Important antimineralocorticoids are spironolactone (Aldactone) and amiloride (Midamor).

Mitotane (Lysodren) is a steroidogenesis inhibitor used in the treatment of Cushing's syndrome and adrenocortical carcinoma. It inhibits the enzymes 11-beta-hydroxylase, 18-beta-hydroxylase, and 3-beta-hydroxysteroidgenase.

Epinephrine

Epinephrine also known as adrenalin is a hormone, a neurotransmitter, and a drug.

It plays a role in the fight-or-flight response by increasing blood flow to muscles, output of the heart, pupil dilation, and blood sugar. It does by binding to adrenergic alpha- and beta-receptors.

As a medication, it is used in anaphylaxis, cardiac arrest, and superficial bleeding by i.m. or i.v. administration. Inhaled epinephrine may be used in croup and asthma. Epinephrine is widely used as a nasal decongestant. In anaphylaxis, it is used as an Epipen Autoclick. Side effects include anxiety, shakiness, sweating, and sometimes broad complex ventricular tachycardia (Kuiper et al. 2016).

Norepinephrine is a sympathomimetic drug. When given by i.v. injection, it increases heart rate and force and constricts blood vessels making it very useful in the treatment of shock. Sympathomimetic and sympatholytic drugs mimic or block the effects of norepinephrine. These are called beta- and alpha-blockers.

Beta-blockers are used in the treatment of hypertension, atrial fibrillation, angina pectoris, congestive heart failure, and performance anxiety (Bangalore 2017). Members of this drug class are atenolol (Tenormin), acebutolol (Sectral), metoprolol (Lopressor), pindolol (Viskeen), and propranolol (Inderal).

Alpha-blockers are used in the treatment of hypertension and benign prostate hyperplasia through their relaxing effect on the muscles of the neck of the bladder and help in the expulsion of bladder stones. They are also used in the treatment of anxiety disorders, panic disorders, and posttraumatic stress disorders (PTSD).

Alpha-blockers include doxazosin mesylate (Cardura), prazosin hydrochlorie (Minipress), and dutasteride (Avodart).

Alpha-1-adrenoceptor agonists act as systemic vasoconstrictors. Members of this class are nose and eye drops (Neo-Synephrine, Mydrin). Alpha-2-adrenoceptor agonists are antihypertensive drugs such as clonidine (Catapresan) and alpha-methyldopa (Aldomet).

Pheochromocytomas are rare adrenal medulla tumors oversecreting the catecholamines epinephrine and norepinephrine. Sometimes they are part of a genetic syndrome of multiple endocrine neoplasia (MEN) type 2 syndrome.

The Renin-Angiotensin-Aldosterone System

The renin-angiotensin-aldosterone system (RAAS) is a hormone system that is involved in the regulation of the plasma sodium concentration and arterial blood pressure. When the plasma sodium concentration is lower than normal or the renal blood flow is reduced, the juxtaglomerular cells in the kidney convert prorenin, an intracellular protein, into renin. Plasma renin then cuts a short ten-amino acid-long peptide off a plasma protein known as angiotensinogen. This short peptide is then converted to angiotensin-2 by cleaving off two amino acids by the angiotensin-converting enzyme (ACE) found in the capillaries

in the body, the lungs, and the epithelial cells of the kidney.

Angiotensin-2 is a potent vasoconstrictor that causes arterioles to constrict resulting in increased arterial blood pressure. Angiotensin-2 also stimulates the secretion of aldosterone from the adrenal cortex.Angiotensin-2 stimulates the release of antidiuretic hormone (ADH), also called vasopressin, a vasoconstrictor too, mainly stimulating water reabsorption in the kidney and the sense of thirst.

ACE inhibitors, angiotensin receptor blockers (ARBs), and direct renin blockers are used in the treatment of hypertension and congestive heart failure, preventing stroke, preventing nephropathy including diabetic nephropathy, or preventing recurrent atrial fibrillation (Boggish et al. 2017).

Members of the class of ACE inhibitors are captopril (Capoten), enalapril (Renitec), quinapril (Accupril), and ramipril (Altace).

Members of the class of ARBs are valsartan (Diovan), losartan (Cozaar), and candesartan (Atacand).

Aliskiren (Tekturna) is an example of a direct renin inhibitor.

The Gonadal Sex Steroids

Testosterone and its more potent metabolites dihydrotestosterone, progesterone, and estradiol are classified as sex steroids. Cholesterol is the precursor for these hormones like it is for cortisol and aldosterone. The nonsteroidal hormones, luteinizing hormone (LH), follicle-stimulating hormone (FSH), and gonadotropin-releasing hormones, are usually not regarded as sex hormones, although they play major sex-related roles.

Natural sex steroids are made by the gonads, ovaries, and testes, by adrenal glands, or by conversion from other sex steroid in other tissues as the liver or fat. Sex steroids include the androgens consisting of androstenedione and dehydroepiandrosterone (DHEA). Estrogens include estradiol, estriol, and estrone. Progesterone belongs to the progestogens. Synthetic sex steroids as synthetic androgens are often referred to as anabolic steroids.

Testosterone induces male secondary sex characteristics. It has an effect on spermatogenesis and has an androgenic anabolic effect in the maintenance of muscle mass.

Androgen agonist steroids (AASs) are synthetic modifications of testosterone that are more or less anabolic or more or less androgenic, having different affinity for the testosterone receptor. The testosterone receptor is an intracellular cytosol receptor. These receptors are located in tissues as the scalp, prostate, and skin.

Androgenic effects are likely mediated under the influence of dihydrotestosterone (DHT), which is produced by the conversion of testosterone by the enzyme 5-alpha reductase. DHT has an affinity for the testosterone receptor three to four times higher than testosterone.

Other mechanisms of direct and indirect anabolic effects include anti-glucocorticoid effects by displacement of glucocorticoids from their receptor and increases in circulating insulin-like growth factor (IGF-1) as well as upregulation of IGF-1 receptors.

Clinically, AASs have been used to treat a host of conditions including many forms of anemia, acute and chronic wounds, protein malnutrition, severe burns, short stature, osteoporosis, primary or secondary hypogonadism, catabolic states due to long-term use of corticosteroids, and HIV wasting syndrome. Testosterone and AASs have been used and abused by individuals to augment their anabolic and androgenic potential. By doing so, these persons aim to boost their physical performance in athletic endeavors or improve their physique (Taylor et al. 2017; Kshirsagar and Wankhede 2017).

Common testosterone esters include testosterone propionate, testosterone enanthate, testosterone cypionate, and methyltestosterone. Methyltestosterone can be aromatized to the potent estrogen 17-alpha-methylestradiol.

Adverse effects are numerous. Cardiovascular side effects include increased heart rate, increased blood pressure, lowering HDL cholesterol and increasing LDL cholesterol, lower left ventricular mass, and ventricular arrhythmias. Nearly all oral AAAs are hepatotoxic in a dose-dependent manner. AAAs use also results in suppression of clotting factors 2, 5, 7, and 10 as well as prolonging prothrombin time.

Gynecomastia can occur as a result of the aromatization of testosterone and AASs. Acne and male baldness are greatly exacerbated by most AASs in susceptible individuals. Worsening benign prostate hyperplasia (BPH) occurs frequently.

AASs as nandrolone decanoate are used in hemodialysis in combination with recombinant erythropoietin (EPO) to treat the anemia of chronic renal disease.

Antiandrogens

Antiandrogens are also known as androgen antagonists or testosterone blockers. They act by blocking the androgen receptor and/or inhibiting or suppressing androgen production.

In men antiandrogens are used in the treatment of prostate cancer, BHP, androgenic alopecia, hypersexuality, paraphilias, and precocious puberty.

In women antiandrogens are used to treat acne, seborrhea, hidradenitis suppurativa, hirsutism, and hyperandrogenism as seen in polycystic ovary syndrome (PCOS). Antiandrogens are also used in the hormone replacement therapy for transgender women.

In males the major side effects are demasculinization and feminization. The side effects of antiandrogens in woman are minimal.

A number of antiandrogens have been associated with hepatotoxicity. Cyproterone acetate, flutamide, and aminoglutethimide are known for this.

Androgen synthesis inhibitors include aminoglutethimide which inhibits the cholesterol side-cleaving enzyme CYP11A1 which is responsible for the conversion of cholesterol into pregnenolone. Ketoconazol and abiraterone are inhibitors of the enzyme CYP17A1 which converts pregnenolone into androgens, mineralocorticoids, and glucocorticoids. To prevent adrenal insufficiency, Addison's disease, all these inhibitors require concomitant treatment with a glucocorticoid.

5-alpha-reductase inhibitors such as finasteride and dutasteride are inhibitors of 5-alpha reductase, the enzyme that converts testosterone to dihydrotestosterone (Woljee et al. 2017).

Antigonadotrofins are drugs which suppress the GnRH-mediated secretion of LH and FSH (Crew et al. 2017).

GnRH agonists and GnRH antagonists are powerful antigonadotrofins that are able to suppress androgen levels by 95% in men. As mentioned before leuprorelin (Lupron) and goserelin (Zoladex) are examples of GnRH agonists. An example of a GnRH antagonist is cetolix (cetrorelix).

Estrogens and Progestogens

Estrogens are the primary female sex hormones and are important in the estrous cycle in females. Natural estrogens are steroid hormones, while some synthetics are nonsteroidal. Enzymatic actions produce estradiol from testosterone while estrone is made from androstenedione.

Phytoestrogens have analogous effects of human estrogens in serving to reduce menopausal symptoms as well the risk of osteoporosis and heart disease.

Progesterone is a steroid hormone involved in the luteal phase of the menstrual cycle. Progesterone is produced in the ovaries and adrenals and in pregnancy and is synthesized from pregnenolone. Progesterone is also called the "pregnancy hormone." It converts the endometrium in the luteal phase to prepare the uterus for implantation. If pregnancy does not occur, progesterone levels drop leading to menstruation.

During implantation and gestation, progesterone appears to decrease the immune response to allow pregnancy. Progesterone decreases contractility of the uterine smooth muscle. Progesterone inhibits lactation during pregnancy. A drop in progesterone levels is possibly one step that facilitates the onset of labor.

The combined oral contraceptive pill contains estrogens and progestogens. Morning-after pills and abortion pills contain only progestogens as Mifepristone and Misoprostol. Mifepristone is an antiprogestin. Antiprogestins bind strongly to both progesterone and glucocorticoid receptors.

Although antiprogestins delay ovulation, this effect is inconsistent unless high doses are given. Under these circumstances, the antiprogestin

effect is associated with unopposed estrogen action.

Antiestrogens are mainly used as a means of estrogen deprivation therapy in the treatment of ER-positive breast cancer. They are also used to treat infertility, male hypogonadism, and gynecomastia in men and in the hormone replacement for transgender men.

Antiestrogens include selective receptor modulators (SERMs) like tamoxifen, clomifene, and raloxifene and selective estrogen receptor degraders (SERDs) such as fulvestrant and aromatase inhibitors (Ais) like anastrozole and progestogens and GnRH analogues.

Conclusion

This endocrinology pharmacodynamic evaluation clearly shows the enormous impact of endocrinology in nearly all areas of medicine. If you could take a look in the medicine cabinet of an older Western patient, you would recognize all classes of medicines that came along in this review.

References and Further Reading

Alagiakrishnan K (2016) Melatonin based therapies for delirium and dementia. Discov Med 21:363

Bangalore S (2017) Renin angiotensin system inhibitors for patients with stable coronary artery disease without heart failure: systematic review and meta-analysis of randomized trials. BMJ 356:j4

Black DM, Rosen CJ (2016) Postmenopausal Osteoporosis. N Engl J Med 374:254–262

Boggish A,L, Weiner RB, Kanayama G et al (2017) Cardiovascular toxicity of illicit anabolic-androgenic steroid use. Circulation 135:1991–2002

Campbell EJ, Campbell GM, Hanley DA (2015) The effect of parathyroid hormone and teriparatide on fracture healing. Expert Opin Biol Ther 15(1):119–129

Christ-Crain M, Fenske W (2016) Copeptin in the diagnosis of vasopressin-dependent disorders of fluid homeostasis. Nat Rev Endocrinol 12:168–176

Crew KD, Albain KS, Horshman Dl et al (2017) How do we do increase uptake of tamoxifen and other anti-estrogens for breast cancer prevention. Breast Cancer 3:20

De Leo S, Lee SY, Braverman LE (2016) Hyperthyroidism. Lancet 388(10047):908–918

Der-Nigoghossian C, Lesch C, Berger K (2017) Effectiveness and tolerability of conivaptan and tolvaptan for the treatment of hyponatremia in neurocritically ill patients. Pharmacotherapy 37(5):52–534. https://doi.org/10.1002/phar.1926

Freda PU, Gordon MB, Kelepouris N et al (2015) Long-term treatment with pegvisomant as monotherapy: experience from Acrostudy. Endocr Pract 2(3):264–274

Hippisley-Cox J, Compland C (2016) Diabetes treatments and risk of heart failure, cardiovascular disease and all cause mortality: cohort study in primary care. BMJ 354:i3477

Karagiannis T et al (2012) Dipeptidyl peptidase-4 inhibitors for treatment of type 2 diabetes mellitus in the clinical setting. Systematic review and meta-analysis. BMJ 344:e1369

Kshirsagar SC, Wankhede UN (2017) Combination of mifepristone and misoprostol: an effective method of medical abortion up to gestational age of 49 days. Int J Reprod Contracept Obstretr Gynaecol 6(7):2923–2925

Kuiper JG, Bezemer ID, Driessen MT (2016) Rates of prostate surgery and acute urinary retention for benign prostate hyperplasia in men with dutasteride or finasteride. BMC Urol 16:53

Laar van de FA (2008) Alpha-glucosidase inhibitors in the early treatment of type 2 diabetes. Vasc Health Risk Manag 4(6):1189–1195

Lispkak J, Hirsch B, Riddle MC (2017) Human insulin for type 2 diabetes. An effective, less expensive option. JAMA 318(1):23–24

Loriaux DL (2017) Diagnosis and differential diagnosis of Cushing's syndrome. N Engl J Med 376:1451–1459

Lovshin JA (2017) Glucagon-like peptide-1 receptor agonists: a class update for treating type 2 diabetes. Can J Diabetes. pii: S1499-2671(16);30532-9

Maralikesevan et al (2016) Myelotoxicity of peptide receptor radionuclide therapy of neuroendocrine tumors: a decade of experience. Cancer Biother Radiopharm. https://doi.org/10.1089/obr.2016.2035

Marathur NM, Tseng E, Hutfless et al (2016) Diabetes medications as monotherapy or metformin-based combination therapy for type 2 diabetes: a systematic review and meta-analysis. Ann Int Med 164:740

Mondoza G, Piroletta A, Jiminez-Ceja LM (2013) The role of nateglinide and repaglinide, derivatives of meglitinide in the treatment of type 2 diabetes mellitus. Arch Med Sci 9:936

Niamh M (2013) Prolactin disorders. Medicine 41(9): 516–518

Posner LS, Camarga CA (2017) Update on the usage of epinephrine autoinjectors. Drug Healthc Patient Saf 9:9–18

Riddle MC (2017) Modern sulfonylureas: dangerous or wrongly accused? Diabetes Care 40(5):629–631

Rossi GP, Ceolotto G, Caroecia B (2017) Genetic screening in arterial hypertension. Nat Rev Endocrinol 13:289–298

Salahudeen MS, Nishtala PS (2017) An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. Sauidi Pharm J 25(2):165–175

Shipley WU, Sefferheld W, Hinnushi R et al (2017) Radiation with or without antiandrogen therapy in recurrent prostate cancer. N Engl J Med 376:417–428

Singh N, Naha Malhotra N et al (2014) Comparison of gonadotropin releasing hormone agonist with GnRH antagonist in polycystic ovary syndrome undergoing in vitro fertilisation cycle. Retrospective analysis from a tertiary center and review of literature. J Hum Reprod Sci 7(1):52–57

Steinman MA, Zullo AR, Lee et al (2017) Association of beta-blockers with functional outcomes, death and rehospitalization in elder nursing home residents after myocardial infarction. JAMA Intern Med (2):254–262

Streeten EA, Mohtasebi Y, Konig M et al (2017) Hypoparathyroidism: less severe hypocalcemia with treatment with vitamin D2 compared with calcitriol. J Clin Endocrinol Metab 102(5):1505–1515

Taylor HS et al (2017) Treatment of endometriosis-associated pain with Elagolix, an oral GnRH antagonist. N Engl J Med 377:28

Thonos L, Gregg B (2017) Metformin, a review of its history and future: from lilac to longevity. Pediatr Diabetes 18(10):10–16

Woljee AK, Rogers MA, Lin P et al (2017) Short term use of oral corticosteroids and related harms among adults in the United States: population based cohort study. BMJ 357:j1415

Zanos P, Georgia P, Weber C et al (2017) Oxytocin and opioid addiction revisited: old drug, new applications. Br J Pharmacol. https://doi.org/10.1111/bph13757

Zou H, Zhou B, Xu G (2017) SGLT 2 inhibitors: a novel choice for the combination therapy in kidney disease. Cardiovasc Diabetol 16:65

Pharmacodynamic Evaluation: Dermatology

13

Liora Bik and Hok Bing Thio

Contents

Abstract

Psoriasis, atopic dermatitis, and skin cancer are the three main dermatological, highly prevalent diseases for which new effective systemic therapies have been developed. For the immune-mediated inflammatory skin diseases, several biologic agents are now available for the chronic plaque psoriasis. These agents target one specific proinflammatory cytokine (TNF-alpha, interleukin (IL)-17, and (IL)-23) or its receptor. The IL-4 receptor blocking dupilumab is the first biologic agent which can be used for the chronic inflammatory skin disease, atopic dermatitis. By interfering in the Hedgehog signaling pathway, vismodegib proves to be effective in metastasized or extensive basal cell carcinoma (BCC) where a surgical procedure is no longer feasible. Immunotherapy with checkpoint inhibitors is a major

L. Bik (✉) · H. B. Thio (✉)
Dermatology, Erasmus University Medical Center, Rotterdam, The Netherlands
e-mail: h.thio@erasmusmc.nl

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_36

breakthrough in melanoma and has been rewarded with the Nobel Prize in Physiology and medicine, 2018. All these new systemic therapies are valuable additions to the already existing armentarium of the dermatologist.

Psoriasis

Pathophysiology

Psoriasis is a chronic immune-mediated inflammatory skin disease. Predisposition is polygenetic and the disease is triggered by trauma, infection, or medication such as beta-blocking agents. Both innate as adaptive immunity are involved in the pathogenesis of psoriasis (Nestle et al. 2009). Dendritic cells, T cells, neutrophils, keratinocytes, and cytokines released by immune cells (especially TNF-alpha, interleukin (IL)-17, and (IL)-23) contribute to the initiation of the disease (Fig. 1).

Clinical Features

Psoriasis has various distinct clinical variants. The major categories include chronic plaque psoriasis, which is the most common, guttate psoriasis, inverse psoriasis, scalp psoriasis, nail psoriasis, pustular psoriasis, and erythrodermic psoriasis (Boehncke and Schön 2015). The typical psoriatic skin lesions are symmetrical, well-demarcated erythematous plaques with silvery scales. Trauma of the skin may provoke new psoriasis lesions. This is called Koebner phenomenon. An associated disorder is psoriatic arthritis in which psoriasis is related to with spondyloarthropathy and/or peripheral arthritis (Fig. 2).

Diagnostics

In clinical practice, diagnosis can be made by proper anamnesis and physical examination.

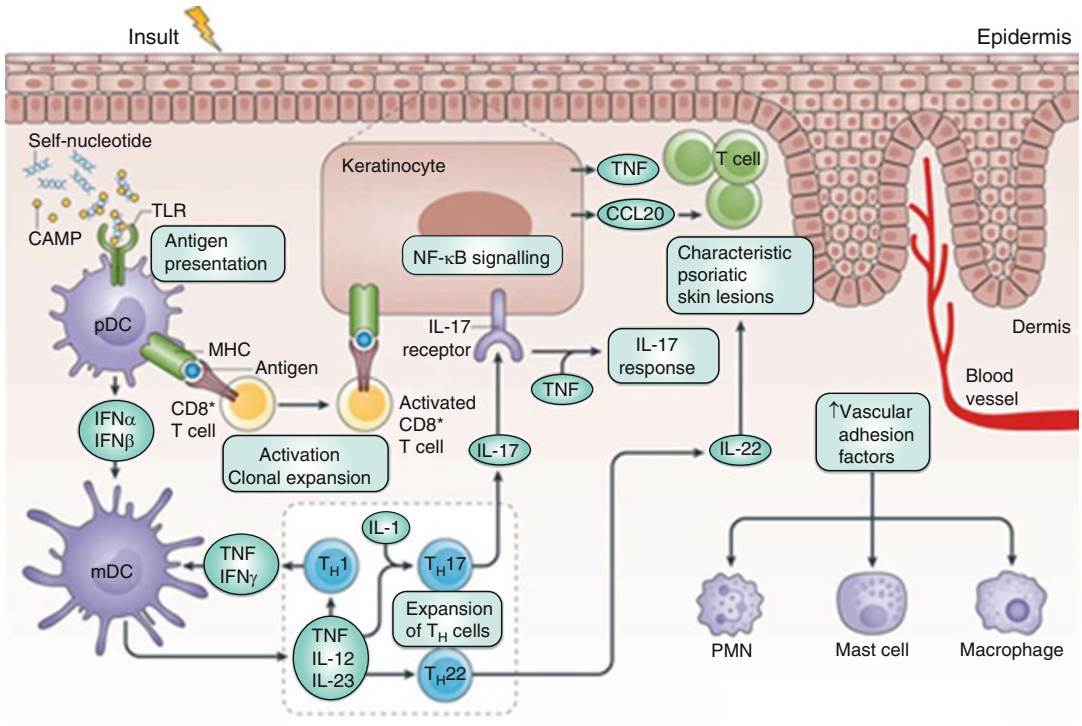


Fig. 1 Pathogenesis of psoriasis

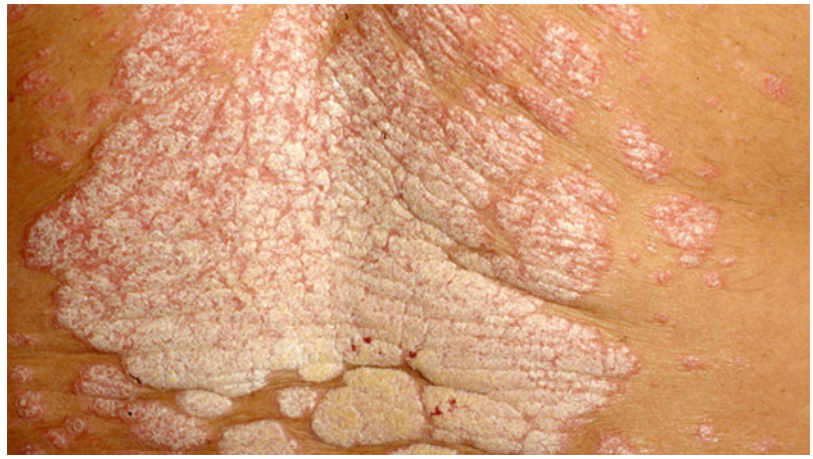


Fig. 2 Chronic plaque psoriasis

However, in some cases, it is necessary to do a skin biopsy to confirm the diagnosis. Histopathologic findings in psoriatic skin are confluent parakeratosis with some neutrophils, regular acanthosis, epidermal hypogranulosis, dermal inflammatory infiltrate, and tortuous dilated dermal capillaries.

Treatment

A broad spectrum of topical and systemic treatment options are available for psoriasis and are discussed below.

Topical Corticosteroids and Vitamin D Analogs

The first choice treatment are topical corticosteroids frequently in combination with a vitamin D analog, such as calcipotriol (Mason et al. 2013). Topical corticosteroids exert anti-inflammatory, antiproliferative, and immunosuppressive effects. Based on vasoconstrictive assays, they are categorized in low, medium, high, or very high potent groups. Vitamin D analog binds to its receptor, which inhibits keratinocyte proliferation and enhances keratinocyte differentiation. Vitamin D has also specific immunomodulatory effects.

Ultraviolet Phototherapy

Moderate to severe psoriasis can be treated with ultraviolet B (UVB), narrowband UVB (nUVB), and photo chemotherapy with ultraviolet A (UVA) following topical or systemic pretreatment with psoralen. All these treatments are antiproliferative and anti-inflammatory by apoptosis of activated inflammatory cells (Weichenthal et al. 2005).

Systemic Medication

Systemic medication is used for moderate to severe psoriasis with large affected body skin surface, but also for nail psoriasis. Regular blood tests should be performed with all systemic medications.

- **Methotrexate** – Methotrexate is a folic antagonist closely related with the substance aminopterin (Menter et al. 2009). It inhibits keratinocyte proliferation and has immunosuppressive effect of activated T cells in psoriatic plaques. Maximum improvement can be seen after 8–10 weeks.
- **Ciclosporin** – Ciclosporin is a calcineurin inhibitor, which prevents T cells from releasing proinflammatory effector cytokines like IL-2. The maximum treatment duration is 1 year due to potential nephrotoxic effects.
- **Acitretin** – Synthetic derivatives of vitamin A normalize keratinocyte proliferation and differentiation, but it has also distinct inhibitory effects on Th-17 cells.
- **Fumarates** – Dimethylfumarate and the biologically active subunit monomethylfumarate act anti-inflammatory and immunosuppressive

by inhibiting proliferation of dendritic and T cells and the release of cytokines like TNF alpha and IL-8. Clinical effects can be seen after 6–8 weeks.

- **Apremilast** – A small molecule inhibitor of the intracellular enzyme phosphodiesterase 4 has anti-inflammatory effects. This oral medication has also a profound antipruritic effect in psoriasis.
- **Biologic agents** – Biologic agents are the latest addition in the treatment of psoriasis and are more expensive compared to other systemic medication (Menter et al. 2008). The mode of action differs per biologic agent, but all biologic agents induce an anti-inflammatory response. Infliximab, adalimumab, etanercept, and certolizumab are TNF-alpha inhibitors. Certolizumab is of special interest due to its pegylated form. This biologic agent can be safely prescribed in pregnant and lactating female psoriasis patients. Secukinumab and ixekizumab are anti-IL-17A monoclonal antibodies and brodalumab blocks the IL-17 receptor. Ustekinumab is a human monoclonal antibody that targets IL-12 and IL-23, while guselkumab is a specific IL-23 antagonist (Fig. 3).

Therapeutic Management

Treatment evaluation can be done by physical examination. To evaluate objectively, physicians use PASI: Psoriasis Area and Severity Index (Schmitt et al. 2005). Points are given for severity of erythema, induration, and scaling respectively on the head, upper limbs, trunk, and lower limbs. A high score means severe psoriasis.

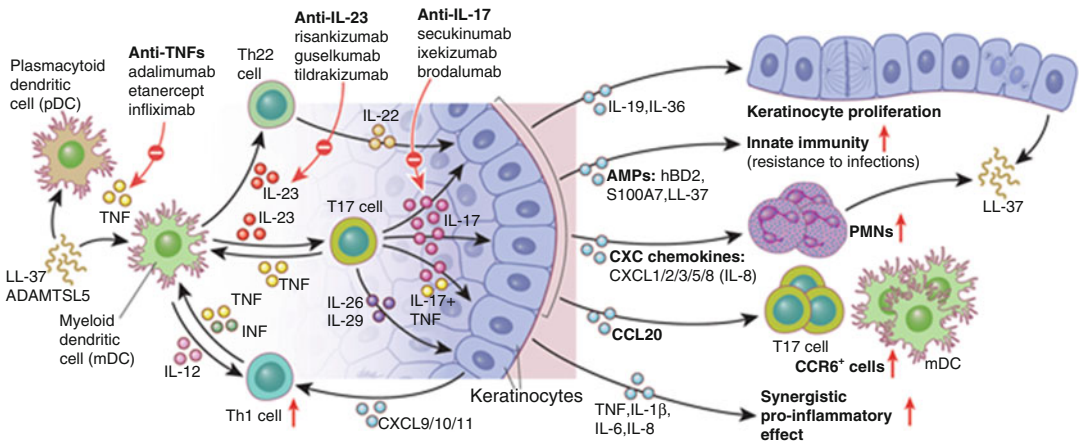
Feature Treatments

Research suggests new treatment options for psoriasis, such as other small molecules that target the intracellular signaling. This includes JAK (Janus kinase) inhibitors: tofacitinib and baricitinib (Ports et al. 2013).

Atopic Dermatitis

Pathophysiology

Atopic dermatitis is a chronic pruritic inflammatory skin disease that occurs mostly in young children that can persist in adulthood. Atopic



Jason E. Hawkes, Tom C. Chan, James C. Krueger
Journal of Allergy and Clinical Immunology, Volume 140, Issue 3, September 2017, Pages 645-653
Psoriasis pathogenesis and the development of novel targeted immune therapies

Fig. 3 Immunopathogenesis of psoriasis and the mode of action of the biologic agents

syndrome stands for a group of IgE-related diseases: atopic dermatitis, allergic rhinoconjunctivitis, and asthma. The pathogenesis of atopic dermatitis is multifactorial and involves genetic predisposition, immune dysregulation, and environmental irritation (Weidinger et al. 2018). This leads to defects of the skin barrier and major skin inflammation.

Genetics

The most important part of the barrier function of the skin is located in het stratum corneum, the outer layer of the epidermis. It prohibits environmental irritants from entering the skin and prevents excessive water loss. Filaggrin is a protein produced by differentiating keratinocytes in the epidermis and is encoded by the FLG gene on the 1q21 epidermal differentiation complex. Filaggrin is needed for correct formation of corneocytes and therefore responsible for an intact barrier function of the skin. A loss-of-function mutation of the filaggrin (FLG) gene leads to skin barrier abnormalities and is a major risk factor for atopic dermatitis (Thyssen and Kezic 2014). Another risk factor is a positive family history of parents with an atopic disease. One atopic parent gives a two- to threefold increased risk of atopic dermatitis. Two atopic parents increase the risk even three- to fivefold.

Immune Dysregulation

The innate immune system in atopic dermatitis is altered and leads to a different skin microbiome and more severe inflammations. Patients with atopic dermatitis have more often colonization and infections with *Staphylococcus aureus*.

The adaptive immune system in atopic dermatitis is also altered by increased expressions of Th2, Th22, and Th17 cytokines leading to immune dysregulation (Boguniewicz and Leung 2011) (Fig. 4).

Environmental Irritation

Irritation factors for the skin vary from excessive contact with water, temperature changes, low humidity, bacterial or viral infection, scratching, and stress (Weidinger et al. 2018). Topical treatment and emollients are needed to prevent these factors from penetrating the skin and causing flares.

Clinical Features

Main characteristics of atopic dermatitis are pruritus and dry skin. Physical examination show morphological differences between acute and chronic atopic dermatitis lesions. In the acute phase, the lesions are erythematous squamous papules and plaques with exudate and vesicles. When this persists, it passes into the chronic form with xerosis and erythematous squamous plaques with lichenification (Rudikoff and Lebwohl 1998).

The form and localization of skin lesions differ per age.

Young children with atopic dermatitis, before the age of two, show the acute form of skin lesions on the cheeks and extensor sites of the body. Napkin area is usually spared. Older children and adolescents, until the age of 18, show more frequently the chronic form of skin lesions on flexor sites of the body, such as antecubital and popliteal fossae, and the neck. Lichenification is more present due to intensive scratching. In adulthood, the chronic form is also seen on flexure sites and to a lesser extent on the hands, neck, and face.

Besides the main characteristics described above, there are several clinical findings that are seen as minor criteria for atopic dermatitis. These are the following: keratotis pilaris, pityriasis alba, periorbital darkening, Dennie-Morgan infraorbital folds, and Hertoghe's sign (hair loss of lateral eyebrows).

Atopic dermatitis can be complicated with a secondary impetiginization, usually with *Staphylococcus aureus*. If so, yellow crusts and exudate are seen at the affected skin lesions (Rudikoff and Lebwohl 1998). Another infectious complication may occur with the herpes simplex virus. This is known as eczema herpeticum and

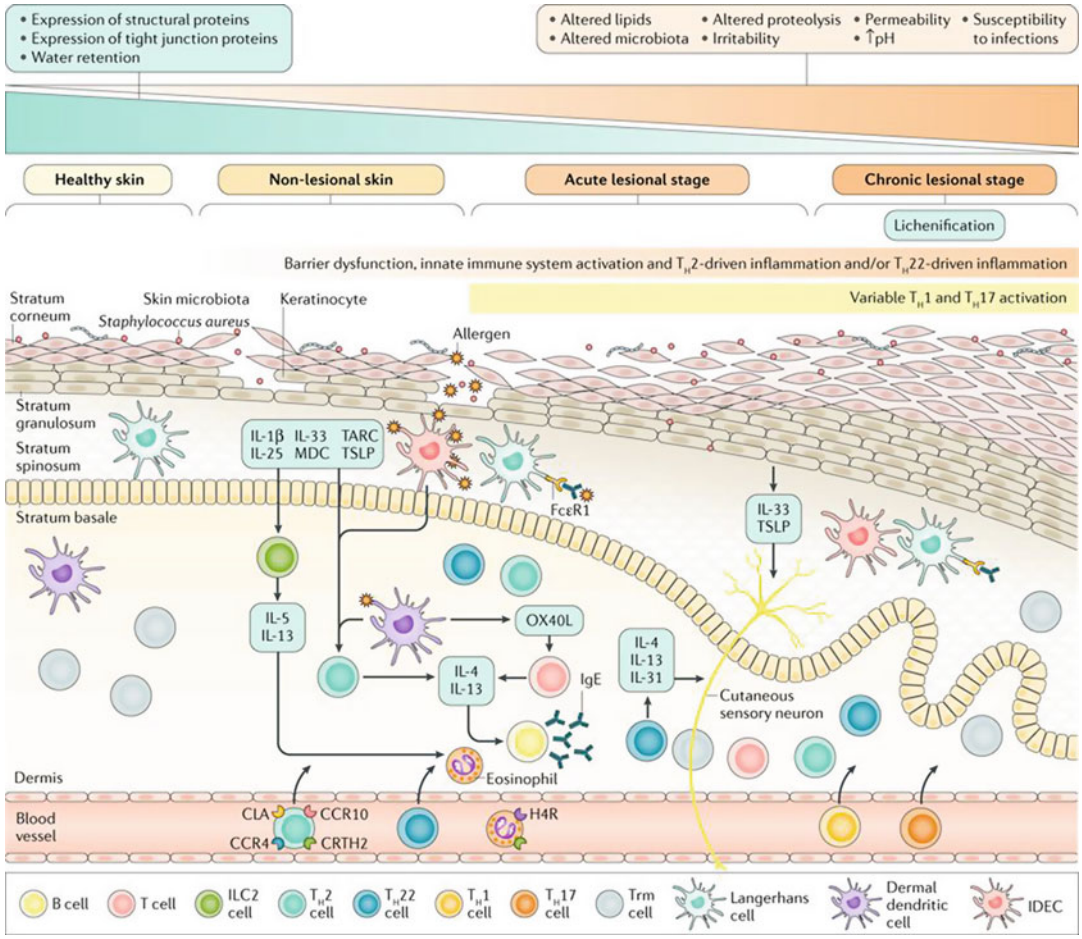


Fig. 4 Immune dysregulation in atopic dermatitis (Weidinger et al. 2018)

needs intravenous treatment with antiviral agents, because of high morbidity (Leung 2013).

Diagnostics

Atopic dermatitis is a clinical diagnosis based on major and minor disease symptoms, family history, and physical examination. A skin biopsy may help for histological confirmation of the diagnosis. Histological findings are hyperkeratosis, epidermal edema (spongiosis) with acanthosis, and lymphohistiocytic dermal infiltration (Weidinger et al. 2018).

An elevated total IgE in blood may support the diagnosis (Fig. 5).

Treatment

Various treatments are available for atopic dermatitis. Often are education, use of emollients, and topical corticosteroid therapy sufficient. In severe disease, ultraviolet phototherapy or systemic medication can be given.

Education and Use of Emollients

First step in treatment of atopic dermatitis is education about pathophysiology of atopic dermatitis, good skin care, and avoidance of skin irritants (Tollefson and Bruckner 2014). It is better to avoid excessive water contact because it causes dehydration of the skin. Triggers of skin irritation need to be avoided. Frequent application of

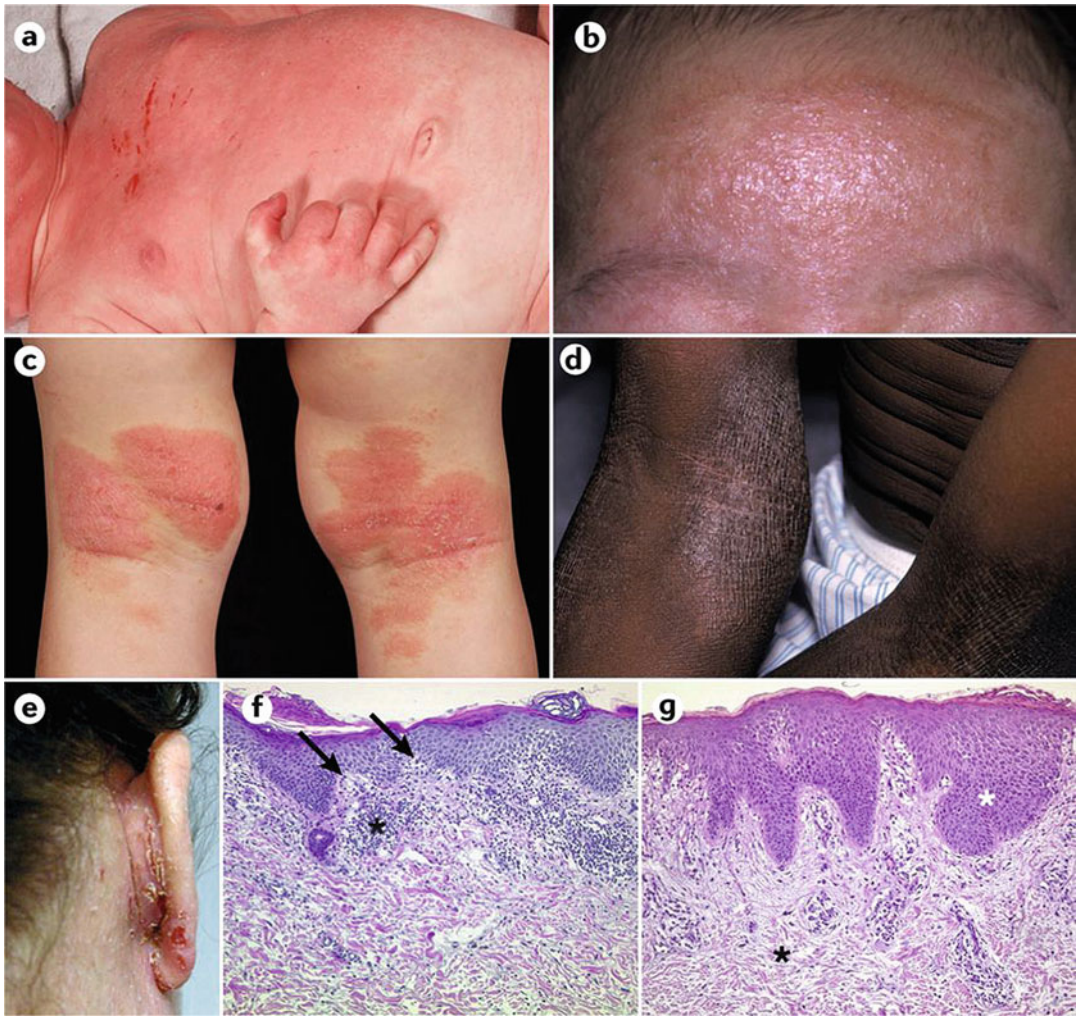


Fig. 5 Clinical and histological features of atopic dermatitis (Weidinger et al. 2018)

skin emollients, two to three times a day, helps recover the stratum corneum lipids. Anti-itch treatment with (sedating) antihistamines can be prescribed to stop scratching and can restore insomnia.

Topical Corticosteroids and Calcineurin Inhibitors

The mainstay in treatment of atopic dermatitis is topical corticosteroids (Eichenfield et al. 2014). Mild atopic dermatitis reacts well to low potent corticosteroids, such as hydrocortisone 1%. Patients with moderate disease need medium to high potent corticosteroids, such as betamethasone dipropionate 0.05%. Mechanism of action is antipruritic, vasoconstrictive, and anti-inflammatory. It suppresses the release and activity of various proinflammatory mediators, such as kinines, histamine, and prostaglandins.

Topical calcineurin inhibitors, tacrolimus and pimecrolimus, are good alternatives for topical corticosteroids. Mechanism of action is binding to FKBP-12, an intracellular protein, and other complexes with calcineurin dependent proteins (Eichenfield et al. 2014). This leads to inhibition of calcineurin phosphatase activity and thereby reduction of T-lymphocytes activation. It can be used in the face, neck, and skinfolds. Most

common side effects are temporal sensations of burning, redness, and itch.

Ultraviolet Phototherapy

Ultraviolet phototherapy is indicated in moderate atopic dermatitis when topical therapy fails. Narrowband ultraviolet B (UVB) with a spectrum of 311–313 nm is preferred. This treatment is only suitable for adults and not for children. The mechanism of action is anti-inflammatory (Weichenthal et al. 2005).

Systemic Medication

Systemic medication is indicated in moderate to severe atopic dermatitis. First choice of systemic medication is cyclosporine due to fast relieve of symptoms. Systemic medication used in atopic dermatitis is described below (Ring et al. 2012).

- **Ciclosporin** – Ciclosporin is a lipophilic cyclic peptide of 11 amino acids that binds to cyclophilins. This drug-receptor complex binds and inhibits calcineurin, causing a reduced release of cytokines IL-2, TNF-alpha, and IL-3. It suppresses T-cell proliferation. The maximum treatment duration is 1 year due to nephrotoxic effects.
- **Systemic corticosteroids** – Short period of prednisone is given in moderate to severe atopic dermatitis when ciclosporin is contraindicated. Prednisone is converted in the liver to an active metabolite, prednisolone. It acts as a glucocorticosteroid, which has immunosuppressive action due to reduction in activity of lymphocytes and suppression of the adrenal function.
- **Methotrexate** – Methotrexate is used as a long-term treatment in patients with severe atopic dermatitis. Remission induction is seen several weeks after start. Folic acid is daily given to prevent methotrexate toxicities. Methotrexate is a folate antimetabolite that interferes with DNA synthesis and repair. Rapidly proliferating tissue is most receptive for methotrexate.
- **Dupilumab** – Dupilumab is most recently accepted as a treatment for moderate atopic dermatitis (Beck et al. 2014). It is a human monoclonal IgG4 antibody and works immunosuppressive by blocking IL-4Ralpha subunits, which leads to inhibition of IL-4 and IL-13. Administration is by subcutaneous injections once every 2 weeks.

Antibacterial and Antiviral Treatment

Atopic dermatitis patients with mild bacterial superinfection, usually with *Staphylococcus aureus*, can be treated with topical mupirocin 2% cream twice daily for 2 weeks. Severe infections are treated with oral antibiotics (Ring et al. 2012). Viral superinfections with herpes simplex need to be treated with an oral antiviral therapy; valaciclovir. Intravenous antiviral treatment during hospital admission is indicated for severe life-threatening cases of eczema herpeticum.

Therapeutic Management

Efficacy of treatment is clinically observed by the treating physician. Objective tools can be used for treatment evaluation, such as EASI, SCORing, and POEM.

The Eczema Area and Severity Index (EASI) is used to measure extend and severity of atopic dermatitis (Hanifin et al. 2001). The score can be calculated by recording the severity of redness, thickness of the skin, scratching, and lichenification at four different body areas. This results in a score 0, for no active lesions, till 72, for very severe atopic dermatitis.

Another tool to measure the extent and severity of atopic dermatitis is the SCORAD (SCORing Atopic Dermatitis) (SCORAD 1993). First, the percentage of affected skin is measured. Next, one region is assessed as none 0, mild 1, moderate 2, or severe 3 for redness, swelling, oozing/crusting, scratch marks, lichenification, and dryness. Subjective symptoms, itching and sleeplessness, are scored 0 for not present to 10 for worst as possible.

Patient-oriented eczema measure (POEM) is a questionnaire filled in by patients or caregivers that result in an anamnestic severity index for atopic dermatitis (Charman et al. 2004).

Future Treatments

Studies of new treatments for atopic dermatitis show promising results for JAK inhibitors and monoclonal anti-IL 31, 13, and 22 antibodies. Tofacitinib is a small-molecule Janus kinase (JAK) inhibitor which inhibits cytokine signaling of interleukin (IL)-4, IL-5, and IL-13 in a 2% ointment topical treatment (Bissonnette et al. 2016). Nemolizumab, lebrikizumab, and fezakinumab are monoclonal antibodies that bind to, respectively, IL-31 receptor A, IL-13, and IL-22 (Ruzicka et al. 2017; Simpson et al. 2018; Guttman-Yassky 2018). Working mechanism is by interfering with the cytokine-mediated immune response and inflammation in affected skin.

Neoplasms of the Skin

Nonmelanoma Skin Cancer: Basal Cell Carcinoma and Squamous Cell Carcinoma

Pathophysiology

Basal Cell Carcinoma

Basal cell carcinoma (BCC) is the most common type of skin cancer and consists of 80% of all skin cancers. It originates from the basaloid cells in the epidermis and grows very slowly. Without treatment it grows invasive and causes tissue destruction. BCC is usually diagnosed in early stages and metastasize rarely.

The most important risk factor for BCC is intense intermittent ultraviolet radiation (UV) in sunlight/sunburn (Gallagher et al. 1995). Extensively, sun light exposure, before the age of 18 years old, has a direct relation with risk of BCC. The use of tanning beds also elevates the risk for BCC (Ferrucci et al. 2012). Related risk factors are fair skin with freckling's and light colored eyes and older age. Prevention of sunburn with sun protection cream under the age of 18 decreases the risk of nonmelanoma skin cancer dramatically.

Other risk factors are a past of radiation therapy, long-term use of immunosuppressants in organ transplant patients, and genetics. Patients with family members with BCC have an increased risk of developing BCC during life.

Basal cell nevus syndrome or Gorlin syndrome is a rare autosomal dominant genetic disorder related with BCC (Gorlin and Goltz 1960). It is caused by a germline mutation of a human patches gene (PTCH). Multiple BCC are presented before the age of 35, mostly at the trunk. Other clinical features are dysmorphic facial characteristics, bifid ribs, palmar and plantar pitting, mandible bone cysts, calcification of flax cerebri, and medulloblastoma.

Squamous Cell Carcinoma

Squamous cell carcinoma (SCC) arises from keratinocytes in the epidermis. SCC represents 20% of all nonmelanoma skin cancers (Alam and Ratner 2001). Patient-specific characteristics and environmental factors contribute to the development of SCC. The incidence for SCC increases strongly with age. In addition, individuals with a light skin develop SCC much more often than people with dark type skin.

As for environmental factors, cumulative lifetime sun exposure is the most important cause of SCC. Other risk factors are radiotherapy, the use of chronic immunosuppressant for organ transplant patients, and positive family history of SCC (Alam and Ratner 2001). Chronic inflammation of the skin as in chronic ulcers or inflammatory dermatoses, such as lichen sclerosus et atrophicus, also increases the risk of SCC.

Clinical Features

Basal Cell Carcinoma

During physical examination, a distinction can be made between nodular and superficial BCC (Marzuka and Book 2015). Nodular BCC is pink or flesh-colored papules or plaques with a pearly reflection and telangiectasia. The border is usually elevated, which is described as rolled border. In more advanced stages, central ulceration is seen within the rolled border. Superficial BCC are most

commonly seen on the trunk. They present as red or pink macules with light scaling. When illuminated, they have a shiny appearance (Fig. 6).

Squamous Cell Carcinoma

Cutaneous SCC is most common at sun-exposed body sites. The head and neck represents about 55% of all SCC. Individuals with dark skin type usually develop SCC on body sites without sun exposure, such as perianal or chronic ulcers.

Most of the SCC (in situ) arises from actinic keratosis. Actinic keratosis is presented as scaly and red skin lesions at sun damaged skin, usually seen in older patients. Differentiation between actinic keratosis and SCC (in situ) can be difficult. Although SCC often arises from actinic keratosis, only 1% of all the lesions progresses into SCC (Criscione et al. 2009).

SCC in situ (Bowen's disease) is a well-demarcated red or skin colored plaque with scaling (Alam and Ratner 2001). SCC in situ is most often asymptomatic. SCC in situ of the penis is called erythroplasia of Queyrat. This is presented as a well demarcated, velvety, and red papule or plaque (Fig. 7).

Invasive SCC has different clinical appearances depending on differentiation level (Alam and Ratner 2001). Well-differentiated SCC present as solid hyperkeratotic papules or plaques with or without ulceration. Poorly differentiated SCC is a more soft and granulomatous papule without hyperkeratosis. Sensation of pain can occur. Neurological problems at the lesion site, such as numbness, paraesthesias, or paralysis, indicate perineural invasion by the tumor, which is a poor prognostic factor.

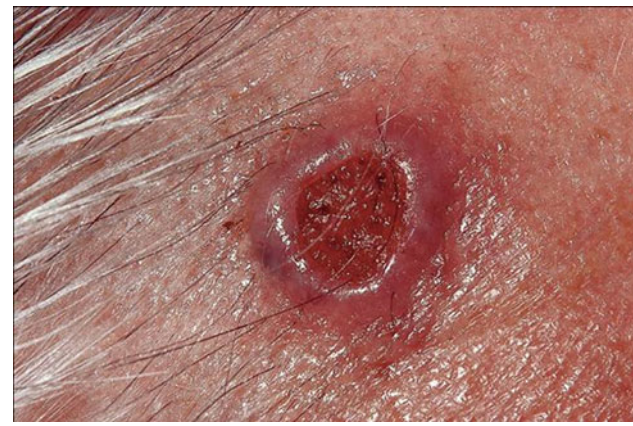


Fig. 6 Nodular basal cell carcinoma

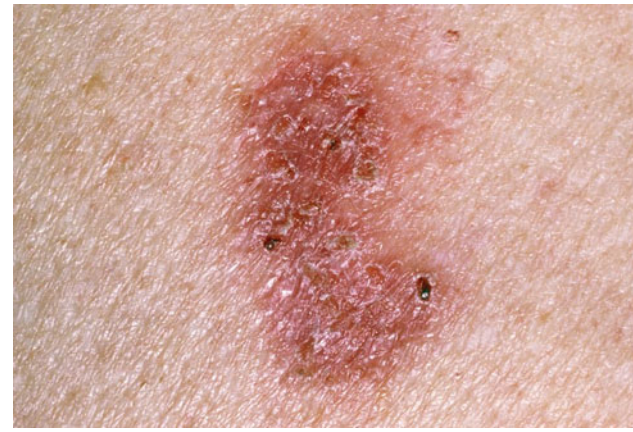


Fig. 7 Squamous cell carcinoma in situ (Bowen's disease)

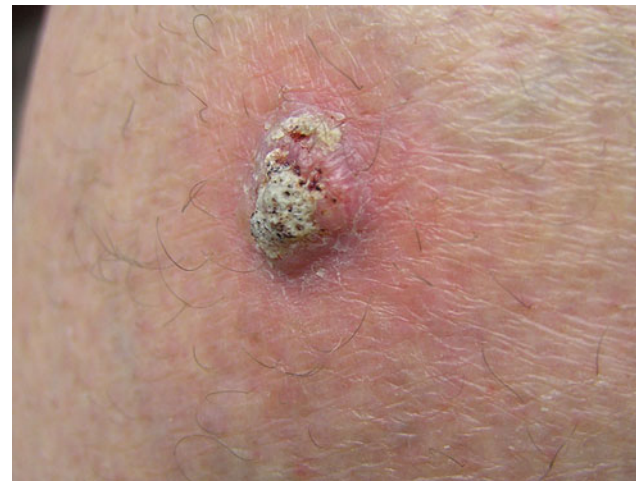


Fig. 8 Squamous cell carcinoma

When SCC metastasizes, this is most often seen in regional lymph nodes. In severe cases, metastasis can spread to lungs, liver, brain, skin, or bone. SCC on ears or mucocutaneous areas such as the lips tends to metastasize more often (Fig. 8).

Diagnostics

Diagnosis of BCC and SCC can be made based on physical examination. However, in most cases a skin biopsy is used for histological confirmation of the diagnosis. Histological findings of nodular BCC are large nests of atypical basaloid tumor cells with peripheral palisading in the dermis and slit-like stroma retractions (Crowson 2006). Superficial BCC are small clusters of atypical basaloid cells attached to the epidermal surface and show as well slit-like stroma retraction.

SCC in situ show keratinocytic dysplasia of the full thickness of the epidermis with no involvement of the dermis (Alam and Ratner 2001). The keratinocytes have intense mitotic activity, pleomorphism, and enlarged nuclei. Acanthosis, hyperkeratosis, and parakeratosis of the stratum corneum are often seen. The epidermis has a windblown look due to loss of maturity and polarity.

In contrast, actinic keratosis shows dysplastic keratinocytes only in parts of the epidermis.

In invasive SCC, the dysplastic keratinocytes are involved in the full epidermis and penetrate the basement membrane into the dermis or further surrounded tissue. Well-differentiated SCC shows large atypical keratinocytes with abundant cytoplasm. Poorly differentiated SCC shows anaplastic keratinocytes without differentiation and with many mitoses. Invasive SCC has a few exclusive histopathological variants, such as spindle cell SCC, clear cell SCC, and desmoplastic SCC (Yanofsky et al. 2011).

Besides diagnosis and tumor differentiation, the tumor depth and perineural invasion are also reviewed. This is necessary for proper tumor staging.

Treatment

Frequently used treatment of nonmelanoma skin cancer is surgical excision. Other treatment options are topical treatments, cryotherapy, photodynamic therapy, and radiation therapy, depending on the type and stage of skin cancer. The latest addition in therapeutic armentarium is systemic therapy with vismodegib for the following treatments:

- **Surgical excision** – Surgical excision is one of the most effective treatment with low recurrence rates for both BCC and SCC. This is performed under local anesthesia. Histologic assessment is used to determine completeness of the procedure. Margins for BCC vary 3–5 mm of normal skin. For SCC, it varies from 4 mm for low risk to 1 cm for high risk tumors (Brodland and Zitelli 1992).

Besides conventional surgical excision, Mohs micrographic surgery can be used to obtain an even lower recurrence rate. Indications are high-risk nonmelanoma skin cancers at complex anatomic sites such as the nose, lips, eyelids, eyebrows, and ears (Van Loo et al. 2014). Mohs micrographic surgery is performed by a specialized dermatology surgeon. After excision of the tissue, the lesion is cut into horizontal sections which make it possible to evaluate the full peripheral and deep margins.

- **Topical treatment** – Topical fluorouracil and imiquimod are topical treatment options which need to be applied 4–8 weeks at the tumor site (Love et al. 2009).

  Topical fluorouracil (FU), usually 5% cream, is a treatment option for actinic keratosis, SCC in situ, and superficial BCC when surgical excision is not preferred. FU is a pyrimidine antimetabolite that interferes with DNA synthesis by inhibiting thymidylate synthesis. Blockage of DNA synthesis leads to prevention of cell proliferation and cell death. Success of treatment depends highly on adequate application, duration, and tumor selection.

  Imiquimod 5% cream is used for actinic keratosis, anogenital warts, and superficial and nodular BCC at low-risk sites. It is a Toll-like receptor 7 agonist and immune cell modifier. It activates immune cells to produce cytokines that stimulate cell-mediated immunity and promotes apoptosis of malignant cells. Imiquimod is more effective in superficial BCC than topical fluorouracil. Oral immunosuppressive medication is a relative contraindication for the use of imiquimod.
- **Cryosurgery** – In actinic keratosis, SCC in situ, small low-risk SCC, and superficial BCC, cryosurgery is an effective modality for treatment (Kuflik 2004). The tumor is one or two times frozen and thawed through application of liquid nitrogen. Malignant cell destruction is due to formation of intra- and extracellular ice crystals, vascular stasis, hypertonic damage, and breakdown of the phospholipid membrane. It is a fast, cost effective treatment and usually needs no anesthesia. Hypopigmentation and scarring are the most common lasting side effects.
- **Photodynamic therapy** – Photodynamic therapy can be used in SCC in situ and superficial BCC (Morton et al. 2013). Mechanism of action is by topical applying a photosensitizing porphyrin, such as 5-aminolevulinic acid (ALA) or methyl aminolevulinate (MAL). Several hours after application, the tumor is exposed to a blue light (400–450 nm) or red light (630–635 nm). The photosensitizer absorbs the light and releases reactive oxygen. This leads to cell damage and cell death.
- **Radiation therapy** – Patients with BCC and SCC who are not candidates for surgical excision or who have multiple recurrences can be treated with radiation therapy (Locke et al. 2001). Disadvantages are need of multiple treatments, lack of histological control, more expensive, and local and long-term side effects.
- **Vismodegib** – Vismodegib is an orally available kinase inhibitor with specific activity against a key step in the Hedgehog signaling pathway. Hedgehog is a key regulator of embryonic development, cell growth, and differentiation. Clinical trials of vismodegib in patients with metastatic or locally advanced basal cell carcinoma reported at least partial responses in up to half of patients. Current indications include metastatic or locally advanced, recurrent, or unresectable basal cell carcinoma.

Regional Lymph Nodes

The extent of the disease in cutaneous SCC is evaluated trough physical examination and ultrasound of locoregional lymph nodes. Patients with enlarged lymph nodes need to undergo biopsy via fine-needle aspiration or surgical excision. When metastasis in lymph nodes occurs, further research is indicated for staging and prognosis.

Therapeutic Management

Follow-up for BCC consists of physical examination by a dermatologist once every 6 months

for 1 year and then annually. When no recurrence or new primary tumor occurs, follow-up will consist of self-examinations at home and visits as needed.

The follow-up interval for cutaneous SCC depends on the tumor stage and varies between 3 and 6 months for 2 years and afterwards yearly. Examination is targeting on recognition of recurrences, new primary skin cancers, and lymph node metastasis. Follow-up for SCC in situ is not necessary.

Malignant Melanoma

Pathophysiology

Malignant melanoma is the most aggressive type of skin cancer which derives from melanocytes in the skin (Rastrelli et al. 2014). Melanocytes are located in the basement layer of the epidermis. They transfer melanosomes to keratinocytes via dendritic processes. Nevi are proliferated nest of melanocytes with loss of dendritic processes. Almost all melanoma starts with a radial growth phase, which are restricted to the epidermis or hair follicle epithelium. These are the melanoma in situ. Next, the vertical growth phase starts and the melanoma grows invasively through the basement membrane into the dermis.

The most common subtypes of malignant melanoma are superficial spreading melanoma, lentigo malign melanoma, acral lentiginous melanoma, and nodular melanoma (Rastrelli et al. 2014). Rare variants of malignant melanoma are nevoid melanoma, desmoplastic melanoma, clear cell sarcoma, and solitary dermal melanoma. Superficial spreading melanoma are the most common and represents 75% of all melanoma.

Risk factors are environmental and genetic. Sunburn in childhood and sunbed usage under the age of 35 increases the risk of melanoma (Marks and Whiteman 1994). This is most important for fair skinned people with red or blond hair. A family history of one or more family members with malign melanoma also increases the risk of melanoma. Familial CDKN2A gene mutation gives an increased risk for malign melanoma and is associated with pancreatic cancer.

Clinical Features

Early detection of melanoma is important. Dermoscopy is used during physical examination to improve recognition of malignant melanocytic lesions. The first step is to determine if a lesion is melanocytic. If so, the following tools can be used for better and earlier detection.

The ABCDE acronym is developed for primary care givers and patients to differentiate nevi from malign melanoma (Abbasi et al. 2004). ABCDE stands for Asymmetry; Border irregularities; Color variation such as brown, black, red, gray, blue, or white; Diameter $\geq$6 mm; and Evolving (change of size, shape, color, or new lesion.)

Another tool is the "ugly duckling" sign (Gaudy-Marqueste et al. 2017). Patients with multiple nevi are screened for pigment lesions that are obviously different from other nevi. This is suspicious and needs further investigation, even when it does not fulfill the ABCDE criteria.

Patients with clinically suspected melanoma need to be quickly referred to a dermatologist (Fig. 9).

Diagnosis

Clinically suspected melanoma need to be histologically confirmed. Excisional biopsy with a margin of 2 mm normal skin and subcutaneous fat is needed (Marsden et al. 2010). The pathologist confirms the diagnosis and determines the Breslow depth, which is the strongest predictor of survival. Breslow tumor thickness is a measurement of micro staging of cutaneous melanoma from the top of the granular layer of the epidermis to the deepest part of the tumor in millimeters.

Melanoma have a combination of histopathological features which varies between melanoma subtypes. Atypical melanocytes and architectural disorder are main criteria for diagnosis. Atypical melanocytes are big cells with hyperchromatic, irregular, and/or polymorphic nuclei. Architectural disorder consists of asymmetry and/or nests of irregular melanocytes in parts of epidermis and dermis. Immunohistochemistry is used in difficult lesions. S-100, MART-1, and HMB-45 are markers that are often used.

After diagnosis, the tumor node metastasis (TNM) staging of the AJCC staging system is

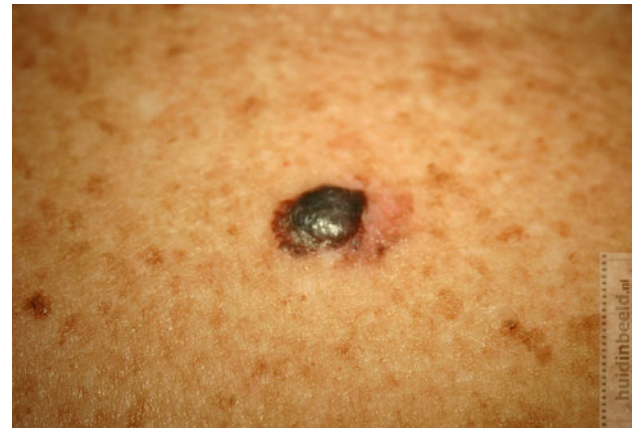


Fig. 9 Malignant melanoma

used for determination of prognosis group. This provides an accurate prognosis of life expectancy (Gershenwald et al. 2017) (Fig. 10).

Treatment

Primary Cutaneous Melanoma

After diagnostic excision, a wider excision of normal tissue is required to ensure complete removal of the primary tumor and micro metastases. The margin of normal tissue removal is based on the Breslow thickness. Melanomas $\leq$1 mm thick are resected with a 1 cm margin, and melanomas $>$1 mm are resected with a 2 cm margin. In situ melanomas are resected with a 0.5 cm margin (Marsden et al. 2010).

Primary melanomas of $>$0.8 mm thickness and $<$0.8 mm with ulceration are indicated for lymphatic mapping and sentinel lymph node biopsy.

Chemotherapy and adjuvant radiation therapy is not indicated for primary cutaneous melanomas without lymph node or distance metastasis (Marsden et al. 2010).

Regional Lymph Nodes

Sentinel lymph node biopsy is recommended for patients without clinical signs of lymph node involvement and with an increased risk of metastasis. This include primary cutaneous melanomas of $>$0.8 mm or $<$0.8 mm with high risk features, such as ulceration. Patients with a positive sentinel lymph node without distant metastasis can be carefully observed with frequent nodal basin ultrasounds or complete lymph node dissection (Wong et al. 2012).

Patients with a clinically positive lymph node which is confirmed with fine needle aspiration cytology need to undergo complete lymph node dissection. Preoperative investigations for distant metastasis need to be performed for accurate staging (Rodrigues et al. 2000).

Adjuvant immunotherapy with anti-programmed cell death 1 (PD-1) agents or target therapy is recommended for patients with positive lymph nodes without distant metastasis and is discussed below.

Metastatic Melanoma

Since the use of new immunotherapy treatments and target therapies, the life expectancy of metastatic melanomas has been improved. Nowadays, chemotherapy and radiotherapy are only used for palliative purposes in a limited number of patients.

Prior to treatment, a full body investigation is performed with CT and/or MRI, and serum lactate dehydrogenase is tested for prognosis. Besides, tumor cells are screened for driver mutations at the V600 site in BRAF gene.

- **Immunotherapy** – Combining checkpoint inhibitors nivolumab (PD-1) and ipilimumab (CTLA-4) results in an enhanced T-cell function against melanoma cells and therefore

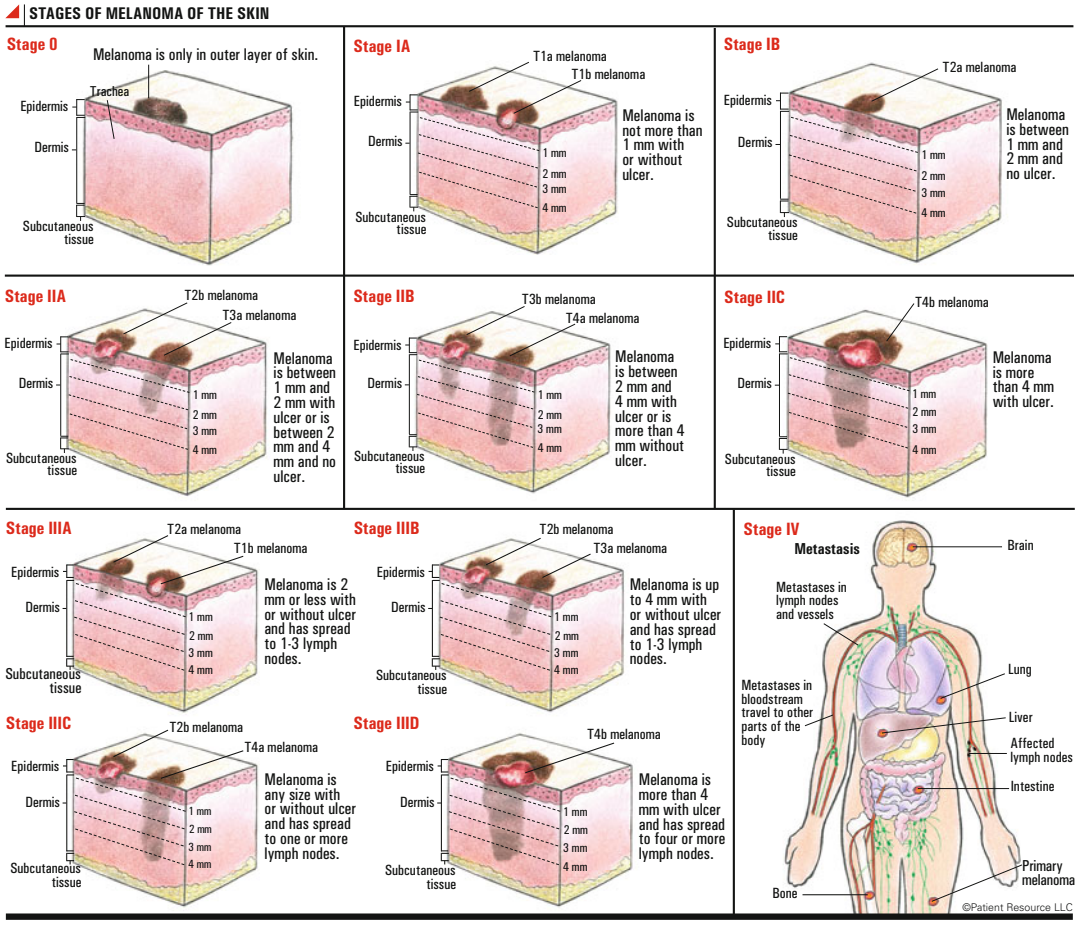


Fig. 10 TNM classification of malignant melanoma (Reprinted with permission from Patient Resource LLC. © 2019 Patient Resource LLC)

provides an antitumor response. Combination therapy provides better response rates and survival-free periods than single therapy (Larkin et al. 2015).

Nivolumab is a human IgG4 monoclonal antibody which inhibits programmed cell death-1 (PD-1) activity by binding to PD-1 receptor on T cells. PD-L1 and PD-L2 ligands are blocked from binding, causing reverse T-cell suppression, and induce antitumor responses. Pembrolizumab is a human monoclonal checkpoint inhibitor of PD-1 as well, but used in a lesser extent.

Ipilimumab is a recombinant human IgG1 monoclonal antibody that works by binding to cytotoxic T-cell associated antigen 4 (CTLA-4). CTL-4 inhibits T-cell activation. Blockage of CTLA-4 leads to T-cell activation and proliferation against melanoma cells.

- **Target therapy** – Patients with BRAF V600 mutation, about fifty percent of cutaneous melanomas, are candidates for target therapy against the mitogen-activated protein kinase (MAPK) pathway. This consists of BRAF-inhibitors (dabrafenib and vemurafenib) as monotherapy or combined with MEK-inhibitors (trametinib and cobimetinib).

BRAF kinase inhibitors (dabrafenib and vemurafenib) inhibit the formation of activated BRAF proteins which are important for cell

proliferation and differentiation in melanoma with the mutation. Therefore, BRAF kinase inhibitors cause tumor regression and or tumor growth inhibition (Chapman et al. 2011).

MEK-inhibitors (trametinib and cobimetinib) inhibit selectively mitogen-activated extracellular kinase (MEK) 1 and 2 activity. MEK 1 and 2 are part of the BRAF pathway. MEK-inhibitors provide an extra tumor growth inhibition when combined with BRAF kinase inhibitors (Flaherty et al. 2012).

- **Surgical metastasectomy** – Surgical resection provides a durable benefit when a small number of metastases occur, usually in the brain.
- **Radiation therapy** – Radiation therapy is used as palliative treatment mostly in brain metastases when surgical resection is not possible (Marsden et al. 2010).
- **Chemotherapy** – Chemotherapy does not contribute to overall survival in metastatic melanomas and is therefore not used as treatment (Marsden et al. 2010).

Follow-Up

Follow-up for nonmetastatic melanoma consists of regular visits to detect local recurrences, regional metastasis, and second primary melanoma. Follow-up interval varies between 3 and 6 months, depending on the recurrence risk. Beside, patients need to be instructed about sun protection and self-examinations at home (Marsden et al. 2010).

Follow-up for metastatic melanoma with immunotherapy consists of CT of chest, abdomen, and pelvis every 3 months for the first 2 years. MRI of the brain is preformed every 6 months. After 2 years, the interval is 6 months to 1 year. Patients who receive BRAF inhibitors undergo CT and MRI scans every 3 months.

Patients with cutaneous melanoma who have two or more first degree family members with melanoma and/or family history of pancreatic cancer should consider genetic testing for genetic mutations, such as CDKN2A mutation.

References and Further Reading

Abbasi NR et al (2004) Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. JAMA 292(22):2771

Alam M, Ratner D (2001) Cutaneous squamous-cell carcinoma. N Engl J Med 344(13):975

Beck LA et al (2014) Dupilumab treatment in adults with moderate-to-severe atopic dermatitis. N Engl J Med 371(2):130

Bissonnette R et al (2016) Topical tofacitinib for atopic dermatitis: a phase IIa randomized trial. Br J Dermatol 175(5):902–911

Boehncke WH, Schön MP (2015) Psoriasis. Lancet 386(9997):983–994

Boguniewicz M, Leung DY (2011) Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. Immunol Rev 242(1):233

Brodland DG, Zitelli JA (1992) Surgical margins for excision of primary cutaneous squamous cell carcinoma. J Am Acad Dermatol 27(2 Pt 1):241

Chapman PB et al (2011) Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med 364(26):2507

Charman CR, Venn AJ, Williams HC (2004) The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. Arch Dermatol 140:1513–1519

Criscione VD et al (2009) Actinic keratoses: natural history and risk of malignant transformation in the Veterans Affairs Topical Tretinoin Chemoprevention Trial. Cancer 115(11):2523

Crowson AN (2006) Basal cell carcinoma: biology, morphology and clinical implications. Mod Pathol 19: S127–S147

Eichenfield LF et al (2014) Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. J Am Acad Dermatol 71(1):116. Epub 2014 May 9

Ferrucci LM et al (2012) Indoor tanning and risk of early-onset basal cell carcinoma. J Am Acad Dermatol 67(4):552. Epub 2011 Dec 9

Flaherty KT et al (2012) Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. N Engl J Med 367(18):1694. Epub 2012 Sep 29

Gallagher RP et al (1995) Sunlight exposure, pigmentary factors, and risk of nonmelanocytic skin cancer. I. Basal cell carcinoma. Arch Dermatol 131(2):157

Gaudy-Marqueste C et al (2017) Ugly duckling sign as a major factor of efficiency in melanoma detection. JAMA Dermatol 153(4):279

Gershenwald JE, Scolyer RA, Hess KR et al (2017) Melanoma of the skin. In: Amin MB (ed) AJCC cancer staging manual, 8th edn. American Joint Committee on Cancer, Chicago, p 563

Gorlin RJ, Goltz RW (1960) Multiple nevoid basal-cell epithelioma, jaw cysts and bifid rib. A syndrome. N Engl J Med 262:908

Guttman-Yassky E (2018) Efficacy and safety of fezakinumab (an IL-22 monoclonal antibody) in adults with moderate-to-severe atopic dermatitis inadequately controlled by conventional treatments: a randomized, double-blind, phase 2a trial. J Am Acad Dermatol 78(5):872–881

Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M (2001) The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. Exp Dermatol 10(1):11–18

Kuflik EG (2004) Cryosurgery for skin cancer: 30-year experience and cure rates. Dermatol Surg 30(2 Pt 2):297

Larkin J et al (2015) Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. N Engl J Med 373(1):23

Leung DY (2013) Why is eczema herpeticum unexpectedly rare? Antiviral Res 98(2):153–157. Epub 2013 Feb 22

Locke J et al (2001) Radiotherapy for epithelial skin cancer. Int J Radiat Oncol Biol Phys 51(3):748–755

Love WE et al (2009) Topical imiquimod or fluorouracil therapy for basal and squamous cell carcinoma: a systematic review. Arch Dermatol 145(12):1431

Marks R, Whiteman D (1994) Sunburn and melanoma: how strong is the evidence? BMJ 308:75–76

Marsden JR et al (2010) Revised U.K. guidelines for the management of cutaneous melanoma 2010. Br J Dermatol 163(2):238. Epub 2010 Jul 1

Marzuka AG, Book SE (2015) Basal cell carcinoma: pathogenesis, epidemiology, clinical features, diagnosis, histopathology, and management. Yale J Biol Med 88(2):167–179. eCollection 2015 Jun

Mason AR et al (2013) Topical treatments for chronic plaque psoriasis. Cochrane Database Syst Rev 3: CD005028

Menter A et al (2008) Guidelines of care for the management of psoriasis and psoriatic arthritis: section 1. Overview of psoriasis and guidelines of care for the treatment of psoriasis with biologics. J Am Acad Dermatol 58(5):826

Menter A et al (2009) Guidelines of care for the management of psoriasis and psoriatic arthritis: section 4. Guidelines of care for the management and treatment of psoriasis with traditional systemic agents. J Am Acad Dermatol 61(3):451

Morton CA et al (2013) European guidelines for topical photodynamic therapy part 1: treatment delivery and current indications – actinic keratoses, Bowen's disease, basal cell carcinoma. J Eur Acad Dermatol Venereol 27(5):536. Epub 2012 Nov 26

Nestle FO et al (2009) Psoriasis. N Engl J Med 361(5):496

Ports WC et al (2013) A randomized phase 2a efficacy and safety trial of the topical Janus kinase inhibitor tofacitinib in the treatment of chronic plaque psoriasis. Br J Dermatol 169(1):137

Rastrelli M et al (2014) Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. In Vivo 28(6):1005–1011

Ring J et al (2012) Guidelines for treatment of atopic eczema (atopic dermatitis) Part II. J Eur Acad Dermatol Venereol 26(9):1176–1193. Epub 2012 Jul 19

Rodrigues LK et al (2000) Fine needle aspiration in the diagnosis of metastatic melanoma. J Am Acad Dermatol 42(5 Pt 1):735

Rudikoff D, Lebwohl M (1998) Atopic dermatitis. Lancet 351(9117):1715

Ruzicka T et al (2017) Anti-interleukin-31 receptor A antibody for atopic dermatitis. N Engl J Med 376(9): 826–835

Schmitt J et al (2005) The psoriasis area and severity index is the adequate criterion to define severity in chronic plaque-type psoriasis. Dermatology 210(3):194–199

Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis (1993) Dermatology 186(1):23–31

Simpson EL et al (2018) Efficacy and safety of lebrikizumab (an anti-IL-13 monoclonal antibody) in adults with moderate-to-severe atopic dermatitis inadequately controlled by topical corticosteroids: a randomized, placebo-controlled phase II trial (TREBLE). J Am Acad Dermatol 78(5):863–871

Thyssen JP, Kezic S (2014) Causes of epidermal filaggrin reduction and their role in the pathogenesis of atopic dermatitis. J Allergy Clin Immunol 134(4):772–799

Tollefson MM, Bruckner AL (2014) Atopic dermatitis: skin-directed management. Pediatrics 134(6):e1735

Van Loo E et al (2014) Surgical excision versus Mohs' micrographic surgery for basal cell carcinoma of the face: a randomised clinical trial with 10 year follow-up. Eur J Cancer 50(17):3011–3020. Epub 2014 Sep 25

Weichenthal M et al (2005) Phototherapy: how does UV work? Photodermatol Photoimmunol Photomed 21(5): 260

Weidinger S, Beck LA, Bieber T, Kabashima K, Irvine AD (2018) Atopic dermatitis. Nat Rev Dis Primers 4(1):1

Wong SL et al (2012) Sentinel lymph node biopsy for melanoma: American Society of Clinical Oncology and Society of Surgical Oncology joint clinical practice guideline. J Clin Oncol 30(23):2912

Yanofsky VR et al (2011) Histopathological variants of cutaneous squamous cell carcinoma: a review. J Skin Cancer 2011:210813. Epub 2010 Dec 29

Pharmacodynamic Evaluation: Inflammation/Immunology

14

Jörg Schüttrumpf and Matthias Germer

Contents

Abstract

In any drug development including those for candidates in acute and chronic inflammation and any other immunological disbalance, distinct types of pharmacological studies are a regulatory requirement. Together clinical and nonclinical data are required to characterize a novel drug candidate adequately. These requirements were summarized, outlined, and exemplified in this chapter as follows:

- Primary pharmacology describes the mode of action of the drug candidate with respect to its therapeutic target.
- Secondary pharmacology: Studies on secondary pharmacology look deeper into the biochemical and physiological effects of a drug and the mode of action that is not directly related to the desired therapeutic target (general pharmacology studies).
- Safety pharmacology has a focus on the impact of a drug candidate in and above the therapeutic range on the function of organ systems.
- Immunological safety assessment: First-in-human studies.
- Immunotoxicology.

As there is no international consensus on definitions of the terms, the attribution of either type of studies may sometimes be ambiguous. It is impossible to separate immunotoxicology studies required by regulatory agencies basically from the pharmacodynamics of a drug. An initial package of such pharmacological studies is a prerequisite for first-in-human studies for novel drugs.

J. Schüttrumpf (✉) · M. Germer
Corporate R&D, Biotest AG, Dreieich, Germany
e-mail: joerg.schuettrumpf@biotest.com; matthias.germer@biotest.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_34

Primary Pharmacology

The intended mode of action of a drug candidate relevant for the intended immunomodulatory use is specific for each substance. Pharmacopoeias sometimes describe the relevant assays for common substances in some detail which are representative for the drug's mode of action and must therefore also be employed in the testing prior to batch release. For the majority of substances and in particular for any novel product, however no standardized and accepted tests are available. Rather the manufacturer starts to investigate the mode of action early in drug development and ideally develops specific functional assays. These should also be indicating stability and the starting point of the development of precise and robust assays for batch release. As every product is different in this respect, there is no general guidance on the required tests and data in preclinical and clinical research. Nevertheless it is important to understand the performance characteristics of any pharmacological test to allow for the interpretation of the data.

Predictive pharmacological in vitro or ex vivo models for disease play a limited role in immunology because this system critically depends on the interactions of a network of different soluble mediators and cell types and on the location of the cellular components in the organs. But also in vivo disease models often are of limited value because of substantial species-specific differences of the immune system in human and laboratory animals even including primates. Many disease models show some similarity to the human disease but differ substantially, e.g., in pathophysiology, disease progression, and therapeutic readout. Models for rheumatoid arthritis as an example for chronic inflammation and for tumors as an example for inadequate immune response are presented here.

Rheumatoid arthritis in humans is a chronic autoimmune disease typically progressing over many years before joints and cartilage become irreversibly damaged. In contrast the best studied animal model is arthritis in mice, rat, or even primates induced by the intradermal injection of collagen type II (Trentham et al. 1978). It has similarities in many clinical, histological, and immunological phenomena and was essential in the development on cytokine inhibitors, the current blockbusters in the arthritis therapy. Nevertheless the differences are also substantial: disease onset in this model is rapid (within weeks) and progression fulminant. The immune response in animal models is dominated from an innate immunity and by anti-collagen antibodies which are of little relevance in patients. In contrast to the response in patients, rodents benefit little from nonsteroidal anti-inflammatory drugs. Reversal of joint destruction is not observed. Alternative models have become available and can dissect various pathomechanisms in the disease initiation and progression. Also the brake of immunological tolerance is still poorly understood. HLA class II transgenic mice stains can differ in predisposition to develop rheumatoid arthritis. Mice with mutations in the T-cell receptor gene, ZAP70 (SKG mice), and with specific variants of the protein tyrosine phosphatase non-receptor type 22 need an additional environmental stimulus. Serum transfer models can be used to study the role of immune complexes in the inflammatory response. TNF-alpha transgenic mice develop chronic progressive polyarthritis spontaneously and very reliably within 6 weeks after birth. Their primary use is the investigation of the pathological production of TNF-alpha. No current model of arthritis can however fully represent the human disease. As all arthritis models are associated with pain and distress, the use of such models is debatable, and the implementation of the 3R principles is obligatory: replacement, reduction, and refinement. Despite of their limitations, the collagen-induced arthritis models have remained popular (Bessis et al. 2017; Benson et al. 2017).

Any progression of tumors can be seen as a failure of the immune system to eradicate aberrant cells. Unfortunately oncology is a field notorious for successful therapy of laboratory animal but failure to predict safety and efficacy in clinical trials (Mak et al. 2014). The process of human carcinogenesis, physiology, and progression is complex. Typical animal models in oncology include subcutaneous implantation of cultured human tumor cells (xenografts) into immunocompromised mice. This procedure is easy and can be

standardized. However the tumor environment is known to have a tremendous impact on treatment susceptibility and often is not the skin. Usually immunecompromised mice have to be employed to allow for adequate tumor cell survival and progression. Therefore, important interactions of the tumor and the host immune system will be missed in animal models. Also tumor cells are typically grown ex vivo before implantation and divide rapidly, whereas human tumor cells divide much slower. As a consequence the susceptibility to cytotoxic agents in the models can be overestimated. Tumor xenograft models remain valuable because of the established technology and the wide variety of tumor types that can be studied. But the ability of such models to mimic more than just single steps of the complex processes in patients and their predictive value are limited.

Individual limitations of the tumor models have been overcome in recent decades. When tumor grafts are directly explanted from patients, not only the malignant cells but also the correct anatomical structure can be transferred (e.g., stromal cells, bone chips). Variability of the response of a specific tumor entity can be assessed better. Because the immune system of mice and men are phenotypically and functionally too different, humanized tumor mouse models have been another major recent progress. By transplantation of CD34-positive human stem cells, a functional human immune system can be established in otherwise immunocompromised mice. When human tumor cells are transplanted in parallel or after establishment of the human immune system, the interaction of these systems can be taken into account. This has been particularly useful in the evaluation of immunomodulatory approaches (antibodies, tumor vaccines, cell therapy, cytokines). Genetically engineered mice may completely represent cancer development (initiation, progression, interaction with stroma and the immune system). Tumors from such animals can be transplanted in syngeneic animals and have been used to study chemotherapeutics and other small molecules and currently are probably the best available tool to study immunotherapeutics. But still species differences can be a major challenge. Future refined models might better address tumor variability, better control of genetic variability of the tumor and the host, clinically relevant endpoints, biomarkers, environmental factors, etc. (Day et al. 2015).

Other immunological indications have seen similar progress in pharmacological animal models. The pathomechanism, the initiation and progression of the disease in the animal model, and the patient should be understood and similar. For validation of an animal model, it can be helpful to demonstrate that a clinically effective intervention is effective in the model as well.

It is important to plan pharmacological animal studies with a clear rational and a well-characterized animal model. Also an adequate study design (e.g., group size) is required to verify or reject a hypothesis, although limited availability of novel drugs, chosen animal species, or ethical considerations make it often necessary to keep the number of animals low. Still the translation of preclinical pharmacological results to the clinical setting can be most reliable if common design flaws are minimized. In clinical practice therapy is started after the onset of symptoms and when diagnosis is clearly established. In animal models often a drug is given shortly after or even before the disease is initiated, and the therapeutic benefit is likely to be overestimated. Misleading results may also be obtained when the animal cohorts do not match the clinical situation in terms of age, gender, environmental factors, or concomitant diseases.

Today the use of irrelevant species in preclinical studies is often discouraged. Studies to proof the relevance of an animal model do not only include demonstration of reactivity with the relevant target structure. Also a good understanding of the target's primary structure, distribution and expression level, pharmacological effects, metabolism, and cross-reactivity with other structures is advisable.

In clinical studies the primary pharmacological effect is often also investigated and correlation to clinical endpoints analyzed. The test systems are specific for each substance because of the unique mode of action of each drug candidate. Due to the complex nature of the immunological, often cell-based, assays, it is not uncommon that only specialized laboratories can perform such assays

reliably. In many cases robust results are only obtained when cellular assays are set up within a few hours after samples are taken. This can limit the use of central laboratory facilities in larger, multinational studies, and relevant data may only become available from a subset of patients/volunteers. Nevertheless they can be very helpful as the basis of pharmacometric models.

Secondary Pharmacology

During drug development secondary pharmaceutical effects are often not fully understood. Typically these studies are started early in development because they help to find an appropriate starting dose for first-in-men clinical trials. But the understanding of their biological relevance increases over time even after drug approval. Some of these effects can be consistently observed in experimental setups but remain inconspicuous in patients. In other cases they can contribute to the adverse event profile of a drug (Amantea and Bagetta 2016). They may also be crucial for upside indications or drug repurposing (for examples see Table 1).

Pharmacological effects are particularly relevant for patient safety in cases of unintended enhancement or suppression of immune responses. Investigations of these aspects are also often described as immunotoxicity studies. Novel chemical entities and biologicals somewhat differ in the complexity of the test strategy. The field has developed rapidly which is often not yet reflected in regulatory guidance documents. These investigations are driven by substance-specific characteristics which are different for every drug candidate. Current characterization of a drug candidate is very much science driven and rarely follow a standardized approach.

Immunological Safety Assessment: First-in-Human Studies

A sound understanding of the pharmacological effects is required in the development of any drug. In 2006 the particular importance of early availability of these data became apparent for novel drugs. A phase I clinical trial with the superagonistic CD28 antibody TGN1412 changed the development path not only for biologicals but also for novel chemical entities. Although a TGN1412 dose 500 times smaller than estimated safe from animal studies was infused in a first-in-men study, all healthy volunteers suffered from a life-threatening cytokine storm. Within an hour all participants were

Table 1 Examples of biologically relevant secondary pharmaceutical effects (Yan et al. 2016; Koo et al. 2010)

Contribution to adverse event profile		
Drug	Indication	Proposed mechanism
Kadcyla (anti-HER2/neu-DM1 antibody-drug conjugate)	Breast cancer	Thrombocytopenia due to cytotoxic effect on megakaryocytes; hepatotoxicity by uptake of the conjugate by hepatocytes
Orthoclone (CD3 antibody OKT3)	Transplant rejection	Cytokine release syndrome due to release by activated leukocytes potentially causing reversible renal function impairment and delayed graft function
TNF antagonists	Rheumatoid arthritis, psoriatic arthritis, Crohn's disease, ankylosing spondylitis	Opportunistic infections due to impaired macrophage activation and differentiation, neutropenia
Repurposing		
Drugs	Original use	New use
Methotrexate Cyclophosphamide	Solid tumors	SLE
Mycophenolic acid	Solid organ transplantation	ITP, pemphigus vulgaris, SLE
Alemtuzumab	Chronic lymphocytic leukemia	Multiple sclerosis

reported to suffer from severe fever, rash, rigors, etc. Within the first 4 h, they developed hypotension, tachycardia, and respiratory distress which rapidly developed into multi-organ failure. Due to massive swelling of the arms and legs, the study is often referred to the "Elephant Man" drug trial in the press. Pharmacology of TGN1412 had been investigated in vitro and in vivo. It induced proliferation of T cells without co-stimulation. Rodent CD28 was hardly bound, but nonhuman primates were reported to have sequence identity in the extracellular part and high binding affinity. Whereas rodents were regarded as irrelevant model, the pharmacology and toxicology of TGN1412 were investigated in rhesus and cynomolgus monkeys. No signs of toxicity and in particular neither immune dysregulation nor hypertension were observed. In addition, a rat CD28 antibody was investigated in relevant species, and the expected pharmacological effect was observed, i.e., an increase in T cells. Importantly, increases in cytokine levels were only moderate, and no signs of cytokine release syndrome, anaphylaxis, or autoimmunity were observed. At the time the first conclusion of the apparently unpredictable outcome of the clinical trial was that the pharmacological studies in animal models may show very different properties in humans.

After the incident extensive research has led to a number of hypotheses why the apparently adequate pharmacological data packages did not predict the activity in humans. Species differences in signal transduction might be responsible. Whereas in rat the mutation of PI3 kinase does not affect signaling of CD28 superagonists in humans, inhibition of this kinase blocked the production completely. Whereas the drug induced a rapid and marked calcium flux in humans, the calcium flux in cynomolgus was only weak. Some Siglec molecules which are receptors belonging to the immunoglobulin superfamily carry inhibition motifs (ITIMs). On human T cells, their expression is low, but in chimpanzees various members are expressed. Pharmacological in vitro models could be identified that may better mimic the situation in humans involving different cell densities, interaction with Fc receptors, and other cell types. Apparently, only rodents CD28 superagonists induced the expansion of regulatory T cells which can rapidly counteract an immune response.

Immediately after the TGN1412 incident, regulatory authorities (e.g., in Germany, France, the UK) gave recommendations on how to minimize the risk of phase I clinical trials with novel substances. Only 1 year after the incident, this was then covered by guidelines by the FDA (US Department of Health and Human Services, Food and Drug Administration 2005) and the EMA (currently updated, Committee for Medicinal Products for Human Use (CHMP) 2017) which are applicable not only to biologicals but also to new chemical entities. Before the TGN1412 incident, the starting dose was usually selected based on toxicological data. On the basis of a no-observed-adverse-effect level (NOAEL) of repeated-dose toxicology studies using the most sensitive species, the human equivalent dose and consequently the maximum recommended starting dose were calculated with the help of conservative conversion and safety factors (e.g., for species or route of application). Since the TGN1412 incident, pharmacological effects are also taken into consideration: the no-observed-effect level (NOEL), i.e., the highest dose tested in mice with no effects detected (e.g., on functional activity, morphology, life span) and the minimal anticipated biological effect level, i.e., the anticipated dose level leading to a minimal biological effect level (MABEL) in humans. For the determination of the MABEL, all relevant in vitro and in vivo pharmacological and if applicable pharmacokinetic data are considered, including, e.g., receptor occupancy, dose-response curves from in vitro and in vivo animal studies with relevant species, and also in vitro studies with human cells. In healthy volunteers the starting dose should be lower than the pharmacologically active dose (PAD). Also the importance of pharmacokinetic/pharmacodynamic models is increasing (Yu et al. 2011; Diao and Meibohm 2015).

Whenever possible, alternatives to the use of live animals should be employed, e.g., in vitro studies using human-derived materials. When studies in animals are required, the relevance of

the model should be demonstrated. Key parameters are target structure and expression and the interaction with the drug (affinities, receptor occupancy, and kinetics), pharmacodynamics, and pharmacokinetics including metabolism. Not only the molecular mechanism related to the intended therapeutic use should be understood, but also secondary pharmacodynamic characteristics are expected to be investigated. Typically, functional responses related to target binding are evaluated, e.g., cellular responses including their dose dependency and duration in relation to the turnover of the drug and its target. Pharmacometric models describing and predicting the relation between pharmacokinetics and pharmacodynamics have become an important routine tool for most stages of drug development.

Whereas the use of one relevant and one other species is acceptable for novel chemical entities, studies in irrelevant animal models should be avoided for biologicals. Instead studies in a single relevant species are often sufficient. In vitro human cellular test systems might be advantageous because of lack of species differences such as affinities for the target, expression patterns, signal transduction, regulation, and metabolism. Still pharmacological studies in animal models including disease models can provide additional understanding of the mode of action in vivo, on appropriate doses and dosing intervals, and safety.

Immunotoxicology

When pharmacological mechanisms of a drug are supposed to provoke dysfunction of the immune system, often the term immunotoxicity is used. The immunotoxic effect refers to any unintended modulation of the immune response and is not limited to exaggerated pharmacological effects. These effects also include, e.g., cytotoxic effects on immune cells and interactions with surface structures of immune cells or on corresponding, interacting tissues. This immunotoxic effect may either manifest as an insufficient or excessive immune response (immunosuppression or, alternatively, allergy, autoimmunity, chronic inflammation). The extent, type, and timing of such studies depend on the specific, functional properties of the drug but also on the characteristics of the patient population (e.g., immunocompromised patients). They are typically required before large patient populations are exposed. Due to the nature of the drugs, the requirements for small molecule and biologicals somewhat differ (see guidelines ICH S8 and ICH S6 (R1), respectively). For biologicals immunogenicity is a concern that may be of major clinical relevance. As human proteins are immunogenic in laboratory animals, relevant immunogenicity data are typically obtained in clinical development.

Discussion

The pharmacodynamic evaluation is essential for the selection of appropriate doses and designs in clinical trials for indications of inflammation and immunological disorders. Understanding of the pharmacodynamic properties is relevant to develop effective and safe drugs. Still the translation of preclinical results to the clinical situation must be made with caution. Testing of molecular pathways, cellular components, cells, or even tissues is only moderately predictive. The complexity of the human organism cannot be reduced to an array of individual in vitro assays. Also in vivo studies in animal models are limited in their predictive value. Especially in the field of immunology, numerous species differences can lead to misinterpretation of the data. Disease models often differ substantially both in the pathomechanism and the progression of the disease from their human counterparts. For biologicals species specificity of the drug is a further major challenge. Last but not least, patients are diverse. Their genetic heterogenicity can have an impact on drug metabolism and transport, disease characteristics, and many other parameters. Furthermore, patients change over their life, e.g., with age or with their diseases. Nevertheless, preclinical pharmacological evaluation is still a scientific valuable basis to design good clinical development strategies to minimize patient safety risks,

to optimize efficacy aiming to maximize the benefit for the patient, and to best show the full potential of a new drug.

References and Further Reading

Amantea D, Bagetta G (2016) Drug repurposing for immune modulation in acute ischemic stroke. Curr Opin Pharmacol 26:124–130

Benson RA, McInnes IB, Garside P, Brewer JM (2017) Model answers: rational application of murine models in arthritis research. Eur J Immunol. https://doi.org/10.1002/eji.201746938

Bessis N, Decker P, Assier E, Semerano L, Boissier MC (2017) Arthritis models: usefulness and interpretation. Semin Immunopathol 39(4):469–486

Committee for Medicinal Products for Human Use (CHMP) (2017) Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products. EMEA/CHMP/SWP/28367/07 Rev. 1

Day CP, Merlino G, Van Dyke T (2015) Preclinical mouse cancer models: a maze of opportunities and challenges. Cell 163(1):39–53

Diao L, Meibohm B (2015) Tools for predicting the PK/PD of therapeutic proteins. Expert Opin Drug Metab Toxicol 11(7):1115–1125

ICH S6(R1) (1997) Preclinical safety evaluation of biotechnology-derived pharmaceuticals. CPMP/ICH/302/95

ICH S8 (2005) Immunotoxicity studies for human pharmaceuticals EMEA/CHMP/167235/2004-ICH

Koo S, Marty FM, Baden LR (2010) Infectious complications associated with immunomodulating biologic agents. Infect Dis Clin N Am 24(2):285–306

Mak IWY, Evaniew N, Ghert M (2014) Lost in translation: animal models and clinical trials in cancer treatment. Am J Transl Res 6(2):114–118

Trentham DE, Townes AS, Kang AH, David JR (1978) Humoral and cellular sensitivity to collagen in type II collagen-induced arthritis in rats. J Clin Invest 61(1):89–96

U.S. Department of Health and Human Services, Food and Drug Administration (2005) Guidance for industry: estimating the maximum safe starting dose in initial clinical trials for therapeutics in adult healthy volunteers

Yan H, Endo Y, Shen Y, Rotstein D, Dokmanovic M, Mohan N, Mukhopadhyay P, Gao B, Pacher P, Wu WJ (2016) Ado-trastuzumab emtansine targets hepatocytes via human epidermal growth factor receptor 2 to induce hepatotoxicity. Mol Cancer Ther 15(3):480–490

Yu J, Karcher H, Feire AL, Lowe PJ (2011) From target selection to the minimum acceptable biological effect level for human study: use of mechanism-based PK/PD modeling to design safe and efficacious biologics. AAPS J 13(2):169–178

Pharmacodynamic Evaluation: Infectious Diseases

15

Smita Bhuyan, Sebastian Felgner, Dino Kocijancic, and Vinay Pawar

Contents

S. Bhuyan
Department of Microbial drugs, Helmholtz Center for Infection Research, Braunschweig, Germany
e-mail: Smita.Bhuyan@helmholtz-hzi.de

S. Felgner · D. Kocijancic
Department of Molecular Immunology, Helmholtz Center for Infection Research, Braunschweig, Germany
e-mail: Sebastian.Felgner@helmholtz-hzi.de; Dino.Kocijancic@helmholtz-hzi.de

V. Pawar (✉)
Department of Molecular Immunology, Helmholtz Center for Infection Research, Braunschweig, Germany

Institute of Immunology, Medical School Hannover, Hanover, Germany
e-mail: vinay.pawar@helmholtz-hzi.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_53

Abstract

The past decades have witnessed an rise in the number of multidrug resistant bacteria, concurrent drug failures, and declining cure rates of several fatal bacterial infectious diseases. This has culminated in evaluation of existing drug regimens. An extensive evaluation of pharmacodynamics and pharmacokinetics of drugs and its application for drug regimen selection has become the cornerstone for successful antibiotic therapy against fatal bacterial infections. Recent years have witnessed an upsurge in the development of diverse models for preclinical and clinical pharmacokinetic-pharmacodynamic (PK-PD) analysis and robust simulation methods. Integration of infection microbiology knowledge and PK-PD analysis in an appropriate model can lead to optimal drug regimens against several acute/chronic bacterial infections. The present chapter provides a comprehensive overview of different models citing their advantages and limitations, along with simulations for optimizing treatment regimens. Furthermore, it describes the applications of pharmacodynamic models for treatment of bacterial infections and, finally, the pathophysiological conditions leading to treatment failures and strategies to overcome them.

Introduction

Rise of multidrug resistant (MDR) bacterial strains and long-term persistence of bacteria in chronic infections pose a critical challenge for public health. Immediate treatment regimen using currently available antibiotics or new formulations is of urgent need. Appropriate choice of antibiotics, doses, and treatment is absolutely essential for optimal therapy. Therefore, large numbers of antibiotics or antibiotic combinations are presently under investigation in clinical trials to treat a variety of acute and chronic fatal infectious diseases. For appropriate selection of antibiotics, it is imperative to gather the pharmacokinetic/pharmacodynamic (PK/PD) information using the different models (Schuck et al. 2005). An excellent background information on the dose-exposure relationship – pharmacokinetics (PK) – and the exposure-response relationship – pharmacodynamics (PD) – can lead to decreased antimicrobial resistance (AMR) and focused clinical trials with improved efficacy and cost-effectiveness. A quantitative representation of the dose–concentration–response relationship provides information that can be utilized to predict the level of response corresponding to a particular drug dose. Models based on different mathematical approaches can be used to describe such relationships. The mathematical relationship is determined by whether single dose or the steady-state measurements are carried out (Pérez-Urizar et al. 2000). Profound knowledge on PK/PD is bifaceted and would help to establish a computable relationship between dose and dosing regimen as well as potency and undesirable drug effects. Additionally, simulating the drug clearance profiles obtained from animal/human studies in different in vitro PK/PD model could lead to a detailed characterization of efficacy of antimicrobials (Vaddady et al. 2010). Meibohm and Derendorf (2002) provide a comprehensive overview on pharmacokinetic/pharmacodynamic studies in drug development. Various studies have been initiated involving extensive PK/PD modeling for a variety of infectious diseases and can be beneficial to implement strategies for successful disease elimination. Recent research in the development of powerful models, statistically robust software tools, and the integration of

pharmacokinetic-pharmacodynamic knowledge has led to excellent decision-making and development of effective dosage regimens for various infectious diseases. In the present chapter, we describe in brief terminologies, methods, models, and issues associated with pharmacodynamics of bacterial infectious diseases.

Pharmacokinetics and Pharmacodynamics

The term pharmacokinetics refers to the relationship between drug dosing and changes in the drug concentration over time in the body (Udy et al. 2008; Nielsen and Friberg 2013; Mouton 2014). The important PK properties include: (i) Drug clearance: Volume of plasma effectively cleared of the drug from each eliminating organ/tissue per unit time (CL); (ii) Volume of distribution: Apparent volume of fluid containing the total amount of administered drug, at the same concentration as that in plasma (Vd); (iii) Peak concentration achieved by a single dose (C_{max}); (iv) Lowest concentration during the dosing period ($C_{\min}$); (v) Area under the plasma concentration-time curve (AUC); and (vi) Plasma half-life ($T_{1/2}$). These PK parameters have a profound influence on the PD parameters of a drug.

The term pharmacodynamics refers to the biochemical and physiological effect of the drug. In case of an infection, it refers to the ability of the drug to either inhibit or abolish the growth of a causative organism. In other words, antibiotic pharmacodynamics integrates the complex relationship between the susceptibility of an organism and the patient pharmacokinetics. The PD indices most correlated to bacteriological eradication are: (i) Time for which the drug concentration remains above the MIC value of the infectious agent during a given dosing period (T > MIC); (ii) Ratio of maximum/peak antibiotic concentration ($C_{\max}$) achieved by a single dose of antibiotic to the MIC of the infectious agent (C_{max}/MIC); and (iii) Ratio of the area under the plasma concentration time curve to the MIC of the infectious agent (AUC_{24}/MIC) (see Fig. 1). The prefix, f, is sometimes introduced to indicate that the free, unbound

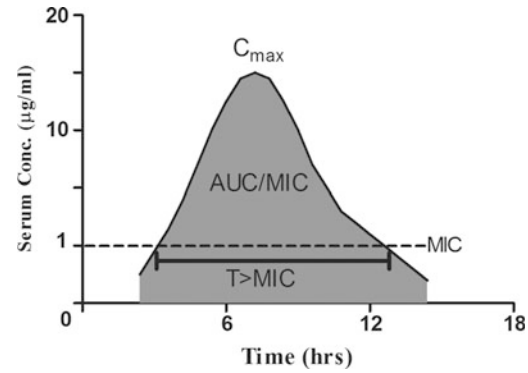


Fig. 1 Pharmacodynamic indices of antibiotics (MIC: Minimum inhibitory concentration; C_{max}: Maximum/peak antibiotic concentration achieved by a single dose of antibiotic; T > MIC: time for which the drug concentration remains above the MIC value of the infectious agent during a given dosing period; AUC/MIC: ratio of the area under the plasma concentration time curve to the MIC of the infectious agent)

fraction of the drug was used in the calculations. When no subscripts are included, it is assumed that the calculations of AUC and T > MIC were based on a 24-h interval at pharmacokinetic steady-state conditions (Mouton et al. 2005; Nielsen and Friberg 2013; Mouton 2014).

MIC values are usually determined using either the agar dilution or broth microdilution methods as specified by the CLSI guidelines. Typically, broth dilution methods use liquid medium in which a specified bacterial inoculum [5×10^5 colony-forming units (CFU)/ml] is exposed to a constant antibiotic concentration generally over an incubation period of 16–20 h. The MIC is defined as the lowest drug concentration that completely inhibits visible growth of the microorganism. The antibiotic concentrations chosen for MIC determinations are typically twofold dilutions of the antibiotic (e.g., 0.5, 1, 2, and 4 concentration units). Depending on the total volume used, the method is either termed macrodilution (1–2 ml) or microdilution ($\leq$ 500 μl). For agar diffusion methods, an agar plate is inoculated with the target organism and the antibiotic diffuses from a disk or a strip into the agar. The results are generally read off after 24 h. An E-test is a semiautomated agar diffusion test, which contains a strip preimpregnated with an exponential

gradient of the antibiotic (Nielsen and Friberg 2013; Wayne 2014). Cumulative fraction of response (CFR) can be calculated using the discrete distribution of MIC values. The AUC_{24} can be obtained directly from previously published literature. In other cases, it can be calculated as follows: $AUC_{24} = Dose/CL$.

A particular efficacy index better corelates to a given antibiotic/group of antibiotics (Zhanel 2001). Though MIC values are the most studied pharamcodynamic parameter, it gives incomplete information regarding activity of the antibiotic over time. Hence, the parameters mentioned above are most widely used as a measure of the PD index. The PK/PD index for a certain drug-bacteria combination is determined by plotting the value of an efficacy endpoint versus the magnitude of each of the three PK/PD indices.

It is evident that the pharmacodynamic properties directly depend on the infectious agent, severity of infection, phase of infection, and the MIC value or resistance pattern of the pathogen. This clearly indicates that the local MIC distribution must be taken into consideration in order to achieve the maximum likelihood of a successful antibiotic treatment regimen for an infectious disease (Canut et al. 2012). Extensive and accurate analysis of PD properties is of foremost importance in determining sustained in vivo efficacy of antimicrobial agents.

Optimization of Drug Dosage Regimen

Determination of PK/PD parameters is the first step in optimization of dose regimen for treatment of infectious diseases. They increase the likelihood of disease eradication and minimize the probability of exposure-related toxicity (Scaglione and Paraboni 2006). Once the population PK model has been developed and the PD parameters determined, they can be transferred to models to design dosing regimens that aim to achieve the recommended PD values affecting antimicrobial response. Monte Carlo simulations have been widely used to reliably predict the probability of achieving the PD parameters in case of antimicrobials. Monte Carlo simulation is a technique that integrates an agent's in vitro potency distribution with the pharmacokinetic profile to achieve a specifically targeted antimicrobial exposure (Nicolau 2003). Laboratory based in vitro or animal data, preclinical, or clinical PK/PD data are used in Monte Carlo simulations to obtain initial PK/PD breakpoints. The MIC distributions of the target populations based on this values are then determined. The robustness of the Monte Carlo simulation, target population, and dose adjustments are then made to determine the final PK/PD breakpoints for disease eradication (Mouton et al. 2012).

Pharmacokinetic and Pharmacodynamic (PK/PD) Models

The determination of PK/PD indices can be carried out using in vitro pharmacodynamic models (IVPM) or animal models. Ethical issues restrict the use of clinical subjects to evaluate PK/PD relationships in the field of anti-infectives. Hence, IVPM are increasingly used to determine PD indices to be utilized as an aid to dose selection and optimization for treatment of infectious diseases. Both preclinical and clinical PK/PD studies are then incorporated into simulations to determine the PK/PD breakpoints for antimicrobials (Mouton et al. 2012).

The in vivo or in vitro kill curves (described below) are generally used to build models to estimate PD parameters. It is imperative that different dosage schemes or concentrations of the anti-infective agent are applied in a PK/PD study. Placebo doses are mandatory and it is essential that varying dose intervals, multiple doses, and information on kinetics of absorption, dissimilation, and excretion are incorporated to guarantee the success of a model. PD parameters could be estimated by using nonlinear regression analysis, and it is absolutely crucial that the model takes all possible survival curves of the organisms under study into account (Czock and Keller 2007).

In Vitro *Models*

The PD in vitro models are broadly divided into static and dynamic models based on whether the concentration of the tested drug is constant or varying over a period of time. The static models usually consist of determination of MIC or continuous monitoring in shake flasks (Vaddady et al. 2010). Static in vitro time-kill studies provide information on killing capacity of a drug and the probability of emergence of antimicrobial resistance, although determination of MIC would be sufficient in clinical practice to determine the efficacy of antibiotics. However, it is of minimal use, as it represents an end-point detection method, therefore, giving no information on pharmacodynamic changes taking place over a period of time. Hence, static models based on MIC determination are of little relevance in development of dosing recommendations. It is therefore imperative that data should be collected using dynamic situations (Czock and Keller 2007; Gloede et al. 2010; Tängdén et al. 2017). Dynamic in vitro models permit continuous adjustment of drug concentrations to mimic the required in vivo PK profile. Target microbes are exposed to the antimicrobials in vessels that are perfused continuously with media to simulate actual conditions during an infection. Sampling to determine PK/PD properties can be done repetitively. This enables the study of various parameters like kill kinetics, drug concentrations, and emergence of resistance. A large number of dynamic models are in vogue, which includes dilution models (i.e., direct contact between infectious agent and the drug) or diffusion/dialysis models (i.e., indirect drug bacteria contact). A detailed description of all models is beyond the scope of this chapter. However, for a comprehensive reading the reader can refer to the published review by Michael et al. 2014, Vaddady et al. 2010, and Czock and Keller 2007. Different models can be used to mimic drug efficacy in humans. These include biofilm models, models for human immune system, multicompartment models, or models to study disease conditions such as otitis media, chronic pneumonia, cystic fibrosis, or tuberculosis (Gloede et al. 2010; Parra-Ruiz et al. 2010; Pawar et al. 2014; Lorenz et al. 2016). In vitro models have some advantages over in vivo animal models as they provide more flexibility and adaptability to the researcher and are comparatively less expensive and resource-intensive. However, they face certain disadvantages like the need of controlled environments and the risk of contamination of the culture vessel. Neither can these models completely mimic all in vivo conditions like the immunological response to and the virulence nor the metabolic behavior of a pathogen. Furthermore, the bacterial growth limits the analysis as it is much faster in vitro than in vivo (Gloede et al. 2010).

In Vivo *Models*

Evaluation of antibiotic therapy using animal models is essential for the evaluation of the therapeutic efficacies of antimicrobial agents. The advantages of animal models over in vitro models are enormous. Animal infection models allow to study drug efficacy with regard to virulence or antimicrobial resistance of an organism. Furthermore, they also relate to the role of the host immune system in response to the infection itself. Thus, it enables to study the antimicrobial effects at the exact site of infection and can mimic/simulate the conditions in humans and thereby human PK/PD profiles in contrast to in vitro models. However, the results from animal models must be interpreted with caution since the antimicrobial PK profile may turn to be extremely different from human subjects. Rodent, rabbit, or more recently porcine hosts are most predominantly used in PK/PD studies. Mice and rats are preferred due to low cost and handling ease in comparison to other animals (Tängdén et al. 2017). Establishment of infection may require the animals to be rendered neutropenic by prior administration of an immunosuppressant like cyclophosphamide in order to appropriately compare these results to those that might be expected in humans. Murine tumor models have been used to study the efficacy of

antimicrobials in patients inflicted by cystic fibrosis and burn wound infections caused by *Salmonella* and *Pseudomonas aeruginosa* (Crull and Weiss 2011; Pawar et al. 2015). Other common models are the murine thigh infection models, pneumonia model, peritonitis/bacteremia models, skin and soft tissue infection model, meningitis models, and endocarditis models (Andes and Craig 1998; Nielsen and Friberg 2013; Rybtke et al. 2015).

Mathematical Approach to Modeling

PK-PD modeling is the mathematical description of the relationships between PK and PD. The choice of the model and the underlying mathematical equation can depend on whether the system under study is in a steady state or an unsteady time-dependant phase. Steady state refers to a condition in that, the concentrations of the active form of the drug at the site of action are constant and the PD parameters are independent of time as in case of long-term intravenous infusions. These models assess how a bacterial culture responds to a constant environment and fixed antibiotic exposure. The growth of the infectious agent is limited by nutrition, space, aeration, and toxic metabolites. When the concentration and response data are in phase or steady state, basic models such as fixed-effect, linear, log-linear, E_{MAX}, and sigmoidal E_{MAX} models are used. When the kinetics and response are out of phase time-variant, pharmacodynamic models which are more complex are applied. These dynamic models utilize time kill curves, where microbial killing is dependent on both time and varying antibiotic concentration. Models can also be referred to as mechanistic, semimechanistic, or nonmechanistic (Pérez-Urizar et al. 2000; Vaddady et al. 2010). A mechanistic model is a model which takes into account the known or hypothesized mechanisms of behavior of an infectious agent. The parameters are in accordance with PK, physicochemical, biophysical, physiological, and pathophysiological principles of the system under consideration and relate drug concentrations to their observed effect. Nonmechanistic models do not take the underlying biological mechanisms into consideration (Vaddady et al. 2010; Felmlee et al. 2012). Semimechanistic models are those in which although mechanistic knowledge is utilized, but are far less complex compared to the mechanistic models. These are also referred to as mechanism-based models. In general, mechanism based PK-PD models contain equations describing microbial growth, effect of antimicrobial drug, and variable drug concentrations (the microorganism submodel, the antimicrobial submodel, and the pharmacokinetic submodel, respectively) (Czock and Keller 2007; Nielson et al. 2011). The final choice of PK/PD model is made based on the pharmacology of the drug and system. Once a model is defined, unknown parameter values are typically estimated using nonlinear regression techniques contained within computer programs such as WinNonlin (Pharsight, Mountain View, CA), Kinetica (Innaphase, Philadelphia, PA), and ADAPT II (Biomedical Simulations Resource, Los Angeles, CA) (Mager et al. 2003). Below we describe some of the commonly applied in vitro models used to rationalize the selection of antibiotics based on the PK/PD characteristics.

Linear Model

This model is based on the assumption that a direct proportionality between drug concentration and its effect exists (1).

$$E = S * C + E_0 \tag{1}$$

Where S is the slope, E_0 the intercept. Pharmacodynamically, S represents the effect induced by one unit of C and E_0 represents the value in the absence of the drug. The parameter estimations are carried out by linear regression, and this model applies to measured effects with physiological baselines such as blood glucose or blood pressure levels.

Log-Linear Model

Log linear model takes into consideration that if the effect of concentration is hyperbolic, the log-concentration-effect relationship would roughly

be linear in the range of 20–80% of maximal effect. This can be considered as a derivation of Eq. 1 where S represents the change elicited by one unit of log *C*.

$$E = S * \log C + E_0 \tag{2}$$

E_{MAX} Model (Hill Equation)

This PK-PD model is extensively used to characterize a wide range of pharmacological effects. The model describes the relationship between concentration of a drug and its elicited effect relationship over a wide range of concentrations. Equation 3 assumes that the plasma drug concentration is in rapid equilibrium with the effect site.

$$E = \frac{E_{max} * C}{EC_{50} + C} + E_0 \tag{3}$$

E_{MAX} describes the maximum effect possible, EC_{50} the concentration required to produce 50% of E_{MAX}, and E_0 is the basal value E. This equation is also referred to as the Hill equation.

Sigmoidal E_{MAX} Model

This model is a derivate of the E_{max} model and allows convenient fitting of different types of PK/PD data. This is the most frequently used model due to the fact that the function asymptotes to an upper limit of stimulation or inhibition by a particular drug on the target infectious agent. Here, γ represents the steepness of the curve also referred to as sigmoidicity factor. $\gamma > 1$ for steep curve, $\gamma < 1$ for a smooth curve, and $\gamma = 1$ for a hyperbolic curve, while other parameters are same as in the E_{max} model.

$$E(t) = E_0 + \frac{E_{max} * C(t)^{\gamma}}{EC^{\gamma}{}_{50} + C(t)^{\gamma}} \tag{4}$$

The Bacterial Submodel

The simplest mechanism-based is the bacterial submodel and involves a single bacterial compartment. It is adapted from a model developed initially for anticancer agents. The model has a first-order rate constant for bacterial multiplication (K_{growth}) and a first-order rate for the death of the bacteria (K_{death}) as shown in Eq. 5.

$$\frac{dB}{dt} = k_{growth} * B - k_{death} * B \tag{5}$$

The equation accounts for observed exponential growth of bacteria as seen in the time-kill curve experiments in absence of drug (control experiments) as the net result of the growth rate and cell death.

The Logistic Growth Model

Most of the modern day antimicrobial models are based on the logistic growth model, which can be used to describe in vitro bacterial population dynamics. It is a very simple yet useful model and is based on the growth rate (r) and the carrying capacity of the environment (K). N is the bacterial population and N_0 the initial bacterial count. It is given by the Eq. 6

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N \tag{6}$$

The logistic growth model is sometimes modified to include the effect of the drug with a new equation as follows

$$\frac{dN}{dt} = k_{growth}N\left(1 - \frac{N}{N_{max}}\right) - f_{death}\,(drug) \tag{7}$$

Where f_{death} (drug) is a function that accounting for death of the bacteria due to the antibiotic and N_{max} is the maximum number of bacteria.

Pharmacokinetic-Pharmacodynamic Model

Pharmacokinetic-pharmacodynamic model is a combination of the bacterial submodel and the PK model. The combined equations characterize the effect that the antibacterial drug has on the bacteria. The effect could be hypothesized to

either inhibit the bacterial growth rate or enhance the bacterial killing rate.

$$\frac{dB}{dt} = k_{growth} * \left(1 - \frac{E_{max} * C^{\gamma}_{(t)}}{E\,C^{\gamma}{}_{50} + C^{\gamma}_{(t)}}\right) * B - k_{death} * B \quad (8)$$

Apart from these models, several other models and variants of the basic model or combinations of model are in use for PK/PD analysis. Detailed descriptions of various models have been described by different researchers. For further reading the reader can refer to the publications from Dayneka et al. (1993), Sharma and Jusko (1998), Czock and Keller (2007), Vaddady et al. (2010), Felmlee et al. (2012), and Mouton (2014).

Simulation Aided PK/PD for Infectious Diseases

Computer simulations have been widely used in diverse fields. In the field of pharmaceuticals, it is used in the discovery of new drugs, optimizing chemical processes, and, most recently, in designing clinical studies for the treatment of several acute and chronic diseases. Molecular modeling is the best-known example of simulation in drug discovery. In recent years, Monte Carlo simulation of clinical trials is the method of choice for appropriate dosing selection (Mouton et al. 2012). Monte Carlo is a different kind of simulation than the traditional ones by the fact that its model parameters are treated as stochastic or random variables, rather than as fixed values. In other words, the variability of the parameters is included in the model and the long-term impact of that variability is examined. Furthermore, by definition, Monte Carlo simulation is a random number generator that incorporates distributions of variability around PK parameters in a population to simulate drug concentration-time profiles for a large number of conjured individuals. Thus, instead of practically studying different concentrations over different periods of time, Monte Carlo simulation allows intensive analysis of outcome of different trails, even when data from single doses are coupled to the simulation (Crandon and Nicolau 2011).

How is a Monte Carlo simulation performed? First, the underlying structural pharmacokinetic model for the given antimicrobial agent against a particular infection is defined. This can be a done in a single compartment, 2-compartment, or multiple compartment models. One-compartment model assumes that elimination is first order and that PK parameters are independent of the dose and that there is immediate distribution and equilibrium of the drug throughout the body. However, 2-compartment model does not follow linear PK and comprise of absorption, along with the distribution and covariance between the pharmacokinetic parameters in the model. Next, a dose administration model and a compliance model are defined. This includes the number of patients being administered with a particular dose, number of patients who skipped the dose, and other such criteria related to the intake of the drug. Once the conceptual model is defined, it is translated into a computer code using different software programs like MATLAB, GAUSS, or the Pharsight Trial Designer. Once this is done, it is verified for accuracy and the simulation run. The number of replications of the simulation must be defined at the start. The number of replicates depends on the nature of analysis. Larger replicates are used when the variability of an outcome is to be studied. This is followed by the generation of a sequence of independent random numbers having a given distribution with finite mean and variance. Depending on the pharmacokinetic profile, the pharmacodynamic end point is simulated. The end point is assessed again after few days of treatment. Once the inputs are defined and the simulation is performed, the outputs are examined in the form of graphical results or summary statistics of a variable or the relationship between variables. This general protocol is used for simulating any drug to be used in clinical practices. Some of the important components among others to be defined in a clinical trial simulation include structural pharmacokinetic model, dose administration model, distribution and covariance of pharmacokinetic parameters, link between PK and PD, pharmacodynamic model, disease progression

model, relationship between pharmacodynamic effect, and outcome and survival model (Bonate 2001).

Pharmacodynamics of Antibiotics for Infectious Diseases

Gram-Positive Bacterial Infections

Large numbers of fatal infections are caused by various Gram-positive bacteria. For instance, *Staphylococcus aureus* is responsible for a wide range of infections including hospital acquired bacterial pneumonia, ventilator associated bacterial pneumonia, complicated skin and skin structure infections, and severe bacteremia. The most common drug of choice for such concurrent *S. aureus* bacteremia is Telavancin. Telavancin is a bactericidal lipoglycopeptide, which is also effective against methicillin susceptible and resistant *Staphylococcus aureus* (MSSA & MRSA) (Wilson et al. 2017). Additionally, Vancomycin and Linezoid have been widely used for the treatment of infections caused by MRSA. Over the years, most bacteria have evolved resistance mechanisms and hence it is imperative that timely preclinical and clinical studies are undertaken to evaluate the efficacy of such antibiotics. The underlying PK data and MIC can either be retrieved from data sets or databases or calculated for the infection under study. In an interesting study, Canut et al. (2012) have evaluated the usefulness of Daptomycin, Tigecycline, and Linezolid for the treatment of MRSA infection and compared it with vancomycin in four western European countries. They have estimated the probability of achieving the recommended value of AUC_{24}/MIC ratio using Monte Carlo simulation technique with 10,000 subjects. They calculated the fC_{min} using Eq. 9.

$$fC_{min} = \frac{f_u D}{V_d} \frac{e^{\frac{0.693}{t_{1/2}} * \tau}}{1 - e^{\frac{0.693}{t_{1/2}} * \tau}} \tag{9}$$

where f_u is the free fraction of drug in plasma, D the administered drug dose, V_d the volume of distribution, and $t_{1/2}$ the half-life elimination. For the analysis, steady-state exposure was evaluated for different intravenous drug dosing regimens, MIC values fixed at a particular concentration and then the probability of target attainment (PTA) calculated from these parameters. A regimen that achieved >90% CFR against bacterial population is considered as optimal. Their studies indicate that 2 g, 3 g, and 4 g daily of vancomycin seem be adequate in Belgium, Spain, and United Kingdom/Ireland, respectively. CFR obtained with 50 mg Tigecycline every 12 h was higher in Spain than in Belgium and the United Kingdom/Ireland. Additionally, a minimum of 8 mg/kg Daptomycin is necessary in United Kingdom/Ireland, while 4 mg/kg may be sufficient in Spain. The authors concluded that differences in the susceptibility of MRSA strains among countries may be responsible for differences in the antibiotic dose selection and suggest use of local MIC values to achieve success of a PK/PD model to achieve eradication of disease condition. As part of preclinical study, the effect of antimicrobials against infections caused by *Staphylococcus aureus* and *S. epidermidis* was evaluated in a novel in vitro PK/PD model of bacterial biofilm by Hall Snyder et al. 2015. Some persistent bacteria that can cause chronic infections are resistant to antibiotics due to their inherent property to form biofilms. Biofilms shield the bacteria against antibiotics and thereby posing an imminent threat especially in chronically infected or immunocompromised patients. Biofilm forming bacteria are difficult to eradicate from severe infection sites such as lungs of cystic fibrosis patients. They are equally notorious and resistant to eradication when associated with medical implants (Taraszkiewicz et al. 2013; Sanchez et al. 2013). Hence, this model by Hall Synder and coworkers could be further extended to study antimicrobials targeting other Gram-positive biofilm forming bacteria. The in vitro model consists of a CDC biofilm reactor (CBR) modified to run PK/PD and simulating human PK in order to evaluate the in vitro activity of antimicrobials. Biofilm conditioning is performed prior to initiation of drug therapy initiation followed by continuous flow with peristaltic pumps in specific media. Upon completion

of conditioning and continuous flow phases, boluses of antibiotics are injected into the reactor. Free drug concentrations were used, and simulated regimens were included. The model was used to study effect of high doses of Daptomycin versus Vancomycin either alone or in combination with Clarithromycin or Rifampin to treat infections caused by *Staphylococcus* species. Biofilm-embedded cell concentrations (mean and standard deviation in CFU/cm^2) were then computed. Time kill curves were plotted to determine total reduction in CFU counts and the therapeutic enhancement of combination regimens was calculated statistically. Pharmacokinetic samples were obtained through the injection port of each model. PK and PD parameters to verify target antibiotic concentrations were obtained at the same time points. PK parameters were estimated using routine procedures. The half-life ($t_{1/2}$), area under the curve (AUC), and fC_{max} were determined by the trapezoidal method utilizing software tools like the PK Analyst software (Hall Snyder et al. 2015). It is essential to maintain appropriate growth controls and replicates for any set of antibiotics to be tested in such model systems. Their study highlights that combinations of Daptomycin + Rifampin and Vancomycin + Rifampin were the most effective against biofilm-associated staphylococcal infections. Daptomycin + rifampin with the best activity emerged as a promising drug combination that could be used as a future regimen to treat resistant biofilm-associated staphylococcal infections. A similar but novel in vitro biofilm model has also been developed by Parra-Ruiz et al. in 2010. The authors used their model to assess the in vitro activities of several antimicrobials alone or in combination against *Staphylococcus aureus* isolates. Daptomycin, Vancomycin, and Moxifloxacin were evaluated either alone or in combination with Clarithromycin or Rifampin. This study was also performed using an in vitro model, which consisted of a CDC biofilm reactor wherein the concentrations of biofilm-embedded bacteria were computed and plotted to graph time-kill curves. The results clearly indicate that combinations of Daptomycin, Moxifloxacin with Clarithromycin were the most effective ($P < 0.01$) regimens. This may represent future regimen to treat persistent infections of biofilm forming Gram-positive bacteria such as *S. aureus* and *Streptococcus pneumoniae*.

Gram-Negative Bacterial Infections

Antimicrobial resistance in Gram-negative bacteria is increasing at an alarming rate. Additionally, the lack of new antibiotics limited treatment option to reappraise of currently available antibiotics. Cefepime, Ceftriaxone, Imipenem, and Piperacillin-Tazobactam antibiotics have been extensively used worldwide during the last three decades to treat acute/chronic hospital acquired infections. These antibiotics are resourceful in treatment of infections caused by Gram-negative bacteria like *Pseudomonas, Acinetobacter, Klebsiella, Enterobacter, Serratia, Stenothermophilus, Proteus,* and *Citrobacter* (Zervos and Nelson 1998; Drago and De Vecchi 2008; Saltoglu et al. 2010). In the past decades, antibiotic resistance is on the surge, and PK/PD modeling based on Monte Carlo simulations can be used reliably to predict the efficacy of antimicrobial regimes against an array of Gram-negative bacteria. This can foster microbial eradication and speed up the recovery rates of infectious diseases (Bonate 2001; Eagye et al. 2007). A two compartment multiple dose model was used by Eagye et al. in 2007 to determine 24-h concentration-time profile at steady-state conditions for different drugs against Gram-negative bacteria causing fatal infections. Patient-derived pharmacokinetic values combined with 5000 trial Monte Carlo simulation was used to determine the predicted cumulative fraction of response (CFR) values. The probability that the selected antibiotic regime will either meet or exceed a predefined pharmacodynamic target at a given MIC dilution (PTA) was calculated for different antibiotics. The target index $fT > MIC$ was selected as the PD property of the drugs. The results of the study indicate that for Carbapenems, 40% $fT > MIC$ was considered bactericidal; 50% $fT > MIC$ for Piperacillin-Tazobactam and Cephalosporins. A CFR of 90% was the set threshold for reliable empirical

therapy. A p-value of ≤ 0.05 indicated a statistically significant outcome. The confidence intervals were calculated at $\alpha = 0.05$ using Newcomb Wilson method without correction for continuity.

A novel standardized time kill-curve assay has been developed by Foerster et al. 2016. This assay and the subsequent pharmacodynamic modeling can be used to evaluate existing and novel antimicrobials against different strains of *Neisseria,* which have developed resistance to first-line empirical monotherapy. GraverWade (GW) medium, which supports the growth of a wide range of *N. gonorrhoeae* auxotypes and clinical isolates, was used. The time kill assay is a useful assay as described earlier and can be used for any antibiotic-pathogen combination. For time-kill curve analyses, *N. gonorrhoeae* was grown in GW medium in the presence of the desired antibiotics, covering a range of dilutions. A 0.5 McFarland inoculum of the test strains has to be prepared. Following this, 30 µl of the inoculum was diluted in 15 ml prewarmed (37 °C) antimicrobial-free GW medium and 90 µl per well was dispensed in round bottom 96-well microtiter plates. Plates were preincubated and 10 µl of the antimicrobial concentrations (or PBS in case of drug free control) was added to each well containing the reincubated bacteria. The growth rate was estimated as the coefficient of a linear regression from the logarithm of the colony counts. Pharmacodynamic model with Eq. 10 as described by Regoes et al. 2004 was used in the study.

$$\psi(a) = \psi_{max} - \frac{\left[(\psi_{max} - \psi_{min})(a/zMIC)^k\right]}{\left[(a/zMIC)^k - (\psi_{max}/\psi_{min})\right]} \quad (10)$$

ψ_{max} describes the maximal growth rate of the bacteria in the absence of antimicrobial agent, while ψ_{min} represents the minimal bacterial growth rate at high antimicrobial concentrations. zMIC is the pharmacodynamic MIC value where the bacterial growth rate is zero (ψ (zMIC) = 0). k is the Hill coefficient, describing the steepness of the sigmoidal relationship between bacterial growth and antimicrobial concentration. All parameters were estimated using the R software package and bacterial growth rates are estimated from the time-kill curves by linear regression. The pharmacodynamic model was finally fitted to the estimated growth rates at different antimicrobial concentrations. The authors suggest that PD parameters based on a wide range of concentrations below and above the MIC can provide valuable information that could be useful for improving future dosing strategies.

PD parameters obtained using time kill analysis coupled to evaluate the potential of Ciprofloxacin as an anti-infective agent (Schuck et al. 2005). Ciprofloxacin has been widely used for the treatment of a variety of nosocomial infections, especially in intensive care units (ICUs). Ciprofloxacin is a second-generation fluoroquinolone, which exhibits rapid concentration-dependent bactericidal activity against most Gram-negative aerobic species like *Pseudomonas aeruginosa* and *Acinetobacter baumannii.* A recent investigation by Khachman et al. (2011) provides an extensive and simplified methodology for optimizing ciprofloxacin dosing in ICU patients mediated by the use of population pharmacokinetic–pharmacodynamic analysis and Monte Carlo simulations. This model can be used as a guideline for studying various other antibiotics active against Gram-negative pathogens. The population PK model was based on a clinical study employing 102 ICU patients. Two sets of PK/PD simulations were carried out with each 10,000 patients simulated per dosage regimen investigated. The first set included complete distribution of MIC values including susceptible and resistant strains, while the second set was carried out across each MIC value according to a geometric progression from 0.002 mg/L to 2 mg/L. This model was also a predictor to determine possible continued usage of ciprofloxacin for the treatment of Gram-negative infections in ICU patients due to their potential to develop resistances. Such population PK/PD modeling could be of great value for studying efficacy of existing and novel antimicrobials against infectious diseases. Colistin or its inactive prodrug Colistin methanesulfonate (CMS) is increasingly used as

a last-line therapy to treat infections caused by multidrug-resistant (MDR) Gram-negative pathogens (Dudhani et al. 2010a). With significant improvements in the understanding of the chemistry, PK/PD characteristics and their interrelationship, substantial progress has been made in optimizing use of CMS in clinics. This includes the first scientifically based dosing algorithm for critically ill patients receiving CMS to generate a desired target steady-state plasma concentration of formed Colistin. Most PD data on Colistin has been generated using in vitro models. Colisitin exhibits a strong concentration-dependent killing against *P. aeruginosa, A. baumannii, and K. pneumoniae* and their MDR strains as indicated time-kill studies in static and dynamic systems. A consistent finding of both in vitro and in vivo studies is regrowth with Colistin monotherapy, even with concentrations above those which can be safely achieved clinically (Nation et al. 2014). Although owing to increased resistance as indicated by regrowth with Colistin monotherapy in in vitro and in vivo studies, combination therapy may be beneficial. It still appears that Colistin is the most promising antimicrobial for severe infections caused by Gram-negative bacteria. Studies have been performed employing dose-fractionation design to investigate the PK/PD index, which best correlates with Colistin efficacy. The overall killing effect was best correlated with fAUC/MIC followed by fT > MIC (Nation et al. 2014). Studies in neutropenic thigh and lung mouse infection models again indicated that the PK/PD index that best correlated with Colistin efficacy was fAUC/MIC. Dose-fractionation studies with Colistin were conducted against *P. aeruginosa* strains, its MDR clinical strain, and a strain from cystic fibrosis patient to determine this index. The relationships between antibacterial effects and PD parameters were examined using an inhibitory sigmoid maximum-effect model (Dudhani et al. 2010a). Similar studies for other antimicrobials can be conducted to define optimum dosage regimens in humans. The models of antimicrobial therapy designed for planktonic infections are often of little importance in treating infections by biofilm forming bacteria. Additionally, models derived from two compartment analysis simulating the blood and various tissues as two compartments may also not be completely mimicking the exact conditions in biofilms. In 2015, Cao et al. have designed a new in vitro model which could mimic the conditions during the biofilm formation by bacteria. This seaweed alginate-embedded biofilm model allows the simultaneous measurement of antibiotics within the matrix and parallel bacterial killing. The authors hypothesized that biofilms may be considered as independent compartments with particular pharmacokinetics. The biofilm model was used to study the antibiotic penetration and concomitant killing of *Pseudomonas aeruginosa* by Tobramycin. Tobramycin is the drug of choice for treating biofilm-associated infections of *P. aeruginosa* like cystic fibrosis (Herrmann et al. 2010). Additionally, Pawar et al. 2015 showed the biofilm associated tolerance toward clinical antibiotics in murine tumor model. Such in vivo biofilm models will be able to simulate the clinical pharmacokinetics of antibiotics and could be an ideal model for testing new treatment strategies.

Pharmacodynamic Indices for Optimal Therapy

As described earlier, for any antimicrobial agent a particular PD parameter correlates best with the successful eradication of an infectious agent. Time above MIC (T > MIC) correlates best with the activity of β lactams. On the hand, AUC_{24}/MIC is adjudged the best for aminoglycosides and fluoroquinolones. However, sometimes in case of these two classes of antibiotic Cp_{max}/MIC is used (Zhanel 2001). A detailed account of PK/PD indices for all antibiotics is beyond the scope of this chapter, so here we describe representatives of the most widely studied values. Additionally, Table 1 briefly describes few antibiotics and their pharmacodynamic parameters.

For β lactams, a minimum T > MIC of 40–50% of the dosing interval is required to exercise bactericidal effects and a corresponding bacteriological cure of greater than 85%. Maximal bactericidal effects are achieved when T > MIC ranges between 60 and 70% of the dosing interval

Table 1 Pharmacodynamics of clinically available antibiotics

Antibiotic	Effective against/ Treatment of	Relevant pharmacodynamic parameter	Setting	Reference
Colistin	*Acinetobacter baumannii*	*f*AUC/MIC	Murine thigh and lung infection model	Dudhani et al. 2010b
Colistin	*Pseudomonas aeruginosa*	*f*AUC/MIC	In vitro model	Bergen et al. 2010
Tobramycin	*Pseudomonas aeruginosa*	*f*AUC/MIC, T > MIC	In patient studies	Mouton et al. 2005
Levofloxacin	*Pseudomonas aeruginosa*	AUC_{24}/MIC, Cp_{max}/MIC	Phase IV clinical trial	Lee et al. 2007
Amikacin	Gram-negative bacteria	T > MIC & log AUC	Thigh infection and pneumonia in mice models	Craig et al. 1991
Cefepime, Ceftriaxone, Imipene, Piperacillin-Tazobactam	Gram-negative bacteria like *Pseudomonas, Acinetobacter, Klebsiella, Enterobacter, Serratia, Stenothermophilus, Proteus, Citrobacter*	T > MIC	In vitro two compartment model	Eagye et al. 2007
Trovafloxacin, Ciprofloxacin	*Staphylococcus aureus*	AUC/MIC	In vitro dynamic model	Firsov et al. 1999
Daptomycin, Vancomycin, Tigecycline, Linezolid	Invasive MRSA infections including *S. aureus* bacteremia	AUC_{24}/MIC	In vitro PK/PD model of bacterial biofilm	Hall Snyder et al. 2015
Ciprofloxacin	Broad spectrum, treatment of acute uncomplicated and complicated urinary tract infections and uncomplicated pyelonephritis	AUC_{24}/MIC	Healthy human volunteers	Schuck et al. 2005
Telavancin, vancomycin plus Aztreonam or Piperacillin-Tazobactam	Mixed infection by *Pseudomonas aeruginosa*, *Escherichia coli,* and methicillin-resistant *Staphylococcus aureus*	Reduction in $\log_{10}$ CFU/ml	Clinical trials	Yim et al. 2016
Ceftriaxone-Sublactam	Bacteria causing complicated urinary tract infections	T > MIC	Clinical trials	Sharma et al. 2016

AUC_{24} – area under the curve at 24 h; Cpmax – maximum serum concentration; MIC – minimal inhibitory concentration; T – time; T > MIC – time over MIC; CFU – colony forming units; ml – milliliter

as suggested by animal studies and clinical trials in otitis media (Craig and Andes 1996; Andes and Craig 1998). Additionally, animal studies indicate that for stasis, a *f*T/MIC of 20% is required for Carbapenams, 30% for Penicillins, and 40% for Cephalosporins. Maximal efficiency also described as 2 log reduction in CFU requires *f*T/MIC 40% for Carbapenams, 50% for Penicillins, and 50–70% for Cephalosporins (Crandon and Nicolau 2011). In a murine thigh infection model to evaluate the effect of Amoxicillin and Amoxicillin-Clavulanate against *Streptococcus pneumoniae,* highest mortality rates (80–100%) were seen when serum levels exceeded the MIC for less than 20% of the dosing interval. Maximal survival was approached when serum levels exceeded the MIC for 40% of the 8-h dosing interval (Andes and Craig 1998). Their studies

demonstrate that a reduction of 1 $\log_{10}$ or greater in CFU/thigh at 24 h is consistently observed when amoxicillin levels exceed the MIC for 25–30% of the dosing interval.

The Aminoglycosides and Fluoroquinolones exhibit a concentration dependent killing. For Fluoroquinolones, an AUC_{24}/MIC ratio of at least 125 is required to successfully treat respiratory tract infections caused by Gram-negative bacteria in terminally ill, elderly patients (Schuck et al. 2005). Based on in vitro, in vivo, and clinical data, it has been suggested that the PK/PD parameter AUC_{24}/MIC describes the activity of Vancomycin the best, therefore giving an exact idea about time related bacteriological outcome in patients with respiratory tract infections caused by *S. aureus*. An AUC_{24}/MIC value $\geq$ 400 is required to improve the outcome of patients with severe staphylococcal infections (Moise-Broder et al. 2004). However, this value is relatively high when compared with other antibiotics.

Pathophysiological Conditions Leading to Treatment Failures

In spite of the in depth analysis of PK/PD parameters for comprehensive assessment of drug efficacy and undue toxic effects, most of the treatment regimen do not result in the optimal outcome often and thus lead to treatment failures. This is a major concern especially in critically ill patients and must be carefully assessed and addressed to realize successful treatment of acute and chronic infectious diseases. There are often different reasons contributing to reduced or incomplete cure in critically ill patients (Bamberger 1997). The major reason is attributed to the fact that PK/PD parameters are mostly studied in infection models/animals. Laboratory based in vitro and animal data are further used to determine antimicrobial PD, which determines the initial dosing regimens to be used in clinical practice. However, it is observed that the PK/PD of antimicrobials in critically ill patients/patients with severe infections differ significantly from the patient groups from whose data the prospective dosing regimens had been developed. This often leads to entirely different/inadequate antimicrobial concentrations at the site of infection in terminally ill patients and thus to poor outcomes. Such data give a general regimen to be used in clinical scenario, but often fail in individual patients due to differences in immunity and other health parameters.

The vast array of pathophysiological changes occurring in patients suffering from infectious disease with varying degree of severity can complicate antibiotic dosing. Pharmacokinetics may be vastly altered in case of critical illness, hepatic dysfunction, sepsis, burns, pregnancy, cystic fibrosis, or other primary and secondary bacterial infections. In such case, dosing may fail and individualized therapy may be necessary. For example, in case of sepsis low plasma concentration of drugs can be a result of either an increase in CL or Vd. This may be a result of increased cardiac output or increased capillary permeability (McKinnon and Davis 2004; Udy et al. 2008). Myocardial depression leading to decrease in organ perfusion and ultimately microvascular circulation failure adds to inefficiency of antibiotic regimen. In other cases, sepsis can induce multiple organ dysfunction, including renal and/or hepatic dysfunction, leading to drastic decrease in antibacterial clearance (Roberts and Lipman 2009). In case of severe burns, serum concentrations of antibiotics may fall below the MICs of infecting pathogens, implicating an increased requirement of the drug or in some cases frequent dosing of the antibiotic (Dryden et al. 2011). Therefore, there is a potential for patients who are critically ill with sepsis or burns to have subinhibitory serum concentrations of Linezolid with standard dosing regimens. These altered PK factors in critically ill patients can have a significant effect attainment of adequate antibiotic doses. Moreover, considerable interpatient variability in drug absorption, distribution, metabolism, and elimination; protein binding; and tissue absorption can have a profound effect on the ability to achieve pharmacodynamic targets at conventional doses (McKinnon and Davis 2004; Vincent et al. 2016). Additionally, it is seen that the antimicrobial PK profile of small animal models can be extremely different to that in humans. The drug

retention and/or clearance may be substantially different. Therefore, careful dosing strategies must be designed with care.

Most importantly, bacteria embedded in biofilms pose a serious challenge for antimicrobial action and is the major reason for failure of dosage regimens in fatal infections. Similarly, most of the chronic infections are due to mixed bacterial species or super-infections which possess an even greater threat. Such diseases are often critical and have extremely low survival rates. Hospital/ventilator-associated pneumonia, sepsis, diabetes foot infection are few such cases which warrant attention. Under such situations, often optimal PD targets cannot be achieved using single antibiotic and combination therapy may be required (McKinnon and Davis 2004; Tamma et al. 2012; Yim et al. 2016).

Overcoming Treatment Failures

To overcome the issue of treatment failures, several factors affecting PK/PD characteristics should be evaluated to achieve the desired pharmacodynamic targets in antimicrobial selection. A mere consideration of MIC values may not be sufficient. Additional consideration of patient-specific pharmacokinetic variation must be made. Concentration–time data of antimicrobials must be studied in terminally ill patients when developing population PK models. These will give a realistic dosing regimen that will account for drastically altered drug concentrations in patients (McKinnon and Davis 2004; Tängdén et al. 2017). Success can be achieved by employing therapeutic drug monitoring (TDM). TDM involves measurement of drug concentrations and dose adjustment based on the observed concentration in relation to a target drug exposure. Antimicrobial TDM can not only minimize drug toxicity, but also maximize drug efficacy leading to an overall therapeutic effect. TDM thus can be used for infected critically ill patients where early dose adaptation to the needs of the individual patient can give better outcome. Furthermore, TDM can be of great importance for such patients where prompt appropriate antibiotic therapy is crucial. Additionally, nomograms can be designed where PK characteristics derived from more than one patient can be used in models to determine dosing regimens (Tängdén et al. 2017). In case of tuberculosis patients with a greater risk of treatment failure, TDM ensures appropriate serum drug concentrations and may assist in the clinical decision-making process (Nuermberger and Grosset 2004). A good outcome of an infection episode may be achieved by early institution of appropriate antimicrobial therapy. Depending on the nature of the etiological agent, either a specific antimicrobial agent or combination of agents can be effective in achieving targeted infection clearance especially in critically ill patients (Bush and Levison 1988). Depending on the type of infection, severity of illness, or possible immunosuppression optimized dosing regimens can be designed for different combination of antibiotics. When studying the combination of drugs and their time–kill curves in vitro PK/PD models, differences in their PK properties can be taken into account. Prediction of PD target attainment can be made based on this study and specifically designed dosing regimens can be effective over normal regimens. However, these should be should be validated in clinical studies before they can be used in general practice.

Combination therapy involving two or more antibiotics can be used to treat complicated fatal/chronic diseases by mixed infections of Gram-positive and Gram-negative bacteria. However, a forehand analysis of the additive, synergistic, or antagonistic effect of the drugs is essential before such regimens can be introduced. In case of TB treatment, combination therapy is in vogue. It benefits from the additive and/or synergistic effects of antimycobacterial agents. This is practiced as there is a high potential for the development of drug resistance using monotherapy. Models similar to the in vitro PK/PD models for simultaneous simulation of serum kinetics of two or more drugs with different half-lives developed against *Streptococcus pneumonia* and *Staphylococcus aureus* could be used for development of anti-TB agents (Nuermberger and Grosset 2004; Vaddady et al. 2010). Combination therapy using Piperacillin/Tazobactam or Imipenem/Cilastatin

can be used to treat mixed bacterial foot infections in diabetic patients (Saltoglu et al. 2010). It has also been found useful in the treatment of chronic biofilm infection by *Pseudomonas aeruginosa* in murine tumor model (Herrmann et al. 2010; Pawar et al. 2015). Clinical trials have been made to compare the efficacy of combination therapy for treatment of hospital-acquired pneumonia. Pharmacodynamic interactions between Telavancin and Aztreonam or Piperacillin/Tazobactam were evaluated against *P. aeruginosa, E. coli,* and *S. aureus* using an in vitro one compartment PK/PD model (Yim et al. 2016).

Conclusion

To counter the antimicrobial resistance and the persistence of bacteria in chronic infections, PK/PD properties should be considered from the point of selection of antibiotics. Additionally, it is of high importance to understand the nature of bacterial infection, the antibiotic resistance profile, and biofilm formation ability of the bacteria. An in depth analysis of PK/PD of antibiotics and use of right models and simulation data is essential for revamping existing treatment regimens. The same procedure should be considered while developing/evaluating new strategies against fatal infectious diseases. Incorporation of region or patient specific PK/PD values would be key factor determining successful outcome. Additionally, choice of particular model or antibiotic/s against infectious diseases should be governed by an extensive investigation of the causative agent/s, resistance mechanisms, underlying patho-physiological stage, and PK/PD parameters.

References and Further Reading

Andes D, Craig WA (1998) In vivo activities of amoxicillin and amoxicillin-clavulanate against Streptococcus pneumoniae: application to breakpoint determinations. Antimicrob Agents Chemother 42:2375–2379

Bamberger D (1997) Antibiotics: why they fail in patients who are critically ill. Crit Care Nurs Q 20:60–68

Bergen PJ, Bulitta JB, Forrest A, Tsuji BT, Li J, Nation RL (2010) Pharmacokinetic/pharmacodynamic investigation of colistin against pseudomonas aeruginosa using an in vitro model. Antimicrob Agents Chemother 54:3783–3789

Bonate PL (2001) A brief introduction to Monte Carlo simulation. Clin Pharmacokinetic 40:15–22

Bush LM, Levison ME (1988) Antibiotic selection and pharmacokinetics in the critically ill. Crit Care Clin 4:299–324

Canut A, Isla A, Betriu C, Gascón AR (2012) Pharmacokinetic – pharmacodynamic evaluation of daptomycin, tigecycline, and linezolid versus vancomycin for the treatment of MRSA infections in four western European countries. Eur J Clin Microbiol Infect Dis 31(9):2227–2235

Cao B, Christophersen L, Thomsen K, Sønderholm M, Bjarnsholt T, Jensen PØ, Høiby N, Moser C (2015) Antibiotic penetration and bacterial killing in a Pseudomonas aeruginosa biofilm model. J Antimicrob Chemother 70:2057–2063

Carlton KK Lee, Michael P Boyle, Marie Diener-West, Lois Brass-Ernst, Michelle Noschese, Pamela L Zeitlin (2007) Levofloxacin Pharmacokinetics in Adult Cystic Fibrosis. Chest 131 (3):796-802

Craig WA, Andes D (1996) Pharmacokinetics and pharmacodynamics of antibiotics in otitis media. Pediatr Infect Dis J 15:944–948

Craig, WA, Redington, J, Ebert, SC Pharmacodynamics of amikacin in vitro and in mouse thigh and lung infections. J. Antimicrob. Chemother. 27 (Suppl.) S29–S40 (1991).

Crandon JL, Nicolau DP (2011) Pharmacodynamic approaches to optimizing beta-lactam therapy. Crit Care Clin 27:77–93

Crull K, Weiss S (2011) Antibiotic control of tumor-colonizing Salmonella enterica serovar Typhimurium. Exp Biol Med 236:1282–1290

Czock D, Keller F (2007) Mechanism-based pharmacokinetic–pharmacodynamic modeling of antimicrobial drug effects. J Pharmacokinet Pharmacodyn 34:727–751

Dayneka NL, Garg V, Jusko WJ (1993) Comparison of four basic models of indirect pharmacodynamic responses. J Pharmacokinet Biopharm 21:457–478

Drago L, De Vecchi E (2008) The safety of cefepime in the treatment of infection. Expert Opin Drug Saf 7:377–387

Dryden M, Johnson AP, Ashiru-Oredope D, Sharland M (2011) Using antibiotics responsibly: right drug, right time, right dose, right duration. J Antimicrob Chemother 66:2441–2443

Dudhani RV, Turnidge JD, Coulthard K, Milne RW, Rayner CR, Li J, Nation RL (2010a) Elucidation of the pharmacokinetic/pharmacodynamic determinant of colistin activity against Pseudomonas aeruginosa in murine thigh and lung infection models. Antimicrob Agents Chemother 54:1117–1124

Dudhani RV, Turnidge JD, Nation RL, Li J (2010b) fAUC/MIC is the most predictive pharmacokinetic/pharmacodynamic index of colistin against Acinetobacter baumannii in murine thigh and lung infection models. J Antimicrob Chemother 65:1984–1990

Eagye KJ, Nicolau DP, Lockhart SR, Quinn JP, Doern GV, Gallagher G, Abramson MA (2007) A pharmacodynamic analysis of resistance trends in pathogens from patients with infection in intensive care units in the United States between 1993 and 2004. Ann Clin Microbiol Antimicrob 6:11

Felmlee MA, Morris ME, Mager DE (2012) Mechanism-based pharmacodynamic modeling BT. In: Reisfeld B, Mayeno AN (eds) Computational toxicology, vol I. Humana Press, Totowa, pp 583–600

Firsov AA, RG Vasilov, SN Vostrov, OV Kononenko, I Yu Lubenko, S H Zinner, (1999) Prediction of the antimicrobial effects of trovafloxacin and ciprofloxacin on staphylococci using an in-vitro dynamic model. Journal of Antimicrobial Chemotherapy 43 (4):483-490

Foerster S, Unemo M, Hathaway LJ, Low N, Althaus CL (2016) Time-kill curve analysis and pharmacodynamic modelling for in vitro evaluation of antimicrobials against Neisseria gonorrhoeae. BMC Microbiol 16(1):1–11

Gloede J, Scheerans C, Derendorf H, Kloft C (2010) In vitro pharmacodynamic models to determine the effect of antibacterial drugs. J Antimicrob Chemother 65:186–201

Hall Snyder AD, Vidaillac C, Rose W, McRoberts JP, Rybak MJ (2015) Evaluation of high-dose daptomycin versus vancomycin alone or combined with clarithromycin or rifampin against Staphylococcus aureus and S. epidermidis in a novel in vitro PK/PD model of bacterial biofilm. Infect Dis Ther 4:51–65

Herrmann G, Yang L, Wu H, Song Z, Wang H, Høiby N, Ulrich M, Molin S, Riethmüller J, Döring G (2010) Colistin-tobramycin combinations are superior to monotherapy concerning the killing of biofilm Pseudomonas aeruginosa. J Infect Dis 202:1585–1592

José Pérez-Urizar, Vinicio Granados-Soto, Francisco J Flores-Murrieta, Gilberto Castañeda-Hernández, (2000) Pharmacokinetic-Pharmacodynamic Modeling. Archives of Medical Research 31 (6):539-545

Lorenz A, Pawar V, Häussler S, Weiss S (2016) Insights into host–pathogen interactions from state-of-the-art animal models of respiratory Pseudomonas aeruginosa infections. FEBS Lett 590:3941–3959

Mager DE, Diversity of Mechanism-Based Pharmacodynamic Models. Drug Metabolism and Disposition 31 (5):510-518

McKinnon PS, Davis SL (2004) Pharmacokinetic and pharmacodynamic issues in the treatment of bacterial infectious diseases. Eur J Clin Microbiol Infect Dis 23:271–288

Meibohm B, Derendorf H (2002) Pharmacokinetic/pharmacodynamic studies in drug product development. J Pharm Sci 91:18–31

Michael J, Barth A, Kloft C, Derendorf H (2014) Pharmacodynamic in vitro models to determine the effect of antibiotics. In: Vinks AA, Derendorf H, Mouton JW (eds) Fundamentals of antimicrobial pharmacokinetics and pharmacodynamics. Springer New York, New York, pp 81–112

Moise-Broder PA, Forrest A, Birmingham MC, Schentag JJ (2004) Pharmacodynamics of vancomycin and other antimicrobials in patients with Staphylococcus aureus lower respiratory tract infections. Clin Pharmacokinet 43:925–942

Mouton JW (2014) Setting clinical MIC breakpoints from a PK/PD point of view: it is the dose that matters. In: Vinks AA, Derendorf H, Mouton JW (eds) Fundamentals of antimicrobial pharmacokinetics and pharmacodynamics. Springer New York, New York, pp 45–61

Mouton J, Dudley M, Cars O, Derendorf H, Drusano G (2005) Standardization of pharmacokinetic/pharmacodynamic (PK/PD) terminology for anti-infective drugs: an update. J Antimicrob Chemother 55:601–607

Mouton JW, Brown DFJ, Apfalter P, Cantón R, Giske CG, Ivanova M, MacGowan AP, Rodloff A, Soussy C-J, Steinbakk M, Kahlmeter G (2012) The role of pharmacokinetics/pharmacodynamics in setting clinical MIC breakpoints: the EUCAST approach. Clin Microbiol Infect 18:E37–E45

Nation RL, Bergen PJ, Li J (2014) Pharmacokinetics and pharmacodynamics of colistin BT. In: Vinks AA, Derendorf H, Mouton JW (eds) Fundamentals of antimicrobial pharmacokinetics and pharmacodynamics. Springer New York, New York, pp 351–380

Nicolau DP (2003) Optimizing outcomes with antimicrobial therapy through pharmacodynamic profiling. J Infect Chemother 9:292–296

Nielsen EI, Friberg LE (2013) Pharmacokinetic-pharmacodynamic modeling of antibacterial drugs. Pharmacol Rev 65:1053–1090

Nuermberger E, Grosset J (2004) Pharmacokinetic and pharmacodynamic issues in the treatment of mycobacterial infections. Eur J Clin Microbiol Infect Dis 23:243–255

Parra-Ruiz J, Vidaillac C, Rose WE, Rybak MJ (2010) Activities of high-dose daptomycin, vancomycin, and moxifloxacin alone or in combination with clarithromycin or rifampin in a novel in vitro model of Staphylococcus aureus biofilm. Antimicrob Agents Chemother 54:4329–4334

Pawar V, Crull K, Komor U, Kasnitz N, Frahm M, Kocijancic D, Westphal K, Leschner S, Wolf K, Loessner H, Rohde M, Häussler S, Weiss S (2014) Murine solid tumours as a novel model to study bacterial biofilm formation in vivo. J Intern Med 276:130–139

Pawar V, Komor U, Kasnitz N, Bielecki P, Pils MC, Gocht B, Moter A, Rohde M, Weiss S, Häussler S (2015) In vivo efficacy of antimicrobials against biofilm producing Pseudomonas aeruginosa. Antimicrob Agents Chemother 59:AAC.00194-15

Regoes RR, Wiuff C, Zappala RM, Garner KN, Baquero F, Levin BR (2004) Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens. Antimicrob Agents Chemothe 48:3670–3676

Roberts J, Lipman J (2009) Pharmacokinetic issues for antibiotics in the critically ill patient pharmacokinetic issues for antibiotics in the critically ill patient. Crit Care Med 37:840–851

Rybtke M, Hultqvist LD, Givskov M, Tolker-Nielsen T (2015) Pseudomonas aeruginosa biofilm infections: community structure, antimicrobial tolerance and immune response. J Mol Biol 427:3628–3645

Saltoglu N, Dalkiran A, Tetiker T, Bayram H, Tasova Y, Dalay C, Sert M (2010) Piperacillin/tazobactam versus imipenem/cilastatin for severe diabetic foot infections: a prospective, randomized clinical trial in a university hospital. Eur Soc Clin Infect Dis 16:1252–1257

Sanchez CJ, Mende K, Beckius ML, Akers KS, Romano DR, Wenke JC, Murray CK (2013) Biofilm formation by clinical isolates and the implications in chronic infections. BMC Infect Dis 13:47

Scaglione F, Paraboni L (2006) Influence of pharmacokinetics/pharmacodynamics of antibacterials in their dosing regimen selection. Expert Rev Anti-Infect Ther 4:479–490

Schuck EL, Dalhoff A, Stass H, Derendorf H (2005) Pharmacokinetic/pharmacodynamic (PK/PD) evaluation of a once-daily treatment using ciprofloxacin in an extended-release dosage form. Infection 33:22–28

Sharma A, Jusko WJ (1998) Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. Br J Clin Pharmacol 45:229–239

Tamma PD, Cosgrove SE, Maragakis LL (2012) Combination therapy for treatment of infections with gram-negative bacteria. Clin Microbiol Rev 25:450–470

Tängdén T, Martín VR, Felton TW, Nielsen EI, Marchand S, Brüggemann RJ, Bulitta JB (2017) The role of infection models and PK/PD modelling for optimising care of critically ill patients with severe infections. Intensive Care Med 43:1021–1032

Taraszkiewicz A, Fila G, Grinholc M, Nakonieczna J (2013) Innovative strategies to overcome biofilm resistance. Biomed Res Int 2013:1–13

Udy A, Roberts J, Boots R, Lipman J (2008) Dose adjustment and pharmacodynamic considerations for antibiotics in severe sepsis and septic shock BT. In: Rello J, Restrepo MI (eds) Sepsis: new strategies for management. Springer Berlin Heidelberg, Berlin/Heidelberg, pp 97–136

Vaddady PK, Lee RE, Meibohm B (2010) In vitro pharmacokinetic/pharmacodynamic models in anti-infective drug development: focus on TB. Future Med Chem 2:1355–1369

Vincent J, Bassetti M, François B, Karam G, Chastre J, Torres A, Roberts JA, Taccone FS, Rello J, Calandra T, De Backer D, Welte T (2016) Advances in antibiotic therapy in the critically ill. Crit Care 20:1–13

Vishnu Dutt Sharma, Aman Singla, Manu Chaudhary, Manish Taneja, (2016) Population Pharmacokinetics of Fixed Dose Combination of Ceftriaxone and Sulbactam in Healthy and Infected Subjects. AAPS PharmSciTech 17 (5):1192-1203

Wayne (2014) Performance standards for antimicrobial susceptibility testing; Twenty-fourth informational supplement, CLSI document M100–S24. In: Clinical and laboratory standards institute report. Clinical and Laboratory Standards Institute, Wayne (PA), pp 1–230

Wilson SE, Graham DR, Wang W (2017) Telavancin in the treatment of concurrent Staphylococcus aureus bacteremia: a retrospective analysis of ATLAS and ATTAIN studies. Infect Dis Ther 6(3):413–422

Yim J, Smith JR, Barber KE, Hallesy JA, Rybak MJ (2016) Evaluation of pharmacodynamic interactions between telavancin and aztreonam or piperacillin/tazobactam against Pseudomonas aeruginosa, Escherichia coli and methicillin-resistant Staphylococcus aureus. Infect Dis Ther 5:367–377

Zervos M, Nelson M (1998) Cefepime versus ceftriaxone for empiric treatment of hospitalized patients with community-acquired pneumonia. The Cefepime study group. Antimicrob Agents Chemother 42: 729–733

Zhanel GG (2001) Influence of pharmacokinetic and pharmacodynamic principles on antibiotic selection. Curr Infect Dis Rep 3:29–34

Role of Clinical Pharmacodynamics Studies in the Era of Precision Medicines Against Cancer

16

Fatih M. Uckun and Sanjive Qazi

Contents

Abstract

Pharmacodynamics (PD) has been integral to the design of rational drug dosing regimens. Detailed PD studies during both the preclinical and clinical stages of the drug development process can also contribute to lead optimization or the selection of the optimal "best-in-class" compound, improve clinical potency estimates, and help predict the drug exposure needed to achieve meaningful clinical responses. There has been a substantial and continued increase in the number of clinical oncology trials with integrated PD studies since 2002. Notably, a significant portion of all interventional clinical trials with PD components are initiated for evaluation of oncology drugs. PD studies frequently play a pivotal role in determining the initial dose level for first-in-human clinical studies of immunooncology drugs. The integration of PD data into the dose safety modeling in early oncology studies

F. M. Uckun (✉)
AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA
e-mail: fatih.uckun@aresmit.com; fatih.uckun@att.net

S. Qazi
AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA

Bioinformatics Program, Gustavus Adolphus College, St. Peter, MN, USA
e-mail: sqazi@gustavus.edu

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_37

may provide accurate predictions of the dose-effect relationships by advancing the understanding of target engagement as well as exposure response and therefore has the potential to improve the decision making regarding the optimal dose and schedules as well as risk-benefit assessments for later stages in clinical development. PD studies also have the potential to provide early clinical proof of concept when drugs with complementary activity profiles are combined in cancer therapy. In recent years, population pharmacokinetics–pharmacodynamics (PK-PD) modeling has become a key tool towards streamlining and optimizing oncologic drug development through early understanding, identification, and quantification of various dose–response relationships in the context of other patient characteristics as well as risk-benefit of different dosing schedules. Finally, a new and exciting strategy known as "Quantitative Systems Pharmacology" is emerging that advances systems level, multiscale models for disease progression and treatment to better characterize the hierarchical, non-linear, dynamic responses at the network level of drug action that may affect both efficacy and toxicity in clinical settings.

Introduction

An improved understanding of the molecular pathology of cancers combined with the development of targeted therapeutics and immuno-oncology drugs that activate the host immunity against cancer cells (Turan et al. 2018; Socinski et al. 2018; Valla et al. 2018) has caused a paradigm-shift in drug development and clinical trial methods by providing the foundation for personalized medicine strategies with specific roadmaps for the patient-tailored rational deployment of specific drugs in biomarker-enriched patient populations (Biankin et al. 2015; Tsimberidou et al. 2017; Phelan et al. 2018; DiNardo et al. 2018; Alsharedi et al. 2018; Drilon et al. 2018; Hidalgo et al. 2018; Mutti et al. 2018; Lu et al. 2016; Torres-Ayuso et al. 2018; Palmirotta et al. 2018; Peck 2015; Jamal et al. 2017; Krebs et al. 2016; Laetsch et al. 2017).

Pharmacodynamics (PD) has been integral to the design of rational drug dosing regimens (FDA 2016). PD is the study of the relationship between drug concentration and its effects at the subcellular, cellular, tissue, organ system, or whole-body level, including all of the pharmacological actions, pathophysiological effects, and therapeutic activities, and adverse side effects of the active drug ingredient, therapeutic moiety, and/or its metabolite(s) (de Man et al. 2018; Derendorf et al. 2000; de Vries et al. 2018). While some PD studies require tissue biopsies, others use surrogate tissues such as blood cells or noninvasive methods such as anatomic or functional imaging. Some drugs result in activation of gene expression which can be leveraged in PD studies. For example, omaveloxolone is a semisynthetic oleanane triterpenoid that potently activates Nrf2 with subsequent antioxidant function. In a recently reported Phase 1 study (NCT02029729), downstream Nrf2 activation was assessed in peripheral blood mononuclear cells by quantification of target gene mRNA expression (Creelan et al. 2017). An increase in select Nrf2 target gene expression was observed during the course of treatment, across multiple dose levels. Mutant IDH1 produces high levels of 2-hydroxyglurate (2HG), thought to initiate oncogenesis through epigenetic modifications of gene expression. Inhibitors of the mutant isocitrate dehydrogenase 1 (IDH1) are being evaluated in patients with brain tumors. Recently, Andronesi et al. described an elegant noninvasive 3D MR spectroscopic neuroimaging method for rapid and easy detection of 2HG to study the PD of IDH305, an orally available, brain penetrant, mutant-selective allosteric high affinity IDH1 inhibitor that acts on both canonical (R132H) and noncanonical (R132C) mutated enzymes (Andronesi et al. 2018). The authors demonstrated the feasibility of image-based 2HG PD serial assessments and demonstrated that the IDH305 treatments of glioma patients during the NCT02381886 Phase 1 clinical study caused a rapid decline of 2HG levels by 70% as expected from an inhibitor of mutant IDH1.

The purpose of this chapter is to discuss the role of PD studies in the drug development process with a focus on the integration of PD studies in contemporary clinical trials of oncology drugs.

Role of PD Studies in Translational Oncology

The drug development strategies should encompass both the nonclinical and clinical stages of the life cycle of a promising new drug candidate. The quality of the nonclinical development, including the identification of robust biomarkers, noninvasive PD assays, well-defined relationships between the PD parameters and pharmacokinetics (PK) parameters, development of laboratory tests for predictive biomarkers, and PD analyses amenable to validation, has a direct and differentiating impact on the success of the early phase clinical development (Brennan et al. 2018). Most drug makers seek earlier decision making about go or no-go plans on the basis of PK and PD characteristics of their promising drug candidates. Detailed PK-PD studies during both the preclinical and clinical stages of the drug development process can also contribute to lead optimization or the selection of the optimal "best-in-class" compound, improve clinical potency estimates and help predict the drug exposure needed to achieve meaningful clinical responses. PD studies frequently play a pivotal role in determining the initial dose level for first-in-human clinical studies of immunostimulatory drugs according to the minimal anticipated biologic effect level (MABEL) approach by integrating all of the available in vitro and in vivo information by PK/PD modeling.

It is also important to note that the insights and lessons learned from nonclinical PD studies often provide the foundation for highly promising combined modality regimens. For example, nonclinical PD studies demonstrated that the FDA-approved 2nd-line anti-chronic lymphocytic leukemia (CLL) drug Venetoclax targeting the anti-apoptotic protein BCL2 potentiated/complemented the activities of and sometimes synergized with the Bruton's tyrosine kinase (BTK) inhibitors Acalabrutinib and Ibrutinib. Tam et al. recently reported the results of a clinical study (NCT02471391) which demonstrated that dual targeting of BTK and BCL2 with Ibrutinib plus Venetoclax as part of an innovative treatment regimen results in significantly improved response rates and treatment outcomes in patients with mantle-cell lymphoma (Tam et al. 2018). The complete response rate at week 16 was 42%, which was markedly higher than the historical result of 9% at this time point with Ibrutinib monotherapy (Tam et al. 2018).

According to the clinicaltrials.gov data repository, a total of 1,930 interventional clinical oncology trials with integrated PD studies were initiated between August 1994 and July 2018. Of these, only 33 (1.7%) were started between August 1994 and July 2002. There has been a substantial and continued increase in the number of clinical oncology trials with integrated PD studies since 2002 (Fig. 1): 224 trials (8.6-fold increase from previous 4 years) were initiated between August 2002 and July 2006, 474 trials (2.1-fold increase from previous 4-years) between August 2006 and July 2010, 564 trials (~19% increase from previous 4 years) between August 2010 and July 2014, and 635 (~13% increase from previous 4 years) between August 2014 and July 2018. Notably, ~62% of the clinical PD studies in oncology were initiated within the last 8 years. Hence, PD studies are playing an increasingly important role in the clinical development path of oncology drugs. Notably, a significant portion of all interventional clinical trials with PD components are initiated for evaluation of oncology drugs. Whereas 635 of the 2,076 clinical PD studies (30.6%) that were initiated between August 2014 and July 2018 were in patients with cancer, only 178 studies (8.6%) were in patients with neurological disorders (Neuro), 177 studies (8.5%) in patients with cardiovascular diseases (CVD), 136 studies (6.6%) in patients with pulmonary disease (PD), 54 studies (2.6%) in patients with allergic disorders (AD), and 129 studies (6.2%) in patients with autoimmune disorders (AI) (Fig. 2). There were more studies in cancer patients (viz.: 635 studies) than in patients with CVD, PD, AD,

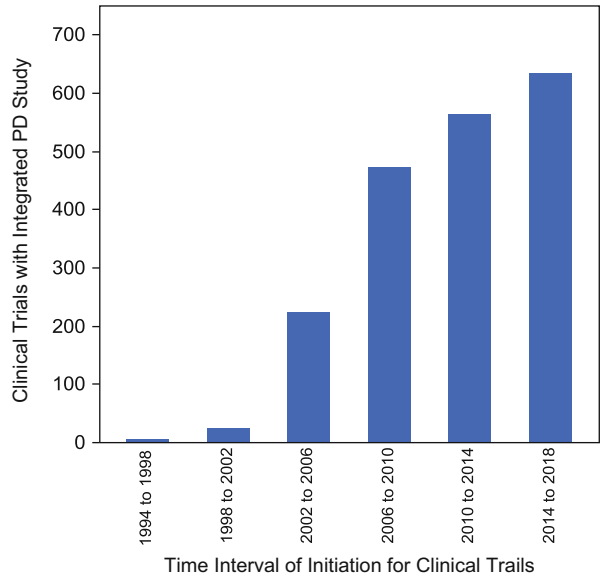


Fig. 1 Clinical trials with integrated PD studies in oncology. We interrogated the clinicaltrials.gov data repository (https://clinicaltrials.gov/) to determine the number of interventional trials that employed pharmacodynamic methods to characterize anti-cancer therapies from 1994 to 2018 in 4 year increments. All Interventional trials that were started over the 4-year period were included in the totals. There was a total of 1,930 trials counted from 1994 to 2018. Search terms to identify the trials were "Pharmacodynamic," "Interventional studies," and "Cancer"

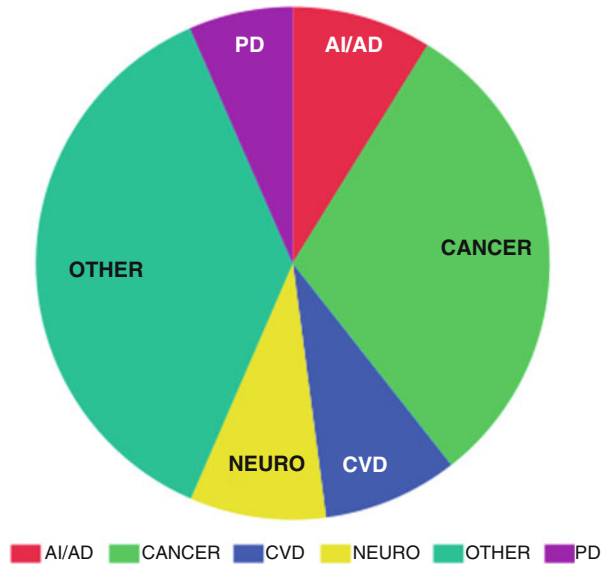


Fig. 2 Patient populations of clinical PD studies initiated between 2014 and 2018. We interrogated the clinicaltrials.gov data repository (https://clinicaltrials.gov/) to determine the number of interventional trials with integrated PD studies that were initiated between August 2014 and July 2018. All Interventional trials that were started over the 4-year period were included in the totals. Search terms to identify the trials were "Pharmacodynamic," "Interventional studies," and either "Cancer," "Cardiovascular Diseases (CVD)," "Pulmonary Disease (PD)," "Autoimmune Diseases (AI)," "Allergic disorders (AD)," or "Neurological Disorder (Neuro)" to stratify the disease types

and AD combined (viz.: 545 studies). Table 1 depicts a select list of actively recruiting clinical oncology trials with integrated PD components and estimated primary completion dates ranging from 01/01/2017 to 07/31/2018.

Role of PD Studies in Early Oncology Trials

The primary objective of Phase I oncology trials is to determine the optimal dose of an agent or combination of agents that can be used as the recommended phase 2 dose (RP2D) (Cook et al. 2015; Caimi et al. 2017). The RP2D levels of anticancer agents are traditionally determined by dose-limiting toxicities (DLT) and correspond to the maximum tolerated dose (MTD) , which is the highest clinically-safe dose that is derived from DLT data obtained most commonly during the first few treatment cycles. Identification of the MTD is still the most commonly used method to identify the RP2D for oncology drugs. There is a need to reconsider the assessment of MTD for some medicinal products as a need for dose reduction is discovered in a high percentage of patients in Phase III trials, despite the absence of dose-limiting toxicity (DLT) conventionally defined by Grade 3 and 4 events (Lavezzi et al. 2018). A recent workshop demonstrated that many FDA-approved anticancer drugs with molecular targets were subject to dose reductions in late-stage registration trials to improve their tolerability (Jänne et al. 2016). As Phase I trials of molecularly targeted agents often do not use toxicity data beyond the first two cycles of treatment to determine the RP2D, it has been suggested that longitudinal relative dose intensity evaluations may be warranted to obtain more robust RP2D levels (Hirakawa et al. 2018).

The comparisons of the on-target PD profiles of targeted therapeutics may help identify the best in class compounds and contribute to the proof of concept obtained early oncology trials involving in biomarker-enriched patient populations. For example, Acalabrutinib and Ibrutinib exhibited comparable on-target PD in regard to changes in CCL3/CCL4 chemokine production, migration assays, and changes in B-cell receptor signaling pathway proteins, and both are associated with high overall response rates and durable remissions in previously treated CLL (Patel et al. 2018).

There is a growing consensus and enthusiasm among oncologists that a more comprehensive evaluation of the drug activity profile including its PD and PK features should be used to identify the RP2D levels. It has been reported that only one third of the studies used toxicity endpoints alone to determine the RP2D (Hansen et al. 2017). That is in part because new-generation targeted anticancer agents exhibit clinically meaningful activity at levels 25% of the MTD (Jain et al. 2010). The potential advantages of including multiple nontoxicity endpoints such as PD, PK, and efficacy with or without toxicity to define RP2D as an alternative to toxicity alone include the identification of better tolerable effective dose levels. For example, using an accelerated titration, 3 + 3 dose-escalation, open-label Phase I trial (NCT01940133) of continuous once-daily dosing (OD), Wicki et al. evaluated the safety, pharmacokinetics (PK), and pharmacodynamics of PQR309 in patients with advanced solid tumors. PQR309 is an orally bioavailable, balanced pan-phosphatidylinositol-3-kinase (PI3K), mammalian target of rapamycin (mTOR) C1, and mTORC2 inhibitor (Wicki et al. 2018). The MTD and RP2D of PQR309 was 80 mg of orally OD. PK was dose-proportional and PD showed PI3K pathway phosphoprotein downregulation in paired tumor biopsies (Wicki et al. 2018).

Notably, "nonclassically" defined RP2Ds were associated with a statistically significant five-fold higher rate of FDA drug approval for individual anticancer drugs (Hansen et al. 2017). The commonly used Phase I/II designs with an expansion phase after determination of the MTD during a dose escalation phase allow for early evaluation of clinical activity across multiple MTD-based vs. PD-based RP2D levels. Adaptive trial designs with randomized evaluation of multiple RP2Ds provide the opportunity to select the "best" RP2D.

There is general consensus among stakeholders that the first-in-human Phase I studies should be designed with focus on pharmacometrics tools and PK/PD-based nonsafety

Table 1 Select list of currently recruiting interventional clinical studies with integrated PD components

NCT study #	Official title	Sponsor	PD/PK study	Patient population
NCT02448589	A Phase I, open-label, nonrandomized, dose-escalating safety, tolerability, pharmacokinetic and pharmacodynamic study of TAS-119 in patients with advanced solid tumors (https://ClinicalTrials.gov/show/NCT02448589)	Taiho Oncology	PD/PK of TAS-119 (Aurora A KI)	Advanced solid tumors
NCT03008018	An open-label ascending dose study evaluating the safety/tolerability, pharmacokinetic and pharmacodynamic effects of KA2507 in patients with solid tumors (https://ClinicalTrials.gov/show/NCT03008018)	Kaxus Therapeutics, Ltd	PD/PK of KA2507 (HDACi)	Advanced solid tumors
NCT03450109	A Randomized and Open-label Study to Assess Pharmacokinetics, Pharmacodynamics and safety of LY01005 versus goserelin comparator (ZOLADEX®) following a single administration in patients with prostate cancer (https://ClinicalTrials.gov/show/NCT03450109)	Luye Pharma Group, Ltd	PD/PK of LY01005 (Goserelin acetate microspheres)	Prostate cancer
NCT02303028	A Phase I and enrichment study of low-dose metronomic Topotecan and Pazopanib in pediatric patients with recurrent or refractory solid tumors including CNS tumors (https://ClinicalTrials.gov/show/NCT02303028)	The Hospital for Sick Children	PD/PK of Pazopanib (TKI)	Pediatric R/R solid tumors and CNS tumors
NCT02619162	Nintedanib plus Letrozole in postmenopausal women with breast cancer: clinical trial phase 0/1 safety and pharmacodynamics (https://ClinicalTrials.gov/show/NCT02619162)	Centro Nacional de Investigaciones Carlos III	PD/PK of Nintedanib (TKI)	Breast cancer
NCT02503709	A Phase 1 trial of the combination of the heat shock protein-90 (HSP90) inhibitor Onalespib (AT13387) and the cyclin-dependent kinase (CDK) inhibitor AT7519M in patients with advanced solid tumors (https://ClinicalTrials.gov/show/NCT02503709)	NCI	PD/PK of Onalespib (Hsp90i) and CDKI AT7519 (CDK1,2,4,6,9 i)	Advanced solid tumors
NCT02679196	An open-label ascending dose study evaluating the safety/tolerability, pharmacokinetic and pharmacodynamic effects of KA2237 In patients with B Cell lymphoma (https://ClinicalTrials.gov/show/NCT02679196)	Karus Therapeutics	PD/PK of KA2237 (PI3Ki)	B-cell lymphoma
NCT02350868	A Phase 1, first-in-human, dose-seeking study evaluating the safety, pharmacokinetics, and pharmacodynamics of orally administered MPT0E028 in subjects with advanced solid malignancies without standard treatment (https://ClinicalTrials.gov/show/NCT02350868)	Taipei Medical University	PD/PK of MPT0E028 (HDACi)	Advanced solid tumors

(continued)

Table 1 (continued)

NCT study #	Official title	Sponsor	PD/PK study	Patient population
NCT01977638	Phase 1 study to assess safety, tolerability, pharmacokinetics, and pharmacodynamics of CXD101 given orally (twice-daily dosing for 5 consecutive days in a 21-day period) in patients with advanced malignancies expressing the biomarker HR23B (https://ClinicalTrials.gov/show/NCT01977638)	Oxford University Hospitals	PD/PK of CXD101 (HDACi)	Advanced solid tumors, lymphoma, MM
NCT02514239	An open-label, Phase I, dose escalation study to characterize the safety, tolerability, pharmacokinetics, and pharmacodynamics of intravenous doses of BI 836909 in relapsed and/or refractory multiple myeloma patients (https://ClinicalTrials.gov/show/NCT02514239)	Boehringer Ingelheim	PD/PK of BI-836909 (anti-BCMAxCD3 BiTE)	R/R MM
NCT02605746	A Phase 0/II study of Ceritinib (LDK378) in preoperative glioblastoma multiforme (GBM) and CNS metastasis patients scheduled for resection to evaluate central nervous system (CNS) penetration (https://ClinicalTrials.gov/show/NCT02605746)	St Joseph Hospital Med Center	PD/PK of Ceritinib (ALKi)	GBM/CNS mets
NCT02510001	A sequential Phase I study of MEK1/2 inhibitors PD-0325901 or Binimetinib combined With cMET inhibitor PF-02341066 in patients with RAS mutant and RAS wild type (with aberrant c-MET) colorectal cancer (https://ClinicalTrials.gov/show/NCT02510001)	University Oxford	PD/PK of Binimetinib (MEKi) and Crizotinib (ALKi/ROS-1i)	Colorectal cancer
NCT02501902	An open-label Phase Ib study of Palbociclib (Oral Cdk 4/6 Inhibitor) plus Abraxane (registered) (nab-paclitaxel) in patients with metastatic pancreatic ductal adenocarcinoma (https://ClinicalTrials.gov/show/NCT02501902)	Pfizer	PD/PK of Palbociclib (CDK4/6 i) and Abraxane	Metastatic pancreatic ductal carcinoma
NCT02940132	Phase 1 study to assess the safety, tolerability, pharmacokinetics/pharmacodynamics and preliminary efficacy of SC10914 in patients with advanced solid tumors (https://ClinicalTrials.gov/show/NCT02940132)	Jiangxi Qingfeng Pharmaceutical Co Ltd	PD/PK of SC10914 (PARPi)	Advanced solid tumors

endpoints to establish a more rationale dose finding paradigm in oncology drug development. The integration of PK/PD data into the dose safety modeling in early oncology studies may provide accurate predictions of the dose–effect relationships by advancing the understanding of target engagement as well as exposure response and therefore has the potential to improve the decision making regarding the optimal dose and schedules as well as risk-benefit assessments for later stages in clinical development (Grisafi et al. 2018). They may also help optimize the benefit–risk profile of oncology drugs through dose adaptation strategies for individualized dosing.

Role of PD Studies in Defining Optimized Treatments Regimens with a New Therapeutic Intervention

PD studies combined with PK and safety evaluations can provide actionable information regarding the alternative dose and schedule to realize the full clinical potential of a new therapeutic intervention. It was discovered that higher systemic exposures of the histone deacetylase inhibitor vorinostat were required than achieved in pediatric Phase I trials with continuous daily dosing for in vivo increased histone acetylation and cytotoxic activity. Consistent histone acetylation in peripheral blood mononuclear cells (PBMC) was only seen at the highest continuous dose level of vorinostat (300 mg/m^2/dose), a dose determined to be too toxic in combination with isotretinoin in a study performed by the Children's Oncology Group (COG) (NCT00217412) (Fouladi et al. 2010). At the continuous vorinostat MTD (230 mg/m^2/dose), only transient histone acetylation was observed. Consequently, Pinto et al. conducted a Phase I trial in children with relapsed/refractory neuroblastoma to determine the MTD of vorinostat on an interrupted schedule, escalating beyond the previously identified pediatric MTD. The maximum intended dose of vorinostat (430 mg/m^2/day) was tolerable when it was combined with isotretinoin. This dose led to increased vorinostat exposures and increased histone acetylation in surrogate tissues (viz., PBMC) when compared to lower doses of vorinostat (Pinto et al. 2018). Overall, the percent change from baseline in histone acetylation levels at 1 h post treatment was significantly greater in dose level 5 compared to dose levels 1–4 and this difference persisted for 24 h.

The metabolism of drugs and therefore the pharmacogenomics (PG) has a substantial impact on systemic exposure levels of the parent compound as well as its metabolites and the risk/severity of toxicities associated with them. The histone deacetylase inhibitors such as abexinostat, panobinostat, romidepsin, and vorinostat are eliminated through glucuronidation by UGT1A1. PD studies combined with PK and PG have demonstrated that polymorphisms (e.g., UGT1A1*28 and UGT1A1*60) that reduce UGT1A1 function cause increased systemic exposure, increased global protein lysine acetylation, and toxicities (e.g., thrombocytopenia) (Goey et al. 2016). Multiparameter modeling combining a population pharmacokinetic (PPK) model and a PD model describing the change in platelet levels in patients with cancer administered belinostat as a 48-h continuous intravenous infusion, along with cisplatin and etoposide, has been employed to optimize the treatment schedule and revealed that a q3week schedule of belinostat allows for sufficient platelet recovery before the next belinostat infusion is optimal (Peer et al. 2018).

Many targeted therapeutics, especially tyrosine kinase inhibitors (TKI), inhibit multiple kinases even if they have been labeled as highly selective inhibitors of one particular tyrosine kinase (Uckun et al. 2002, 2007, 2010; Uckun and Qazi 2010). For example, the BTK inhibitor Ibrutinib has been shown to inhibit other tyrosine kinases, including SRC, LYN, FYN, HCK, LCK, YES1, and FGR at nanomolar concentrations (Uckun and Qazi 2010; Honigberg et al. 2010). Therefore, PD evaluations of such compounds should not be limited to a single target kinase occupancy or inhibition in order to better understand its on-target and off-target effects and design appropriate and data-driven risk mitigation strategies. Ilorasertib (ABT-348) inhibits Aurora and VEGF receptor (VEGFR) kinases. In patients with advanced solid tumors, PD studies indicated that ilorasertib treatment engages both of these intended targets, but with maximum inhibition of VEGFR family kinases occur at lower exposures than typically required for inhibition of Aurora B in tissue. In agreement with the PD data, the DLTs in the NCT01110486 clinical trial were predominantly related to VEGFR inhibition (Maitland et al. 2018).

The combined PK and PD evaluations help determine clinical strategies for effective treatment of target patient populations (Stein et al. 2018; Tan et al. 2018; Tham et al. 2008). For example, for the BTK inhibitor Acalabrutinib that has an elimination half-life of 1 h., a twice daily (BID) dosing is used because it has been

shown to maintain plasma concentrations that are associated with >95% target BTK occupancy over the treatment interval and inhibition of BTK phosphorylation and activity in peripheral blood circulating CLL cells (Byrd et al. 2016). Likewise, recent analyses of the quantitative relationship between duration of severe neutropenia (the efficacy endpoint) and area under effect curve of absolute neutrophil counts, the PD endpoint, based on data from filgrastim products, a human granulocyte colony-stimulating factor (G-CSF), have provided useful information regarding the relationship between ANC and duration of severe neutropenia that can be used for dose selection and optimization of clinical trial design for G-CSF (Li et al. 2018).

Pegfilgrastim is a long-acting G-CSF indicated for prevention of febrile neutropenia in patients receiving myelosuppressive chemotherapy by promoting neutrophil recovery. In a Phase I, randomized, double-blind, three-way crossover trial in healthy volunteers, Waller et al. evaluated the PK, PD, safety, and tolerability of the proposed biosimilar, comparing MYL-1401H, reference pegfilgrastim (Neulasta®, Amgen Inc., Thousand Oaks, CA, USA) sourced from the European Union, and reference pegfilgrastim sourced from the USA. The primary PK and PD end points were similar across all groups. MYL-1401H demonstrated similar PK, PD, and safety to reference pegfilgrastim in healthy volunteers and may be an equivalent option for the prevention of febrile neutropenia (Waller et al. 2018). Likewise, combined PK/PD studies are often critical in determining if different administration routes of the same compound are equally effective and safe. For example, PK and PD (viz.: 20S proteasome inhibition) parameters of the proteasome inhibitor bortezomib following subcutaneous versus intravenous administration were very similar and this information together with the similar efficacy of subcutaneous versus intravenous bortezomib supports the approved routes of administration for bortezomib (i.e., intravenous and subcutaneous injection) (Moreau et al. 2012).

That being said, there are multiple challenges in incorporating PK and PD endpoints, including but not limited to increased labor, resource utilization, and trial complexity; increased burden of multiple blood draws, tumor biopsies, and imaging for the patient populations with advanced cancer; and commonly the absence of validated robust assays that can be used to obtain reliable PD endpoints in clinical settings.

Role of PD for Identification and Development of Combined Treatment Modalities

PD studies have the potential to provide early clinical proof of concept when drugs with complementary activity profiles are combined in cancer therapy (Rocchetti et al. 2009). The addition of the base excision repair inhibitor methoxyamine to fludarabine increases DNA double-strand breaks (Bulgar et al. 2010). Caimi et al. determined the safety, PK, PD, and RP2D of the base excision repair blocker methoxyamine combined with fludarabine in adult patients with relapsed/refractory hematologic malignancies (Caimi et al. 2017). They reported that this drug combination resulted in increased DNA damage measured with the Comet assay, as documented by cumulative increases in comet tail length throughout the first week of the combined methoxyamine + fludarabine therapy, indicating progressive DNA damage. The highly significant correlation between decreases in circulating malignant lymphocytes and comet tail length highlighted the relevance of DNA double-strand break measurements as a surrogate PD marker of the antineoplastic effect of methoxyamine and fludarabine. Notably, methoxyamine combined with fludarabine was safe and well tolerated. Hematologic toxicity was comparable to single agent fludarabine. The PD studies therefore demonstrated the potential of this combination as part of conditioning regimens of stem cell transplant and use of methoxyamine as fludarabine dose-sparing agent.

Sometimes, combined use of multiple anticancer drugs results in excellent treatment outcomes, and the question arises if the therapeutic benefits could be further improved by reducing toxicities with fewer cycles of therapy. Functional/

metabolic imaging using PET scans has been frequently applied as a PD measure of the activity of the multiagent regimen. For example, the intensive polychemotherapy regimen eBEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone in escalated doses) is highly active in patients with advanced-stage Hodgkin's lymphoma, but it is also associated with toxicities. Borchmann et al. investigated in a randomized Phase 3 study whether metabolic tumor response as determined by PET after two cycles of standard regimen eBEACOPP would allow response-directed adjustments of treatment intensity, increasing it for PET-positive patients or reducing it for PET-2-negative patients (NCT00515554) (Borchmann et al. 2018). PET negativity after two cycles allowed reduction to only four cycles of eBEACOPP without loss of tumor control. PET-guided eBEACOPP provided outstanding efficacy for all patients and increased the overall survival by reducing treatment-related risks for patients with excellent PET responses. PET-guided personalized reduced intensity treatment strategies should be considered in patients undergoing treatment with highly active regimens for metabolically active tumors such as non-Hodgkin's lymphoma, Hodgkin's lymphoma, NSCLC, breast cancer.

Based on clinical safety and activity data in the NCT01063816 trial, poly-ADP-ribose polymerase (PARP) inhibitors like Veliparib (Niu et al. 2017) or Rucaparib will likely be used in combination with standard chemotherapy drugs especially for treatment of ovarian cancer (Gray et al. 2018). It will be important to obtain detailed PK and PD data using parent-metabolite PK modeling and PD of these targeted therapeutics to gain insights into drug–drug interactions in order to optimize the administration schedules for the various components of the treatment regimens. It is also important to evaluate the impact of food on the PK and PD of these drugs. For example, Rucaparib can be taken with or without food but has different PK parameters when taken with food (versus fasting) probably due to solubility in the small intestine (Dal Molin et al. 2018).

Tegafur/gimeracil/oteracil (S-1) and irinotecan combination is attractive for breast cancer refractory to anthracyclines and taxanes. A reduction in circulating endothelial cell progenitors (CEPs) used to monitor the PD of S-1 is strongly correlated with antiangiogenic effects. Because vascular endothelial growth factor-A-driven tumor angiogenesis for the formation of a functional vascular bed and the subsequent tumor growth partly depend on the mobilization of CEPs, a change in the CEP level may be a predictive marker for antiangiogenesis therapy. The $CD34^+$ circulating endothelial cell (CEC) level was closely associated with the treatment response to chemotherapy, including S-1. Pharmacokinetics and reductions of $CD34^+$ CECs as pharmacodynamics were also analyzed. There was an association between clinical benefit and reduction in baseline $CD34^+$ CECs (4,6-diamino-2-phenylindole $(DAPI)^+$, $CD45^-$, $CD146^+$, or $CD105^+$ and $CD34^+$) by S-1. These results provided the foundation for combined use of irinotecan and S-1 in advanced GI malignancies (Ishiguro et al. 2017).

PD studies in early oncology trials also provide the first clinical proof of concept for further development of a new clinical strategy for difficult-to-treat cancers. Pelareorep, an oncolytic virus and an isolate of reovirus Type 3 Dearing showed single-agent antitumor activity. A recent PD study in patients with advanced pancreatic cancer demonstrated reovirus replication within pancreatic tumor and associated apoptosis, thereby providing the first proof of concept that the high frequency of RAS mutations in pancreas cancer would promote selective reovirus replication in pancreatic tumors and enhance the anticancer activity of gemcitabine (Mahalingam et al. 2018).

Likewise, it will be very important to emphasize PK/PD analyses in clinical trials involving older adults with cancer, especially those over age 75. The careful analysis of both chronological and functional age and comorbidities on PK/PD of new drugs in relationship to the safety- as well as efficacy-related clinical outcome parameters will help identify subsets of older adults who are likely to benefit from specific therapeutic interventions

as well as those who are most vulnerable to morbidity and/or mortality (Nightingale et al. 2018).

Pharmacodynamics Modeling for Development of Oncology Drugs

In recent years, population pharmacokinetics–pharmacodynamics (PK-PD) modelling has become a key tool toward streamlining and optimizing oncologic drug development through early understanding, identification, and quantification of various dose–response relationships in the context of other patient characteristics as well as risk-benefit of different dosing schedules (Garralda et al. 2017; Nightingale et al. 2018; Owonikoko et al. 2018; Sato et al. 2017). The development of nonclinical models that can predict the clinical toxicities of immuno-oncology drugs, including immune checkpoint inhibitors and stimulators, is a focal point of emphasis in contemporary translational cancer research and regulatory science/policy workshops (e.g., FDA-AACR Workshop on nonclinical Models for Safety Assessment of Immuno-oncology Products. September 6th, 2018 Marriott Wardman Park, Washington, DC).

PK/PD modeling is a useful tool throughout all stages of drug development, and applications differ during the preclinical and clinical stage. Modeling strategies can accelerate the clinical development process by (i) providing the foundation for an early analysis of the safety and tolerability profile of drug candidates, (ii) early definition of the risk-benefit ratio and the therapeutic index and (ii) supporting the design of optimal treatment regimens (Meille et al. 2017; Zamboni et al. 2001; Zhou and Gallo 2011). Modeling has been extensively used in anticancer drug development to individualize dosing strategies based on patient characteristics, and design optimal and sometimes personalized dosing regimens (Ait-Oudhia and Mager 2016; Block 2015; Buil-Bruna et al. 2016; Ciccolini et al. 2017; Claret et al. 2009, 2015).

Early understanding of toxicities and PK determination of the oral pan-histone deacetylase inhibitor Abexinostat allowed Fouliard et al. to build a PK/PD model of thrombocytopenia, which predicted the optimal administration schedule allowing higher doses with minimal thrombocytopenia (Fouliard et al. 2013). This optimized schedule is currently used in the trials in solid tumors with abexinostat. Exposure to anthracycline and trastuzumab was simulated based on available dosing records and by using a kinetic-pharmacodynamics (K-PD) and a fixed PK model from literature, respectively. PD models for troponin T and LVEF were successfully developed, identifying maximum troponin T concentration after anthracycline treatment as a significant determinant for trastuzumab-induced LVEF decline. These models can help identify patients at risk of drug-induced cardiotoxicity and optimize cardiac-monitoring strategies.

One of the contributing factors to the high attrition rate for developmental therapeutics in oncology is the inadequate dose and regimen selection combined with an insufficient understanding of the pharmacology to design an optimal drug development program (Postel-Vinay et al. 2016). The US Food and Drug Administration (FDA), European Medicines Agency (EMA), as well as Japan's PMDA consider quantitative modeling and simulation (M&S), including population PK analyses, population PK and PD model analyses, exposure–response analyses, and physiologically based pharmacokinetic (PBPK) model analyses, as useful tools that can provide actionable insights that inform the decision-making process in early-stage as well as late-stage oncology drug development programs. PK/PD models for anticancer agents have been developed and successfully applied to: (1) provide insights into fundamental mechanisms implicated in tumor growth, (2) assist in dose selection for first-in-human phase I studies (e.g., effective dose, escalating doses, and maximal tolerated doses), (3) design and optimize combination drug regimens, (4) design clinical trials, and (5) establish links between drug efficacy and safety and the concentrations of measured biomarkers (Eigenmann et al. 2017; Garralda et al. 2017; Gallo and Birtwistle 2015). The emergent field of pharmacometrics, defined as "the science of

developing and applying mathematical and statistical methods to (a) characterize, understand, and predict a drug's pharmacokinetic and PD behavior, (b) quantify uncertainty of information about that behavior, and (c) rationalize data-driven decision making in drug development process and pharmacotherapy," combines principles from pharmacology (PK and PD), statistics, and computational modeling to support drug development and optimize the use of already marketed drugs. Integrated population PK/PD/disease progression models as part of the pharmacometrics platform provide a powerful tool to predict outcomes so that the right dose can be given to the right patient to maximize drug efficacy and reduce drug toxicity (Buil-Bruna et al. 2016; Musuamba et al. 2017; Manolis et al. 2017).

The importance of PK-PD modeling for the drug development process is best illustrated in the example of pembrolizumab (Turner et al. 2018; Elassaiss-Schaap et al. 2017; Freshwater et al. 2017; Patnaik et al. 2015). Modeling helped identify the FDA-approved 2 mg/kg every 3 weeks dose schedule for nonsmall cell lung cancer therapy by predicting that this dose level which is much lower than the 10 mg/kg dose level studied in patients would have robust clinical activity due to intratumor drug exposure estimations. Notably, pembrolizumab was approved just 4 years after the phase I clinical trial started, through breakthrough designation by the FDA. This timeframe clearly contrasts with the 10 or greater years that former drugs traditionally took to be approved.

Model-informed drug development (MIDD) employs mathematical and statistical models to describe disease progression, PK, and PD to improve the clinical trial design and clinically relevant predictions. Advancing MIDD in oncology and identifying regulatory-acceptable best practices pertaining to MIDD will require close collaboration between drug makers and regulatory agencies as well as multistakeholder workshops, such as the 2014 EMA/European Federation of Pharmaceutical Industries and Associations (EFPIA) Workshop on Dose Finding and the 2018 public FDA-International Society of Pharmacometrics (ISOP) Workshop on Model-Informed Drug Development (https://www.fda.gov/Drugs/NewsEvents/ucm589449.htm) co-sponsored by the FDA's Center for Drug Evaluation and Research (CDER) and ISOP (Musuamba et al. 2017; Manolis et al. 2017; Schindler et al. 2018). The EMA Modelling and Simulation Working Group (MSWG) in collaboration with the FDA-Office of Clinical Pharmacology (OCP) pharmacometrics group strive to facilitate the much-needed harmonization on good M&S/MIDD practices through dialog and collaboration across all stakeholders. The desired goal is to develop "best practices in integrating PK, PD, efficacy, and safety data into models to best inform oncology drug development, evaluate disease- and mechanism-specific early endpoints to predict long-term efficacy, and discuss potential regulatory implications of model-informed decisions in drug development."

Emerging Role of Quantitative Systems Pharmacology for Model-Informed Drug Discovery and Development

Traditional PD models attempt to determine drug effects by integrating specific and confirmatory similar datasets and then predicting results in related scenarios. In this paradigm, the models are parsimonious, the parameters of the models can be identified, and therefore the models can incorporate population variability and define parameter uncertainties. An emerging new approach is termed "Quantitative Systems Pharmacology" that advances systems level, multiscale models for disease progression and treatment (Iyengar et al. 2012; Lai et al. 2018; Musante et al. 2017; Ribba et al. 2017). Systems level consideration of drug responses in these models attempt to better characterize the hierarchical, nonlinear, dynamic responses at the network level of drug action that may affect both efficacy and toxicity in clinical settings. These PD models are highly mechanistic and take into consideration the effects of drug actions spanning from the scale of molecular interactions to organ-level responses. Since these interactions are

nonlinear at the multiscale levels, the effects of drugs exhibit emergent behaviors relating to pronounced on-target and off-target PD actions of drug treatments. These models strive to integrate data from diverse datasets and hence have required the development of model exchange platforms such as PharmML (Bizzotto et al. 2017) and sophisticated toolboxes to perform multiscale simulations and apply nonlinear statistical analyses (Cheng et al. 2017; Eissing et al. 2011). These types of models prioritize biological detail over parameter identifiability and the simulations enable rich exploration of mechanistic variabilities to better identify on-target and off-target PD effects of drug action. An important aspect of these models relating to precision medicine goals is to explain differences in drug efficacy and toxicity in heterogeneous populations that display genetic or biomarker profile differences.

References and Further Reading

Ait-Oudhia S, Mager DE (2016) Array of translational systems pharmacodynamic models of anti-cancer drugs. J Pharmacokinet Pharmacodyn 43:549–565

Alsharedi M, Bukamur H, Elhamdani A (2018) Osimertinib for the treatment of patients with *EGFR* mutation-positive non-small cell lung cancer. Drugs Today (Barc) 54:369–379. https://doi.org/10.1358/dot.2018.54.6.2817668

Andronesi OC, Arrillaga-Romany IC, Ly KI, Bogner W, Ratai EM, Reitz K, Iafrate AJ, Dietrich J, Gerstner ER, Chi AS, Rosen BR, Wen PY, Cahill DP, Batchelor TT (2018) Pharmacodynamics of mutant-IDH1 inhibitors in glioma patients probed by in vivo 3D MRS imaging of 2-hydroxyglutarate. Nat Commun 9(1):1474. https://doi.org/10.1038/s41467-018-03905-6

Biankin AV, Piantadosi S, Hollingsworth SJ (2015) Patient-centric trials for therapeutic development in precision oncology. Nature 526:361–370. https://doi.org/10.1038/nature15819

Bizzotto R, Blaudez E, Borella E, Carrara L, Chan P, Chenel M, Comets E, Gieschke R, Harling K, Harnisch L, Hartung N, Hooker AC, Karlsson MO, Kaye R, Kloft C, Kokash N, Lavielle M, Lestini G, Magni P, Mari A, Mentré F, Muselle C, Nordgren R, Nyberg HB, Parra-Guillén ZP, Pasotti L, Rode-Kristensen N, Sardu ML, Smith GR, Swat MJ, Terranova N, Yngman G, Yvon F, Holford N, DDMoRe consortium (2017) Model description language (MDL): a standard for modeling and simulation. CPT Pharmacometrics Syst Pharmacol 6:647–650. https://doi.org/10.1002/psp4.12222

Block M (2015) Physiologically based pharmacokinetic and pharmacodynamic modeling in cancer drug development: status, potential and gaps. Expert Opin Drug Metab Toxicol 11:743–756. https://doi.org/10.1517/17425255.2015.1037276

Borchmann P, Goergen H, Kobe C, Lohri A, Greil R, Eichenauer DA, Zijlstra JM, Markova J, Meissner J, Feuring-Buske M, Hüttmann A, Dierlamm J, Soekler M, Beck HJ, Willenbacher W, Ludwig WD, Pabst T, Topp MS, Hitz F, Bentz M, Keller UB, Kühnhardt D, Ostermann H, Schmitz N, Hertenstein B, Aulitzky W, Maschmeyer G, Vieler T, Eich H, Baues C, Stein H, Fuchs M, Kuhnert G, Diehl V, Dietlein M, Engert A (2018) PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. Lancet 390:2790–2802. https://doi.org/10.1016/S0140-6736(17)32134-7

Brennan FR, Andrews L, Arulanandam AR, Blumel J, Fikes J, Grimaldi C, Lansita J, Loberg LI, MacLachlan T, Milton M, Parker S, Tibbitts J, Wolf J, Allamneni KP (2018) Current strategies in the non-clinical safety assessment of biologics: new targets, new molecules, new challenges. Regul Toxicol Pharmacol. https://doi.org/10.1016/j.yrtph.2018.07.0$32#09. pii: S0273-2300(18)30187-9

Buil-Bruna N, López-Picazo JM, Martín-Algarra S, Trocóniz IF (2016) Bringing model-based prediction to oncology clinical practice: a review of pharmacometrics principles and applications. Oncologist 21:220–232. https://doi.org/10.1634/theoncologist.$32#2015-0322

Bulgar AD, Snell M, Donze JR, Kirkland EB, Li L, Yang S, Xu Y, Gerson SL, Liu L (2010) Targeting base excision repair suggests a new therapeutic strategy of fludarabine for the treatment of chronic lymphocytic leukemia. Leukemia 4:1795–1799

Byrd JC, Harrington B, O'Brien S, Jones JA, Schuh A, Devereux S, Chaves J, Wierda WG, Awan FT, Brown JR, Hillmen P, Stephens DM, Ghia P, Barrientos JC, Pagel JM, Woyach J, Johnson D, Huang J, Wang X, Kaptein A, Lannutti BJ, Covey T, Fardis M, McGreivy J, Hamdy A, Rothbaum W, Izumi R, Diacovo TG, Johnson AJ, Furman RR (2016) Acalabrutinib (ACP-196) in relapsed chronic lymphocytic leukemia. N Engl J Med 374:323–332

Caimi PF, Cooper BW, William BM, Dowlati A, Barr PM, Fu P, Pink J, Xu Y, Lazarus HM, de Lima M, Gerson SL (2017) Phase I clinical trial of the base excision repair inhibitor methoxyamine in combination with fludarabine for patients with advanced hematologic malignancies. Oncotarget 8:79864–79875. https://doi.org/10.18632/oncotarget.20094

Cheng Y, Thalhauser CJ, Smithline S, Pagidala J, Miladinov M, Vezina HE, Gupta M, Leil TA, Schmidt BJ (2017) QSP toolbox: computational implementation of integrated workflow components for deploying multi-scale mechanistic models. AAPS J 19:1002–1016

Ciccolini J, Barbolosi D, Meille C, Lombard A, Serdjebi C, Giacometti S, Padovani L, Pasquier E, André N (2017) Pharmacokinetics and pharmacodynamics-based mathematical modeling identifies an optimal protocol for metronomic chemotherapy. Cancer Res 77(17):1–11. https://doi.org/10.1158/0008-5472.CAN-16-3130

Claret L, Girard P, Hoff PM, Van Cutsem E, Zuideveld KP, Jorga K, Fagerberg J, Bruno R (2009) Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. J Clin Oncol 27(25):4103–4108. https://doi.org/10.1200/JCO.2008.21.0807

Claret L, Mercier F, Houk BE, Milligan PA, Bruno R (2015) Modeling and simulations relating overall survival to tumor growth inhibition in renal cell carcinoma patients. Cancer Chemother Pharmacol 76(3):567–573. https://doi.org/10.1007/s00280-015-2820-x

Cook N, Hansen AR, Siu LL, Abdul Razak AR (2015) Early phase clinical trials to identify optimal dosing and safety. Mol Oncol 9:997–1007

Creelan BC, Gabrilovich DI, Gray JE, Williams CC, Tanvetyanon T, Haura EB, Weber JS, Gibney GT, Markowitz J, Proksch JW, Reisman SA, McKee MD, Chin MP, Meyer CJ, Antonia SJ (2017) Safety, pharmacokinetics, and pharmacodynamics of oral omaveloxolone (RTA 408), a synthetic triterpenoid, in a first-in-human trial of patients with advanced solid tumors. Onco Targets Ther 10:4239–4250. https://doi.org/10.2147/OTT.S136992

Dal Molin GZ, Omatsu K, Sood AK, Coleman RL (2018) Rucaparib in ovarian cancer: an update on safety, efficacy and place in therapy. Ther Adv Med Oncol 10. https://doi.org/10.1177/1758835918778483

de Man FM, Goey AKL, van Schaik RHN, Mathijssen RHJ, Bins S (2018) Individualization of irinotecan treatment: a review of pharmacokinetics, pharmacodynamics, and pharmacogenetics. Clin Pharmacokinet 57:1229. https://doi.org/10.1007/s40262-018-0644-7

de Vries Schultink AHM, Boekhout AH, Gietema JA, Burylo AM, Dorlo TPC, van Hasselt JGC, Schellens JHM, Huitema ADR (2018) Pharmacodynamic modeling of cardiac biomarkers in breast cancer patients treated with anthracycline and trastuzumab regimens. J Pharmacokinet Pharmacodyn 45:431–442. https://doi.org/10.1007/s10928-018-9579-8

Derendorf H, Lesko LJ, Chaikin P, Colburn WA, Lee P, Miller R, Powell R, Rhodes G, Stanski D, Venitz J (2000) Pharmacokinetic/pharmacodynamic modeling in drug research and development. J Clin Pharmacol 40:1399–1418

DiNardo CD, Stein EM, de Botton S, Roboz GJ, Altman JK, Mims AS, Swords R, Collins RH, Mannis GN, Pollyea DA, Donnellan W, Fathi AT, Pigneux A, Erba HP, Prince GT, Stein AS, Uy GL, Foran JM, Traer E, Stuart RK, Arellano ML, Slack JL, Sekeres MA, Willekens C, Choe S, Wang H, Zhang V, Yen KE, Kapsalis SM, Yang H, Dai D, Fan B, Goldwasser M, Liu H, Agresta S, Wu B, Attar EC, Tallman MS, Stone RM, Kantarjian HM (2018) Durable remissions with Ivosidenib in IDH1-mutated relapsed or refractory AML. N Engl J Med 378:2386–2398. https://doi.org/10.1056/NEJMoa1716984

Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, Nathenson M, Doebele RC, Farago AF, Pappo AS, Turpin B, Dowlati A, Brose MS, Mascarenhas L, Federman N, Berlin J, El-Deiry WS, Baik C, Deeken J, Boni V, Nagasubramanian R, Taylor M, Rudzinski ER, Meric-Bernstam F, Sohal DPS, Ma PC, Raez LE, Hechtman JF, Benayed R, Ladanyi M, Tuch BB, Ebata K, Cruickshank S, Ku NC, Cox MC, Hawkins DS, Hong DS, Hyman DM (2018) Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. N Engl J Med 378:731–739. https://doi.org/10.1056/NEJMoa1714448

Eigenmann MJ, Frances N, Lave T, Walz A-C (2017) PKPD modeling of acquired resistance to anti-cancer drug treatment. J Pharmacokinet Pharmacodyn 44:617–630. https://doi.org/10.1007/s10928-017-955$32#3-x

Eissing T, Kuepfer L, Becker C, Block M, Coboeken K, Gaub T, Goerlitz L, Jaeger J, Loosen R, Ludewig B, Meyer M, Niederalt C, Sevestre M, Siegmund HU, Solodenko J, Thelen K, Telle U, Weiss W, Wendl T, Willmann S, Lippert J (2011) A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. Front Physiol 2:4

Elassaiss-Schaap J, Rossenu S, Lindauer A, Kang SP, de Greef R, Sachs JR, de Alwis DP (2017) Using model-based "learn and confirm" to reveal the pharmacokinetics-pharmacodynamics relationship of pembrolizumab in the KEYNOTE-001 trial. CPT Pharmacometrics Syst Pharmacol 6:21–28

FDA Guidance on Clinical Pharmacology Section of Labeling for Human Prescription Drug and Biological Products (2016) https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm109739.pdf

Fouladi M, Park JR, Stewart CF, Gilbertson RJ, Schaiquevich P, Sun J, Reid JM, Ames MM, Speights R, Ingle AM, Zwiebel J, Blaney SM, Adamson PC (2010) Pediatric phase I trial and pharmacokinetic study of vorinostat: a children's oncology group phase I consortium report. J Clin Oncol 28:3623–3629

Fouliard S, Robert R, Jacquet-Bescond A, Rieu C, Balasubramanian RS, Loury D, Loriot Y, Hollebecque A, Kloos I, Soria JC, Chenel M, Depil S (2013) Pharmacokinetic/pharmacodynamic modelling-based optimisation of administration schedule for the histone deacetylase inhibitor abexinostat (S78454/PCI-24781) in phase I. Eur J Cancer 49:2791–2797. https://doi.org/10.1016/j.ejca.2013.05.009. Get rights and content

Freshwater T, Kondic A, Ahamadi M, Li CH, de Greef R, de Alwis D, Stone JA (2017) Evaluation of dosing strategy for pembrolizumab for oncology indications.

J Immunother Cancer 5:43. https://doi.org/10.1186/s40425-017-0242-5. eCollection 2017

Gallo JM, Birtwistle MR (2015) Network pharmacodynamic models for customized cancer therapy. WIREs Syst Biol Med 7:243–251. https://doi.org/10.1002/wsbm.1300

Garralda E, Dienstmann R, Tabernero J (2017) Pharmacokinetic/pharmacodynamic modeling for drug development in oncology. Am Soc Clin Oncol Educ Book 37:210–215. https://doi.org/10.14694/EDBK_180460

Goey AK, Sissung TM, Peer CJ, Trepel JB, Lee MJ, Tomita Y, Ehrlich S, Bryla C, Balasubramaniam S, Piekarz R, Steinberg SM, Bates SE, Figg WD (2016) Effects of UGT1A1 genotype on the pharmacokinetics, pharmacodynamics, and toxicities of belinostat administered by 48-hour continuous infusion in patients with cancer. J Clin Pharmacol 56:461–473. https://doi.org/10.1002/jcph.625

Gray HJ, Bell-McGuinn K, Fleming GF, Cristea M, Xiong H, Sullivan D, Luo Y, McKee MD, Munasinghe W, Martin LP (2018) Phase I combination study of the PARP inhibitor veliparib plus carboplatin and gemcitabine in patients with advanced ovarian cancer and other solid malignancies. Gynecol Oncol 148:507–514. https://doi.org/10.1016/j.ygyno.2017.$32#12.029

Grisafi D, Ceschi A, Sava G, Avalos Clerici V, Scaglione F (2018) Reconsidering clinical pharmacology frameworks as a necessary strategy for improving the health care of patients: a systematic review. Eur J Clin Pharmacol. https://doi.org/10.1007/s00228-018-2511-8. [Epub ahead of print]

Hansen AR, Cook N, Amir E, Siu LL, Razak ARA (2017) Determinants of the recommended phase 2 dose of molecular targeted agents. Cancer 123:1409–1415

Hidalgo MM, Epelbaum R, Semenisty V, Geva R, Golan T, Borazanci EH, Stemmer SM, Borad MJ, Park JO, Pedersen K, Wolpin BM, Ramirez RA, Becerra C, Shaw S, Oberkovitz G, Haras AV, Sorani E, Lustig TM, Peled A, Von Hoff DD (2018) Evaluation of pharmacodynamic (PD) biomarkers in patients with metastatic pancreatic cancer treated with BL-8040, a novel CXCR4 antagonist. J Clin Oncol 36(Suppl 5):88

Hirakawa A, Yonemori K, Kinoshita F, Kobayashi Y, Okuma HS, Kawachi A, Tamura K, Fujiwara Y, Rubinstein L, Harris PJ, Takebe N (2018) Potential utility of a longitudinal relative dose intensity of molecularly targeted agents in phase 1 dose-finding trials. Cancer Sci 109:207–214. https://doi.org/10.1111/cas.$32#13436

Honigberg LA, Smith AM, Sirisawad M, Verner E, Loury D, Chang B, Li S, Pan Z, Thamm DH, Miller RA, Buggy JJ (2010) The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. Proc Natl Acad Sci U S A 107:13075–13080

Ishiguro H, Saji S, Nomura S, Tanaka S, Ueno T, Onoue M, Iwata H, Yamanaka T, Sasaki Y, Toi M (2017) A phase I/II pharmacokinetics/pharmacodynamics study of irinotecan combined with S-1 for recurrent/metastatic breast cancer in patients with selected UGT1A1 genotypes (the JBCRG-M01 study). Cancer Med 6:2909–2917. https://doi.org/10.1002/cam4.1258

Iyengar R, Zhao S, Chung SW, Mager DE, Gallo JM (2012) Merging systems biology with pharmacodynamics. Sci Transl Med. https://doi.org/10.1126/scitranslmed.3003563

Jain RK, Lee JJ, Hong D, Markman M, Gong J, Naing A, Wheler J, Kurzrock R (2010) Phase I oncology studies: evidence that in the era of targeted therapies patients on lower doses do not fare worse. Clin Cancer Res 16:1289–1297

Jamal R, Lapointe R, Cocolakis E, Thébault P, Kazemi S, Friedmann JE, Dionne J, Cailhier JF, Bélanger K, Ayoub JP, Le H, Lambert C, El-Hajjar J, van Kempen LC, Spatz A, Miller WH Jr (2017) Peripheral and local predictive immune signatures identified in a phase II trial of ipilimumab with carboplatin/paclitaxel in unresectable stage III or stage IV melanoma. J Immunother Cancer 5:83. https://doi.org/10.1186/s40425-017-0290-x

Jänne PA, Kim G, Shaw AT, Sridhara R, Pazdur R, McKee AE (2016) Dose finding of small-molecule oncology drugs: optimization throughout the development life cycle. Clin Cancer Res 22:2613–2617

Krebs M, Dive C, Dean EJ, Rothwell DG, Brognard J, Wallace A (2016) TARGET trial: molecular profiling of circulating tumour DNA to stratify patients to early phase clinical trials. J Clin Oncol 34:TPS11614. http://www.ascopubs.org/doi/abs/10.1200/JCO.2016.34.15_suppl.TPS11614#affiliationsContainer

Laetsch TW, DuBois SG, Nagasubramanian R, Turpin B, Mascarenhas L, Federman N (2017) A pediatric phase I study of larotrectinib, a highly selective inhibitor of the tropomyosin receptor kinase (TRK) family. J Clin Oncol 35(Suppl 15):10510. http://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15_suppl.10510

Lai X, Geier O, Fleischer T, Garred O, Borgen EF, Funke S, Kumar S, Rognes ME, Seierstad T, Boerresen-Dale A-L, Kristensen VN, Engebraaten O, Kohn-Luque A, Frigessi A (2018) Towards personalized computer simulation of breast cancer treatment: a multi-scale pharmacokinetic and pharmacodynamic model informed by multi-type patient data. bioRxiv. https://doi.org/10.1101/371369

Lavezzi SM, Borella E, Carrara L, De Nicolao G, Magni P, Poggesi I (2018) Mathematical modeling of efficacy and safety for anticancer drugs clinical development. Expert Opin Drug Discovery 13(1):5–21. https://doi.org/10.1080/17460441.2018.1388369

Li L, Ma L, Schrieber SJ, Rahman NA, Deisseroth A, Farrell AT, Wang Y, Sinha V, Marathe A (2018) Quantitative relationship between AUEC of absolute neutrophil count and duration of severe neutropenia for G-CSF in breast cancer patients. Clin Pharmacol Ther. https://doi.org/10.1002/cpt.991. [Epub ahead of print]

Lu D, Lu T, Stroh M, Graham RA, Agarwal P, Musib L, Li CC, Lum BL, Joshi A (2016) A survey of new

oncology drug approvals in the USA from 2010 to 2015: a focus on optimal dose and related post-marketing activities. Cancer Chemother Pharmacol 77:459–476

Mahalingam D, Goel S, Aparo S, Patel Arora S, Noronha N, Tran H, Chakrabarty R, Selvaggi G, Gutierrez A, Coffey M, Nawrocki ST, Nuovo G, Mita MM (2018) A phase II study of Pelareorep (REOLYSIN®) in combination with gemcitabine for patients with advanced pancreatic adenocarcinoma. Cancers (Basel) 10(6). https://doi.org/10.3390/cancers10060160. pii: E160

Maitland ML, Piha-Paul S, Falchook G, Kurzrock R, Nguyen L, Janisch L, Karovic S, McKee M, Hoening E, Wong S, Munasinghe W, Palma J, Donawho C, Lian GK, Ansell P, Ratain MJ, Hong D (2018) Clinical pharmacodynamic/exposure characterisation of the multikinase inhibitor ilorasertib (ABT-348) in a phase 1 dose-escalation trial. Br J Cancer 118:1042–1050. https://doi.org/10.1038/s4141$32#6-018-0020-2

Manolis E, Brogren J, Cole S, Hay JL, Nordmark A, Karlsson KE, Lentz F, Benda N, Wangorsch G, Pons G, Zhao W, Gigante V, Serone F, Standing JF, Dokoumetzidis A, Vakkilainen J, van den Heuvel M, Mangas Sanjuan V, Taminiau J, Kerwash E, Khan D, Musuamba FT, Skottheim Rusten I, EMA Modelling and Simulation Working Group (2017) Commentary on the MID3 good practices paper. CPT Pharmacometrics Syst Pharmacol 6:416–417. https://doi.org/10.1002/psp4.12223. Epub 2017 Jul 22

Meille C, Guerreiro N, Jullion A, Bauer S, Chatterjee M, de Vos F, Lin CC, Suárez Rodríguez C, Tai D, Cassier P, Yamamoto N, de Weger VA, Klopfenstein M, Mariconti L, Kuriakose ET, Hyman DM (2017) Optimization of the dose and schedule of an HDM2 inhibitor NVP-HDM201 in a first-in-human phase I study using a mechanism-based PK/PD model [abstract]. In: Proceedings of the American Association for Cancer Research annual meeting 2017; 2017 Apr 1–5, Washington, DC. AACR, Philadelphia. https://doi.org/10.1158/1538-7445.AM2017-CT154. Cancer Res 2017;77(13 Suppl): Abstract No: ct 154

Moreau P, Karamanesht II, Domnikova N, Kyselyova MY, Vilchevska KV, Doronin VA, Schmidt A, Hulin C, Leleu X, Esseltine DL, Venkatakrishnan K, Skee D, Feng H, Girgis S, Cakana A, van de Velde H, Deraedt W, Facon T (2012) Pharmacokinetic, pharmacodynamic and covariate analysis of subcutaneous versus intravenous administration of bortezomib in patients with relapsed multiple myeloma. Clin Pharmacokinet 51:823–829. https://doi.org/10.1007/s40262-012-0010-0

Musante CJ, Ramanujan S, Schmidt BJ, Ghobrial OG, Lu J, Heatherington AC (2017) Quantitative systems pharmacology: a case for disease models. Clin Pharmacol Ther 101:24–27

Musuamba FT, Manolis E, Holford N, Cheung S, Friberg LE, Ogungbenro K, Posch M, Yates J, Berry S, Thomas N, Corriol-Rohou S, Bornkamp B, Bretz F, Hooker AC, Van der Graaf PH, Standing JF, Hay J, Cole S, Gigante V, Karlsson K, Dumortier T, Benda N, Serone F, Das S, Brochot A, Ehmann F, Hemmings R, Rusten IS (2017) Advanced methods for dose and regimen finding during drug development: summary of the EMA/EFPIA workshop on dose finding (London 4–5 December 2014). CPT Pharmacometrics Syst Pharmacol 6:418–429. https://doi.org/10.1002/psp4.12196. Epub 2017 Jul 19

Mutti L, Peikert T, Robinson BWS, Scherpereel A, Tsao AS, de Perrot M, Woodard GA, Jablons DM, Wiens J, Hirsch FR, Yang H, Carbone M, Thomas A, Hassan R (2018). State of the art: concise review scientific advances and new frontiers in mesothelioma therapeutics. J Thorac Oncol. https://doi.org/10.1016/j.jtho.2018.06.011 . pii: S1556-0864(18)30720-2. [Epub ahead of print]

Nightingale G, Schwartz R, Kachur E, Dixon BN, Cote C, Barlow A, Barlow B, Medina P (2018) Clinical pharmacology of oncology agents in older adults: a comprehensive review of how chronologic and functional age can influence treatment-related effects. J Geriatr Oncol. https://doi.org/10.1016/j.jgo.2018.06.008. pii: S1879-4068(18)30161-9

Niu J, Scheuerell C, Mehrotra S, Karan S, Puhalla S, Kiesel BF, Ji J, Chu E, Gopalakrishnan M, Ivaturi V, Gobburu J, Beumer JH (2017) Parent-metabolite pharmacokinetic modeling and pharmacodynamics of veliparib (ABT-888), a PARP inhibitor, in patients with BRCA 1/2-mutated cancer or PARP-sensitive tumor types. J Clin Pharmacol 57:977–987. https://doi.org/10.1002/jcph.892

Owonikoko TK, Busari AK, Kim S, Chen Z, Akintayo A, Lewis C, Carthon BC, Alese OB, El-Rayes BF, Ramalingam SS, Harvey RD (2018) Race-, age-, and gender-based characteristics and toxicities of targeted therapies on phase I trials. Oncology 95:138. https://doi.org/10.1159/000488763. [Epub ahead of print]

Palmirotta R, Carella C, Silvestris E, Cives M, Stucci SL, Tucci M, Lovero D, Silvestris F (2018) SNPs in predicting clinical efficacy and toxicity of chemotherapy: walking through the quicksand. Oncotarget 9:25355–25382. https://doi.org/10.18632/oncotarget.25256

Patel VK, Lamothe B, Ayres ML, Gay J, Cheung JP, Balakrishnan K, Ivan C, Morse J, Nelson M, Keating MJ, Wierda WG, Marszalek JR, Gandhi V (2018) Pharmacodynamics and proteomic analysis of acalabrutinib therapy: similarity of on-target effects to ibrutinib and rationale for combination therapy. Leukemia 32:920–930. https://doi.org/10.1038/leu.2017.321

Patnaik A, Kang SP, Rasco D, Papadopoulos KP, Elassaiss-Schaap J, Beeram M, Drengler R, Chen C, Smith L, Espino G, Gergich K, Delgado L, Daud A, Lindia JA, Li XN, Pierce RH, Yearley JH, Wu D, Laterza O, Lehnert M, Iannone R, Tolcher AW (2015) Phase I study of pembrolizumab (MK-3475; anti-PD-1 monoclonal antibody) in patients with advanced solid tumors. Clin Cancer Res 21:4286–4293

Peck RW (2015) The right dose for every patient: a key step for precision medicine. Nat Rev Drug Discov 15:145–146. Available from http://www.nature.com/doifinder/10.1038/nrd.2015.22

Peer CJ, Hall OM, Sissung TM, Piekarz R, Balasubramaniam S, Bates SE, Figg WD (2018) A population pharmacokinetic/toxicity model for the reduction of platelets during a 48-h continuous intravenous infusion of the histone deacetylase inhibitor belinostat. Cancer Chemother Pharmacol 82:565–570. https://doi.org/10.1007/s00280-018-3631-7

Phelan JD, Young RM, Webster DE, Roulland S, Wright GW, Kasbekar M, Shaffer AL 3rd, Ceribelli M, Wang JQ, Schmitz R, Nakagawa M, Bachy E, Huang DW, Ji Y, Chen L, Yang Y, Zhao H, Yu X, Xu W, Palisoc MM, Valadez RR, Davies-Hill T, Wilson WH, Chan WC, Jaffe ES, Gascoyne RD, Campo E, Rosenwald A, Ott G, Delabie J, Rimsza LM, Rodriguez FJ, Estephan F, Holdhoff M, Kruhlak MJ, Hewitt SM, Thomas CJ, Pittaluga S, Oellerich T, Staudt LM (2018) A multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature 560:387–391. https://doi.org/10.1038/s41586-018-0290-0

Pinto N, DuBois SG, Marachelian A, Diede SJ, Taraseviciute A, Glade Bender JL, Tsao-Wei D, Groshen SG, Reid JM, Haas-Kogan DA, Reynolds CP, Kang MH, Irwin MS, Macy ME, Villablanca JG, Matthay KK, Park JR (2018) Phase I study of vorinostat in combination with isotretinoin in patients with refractory/recurrent neuroblastoma: a new approaches to neuroblastoma therapy (NANT) trial. Pediatr Blood Cancer 65:e27023. https://doi.org/10.1002/pbc.27023

Postel-Vinay S, Aspeslagh S, Lanoy E, Robert C, Soria JC, Marabelle A (2016) Challenges of phase 1 clinical trials evaluating immune checkpoint-targeted antibodies. Ann Oncol 27:214–224

Ribba B, Grimm HP, Agoram B, Davies MR, Gadkar K, Niederer S, van Riel N, Timmis J, van der Graaf PH (2017) Methodologies for quantitative systems pharmacology (QSP) models: design and estimation. CPT Pharmacometrics Syst Pharmacol 6:496–498

Rocchetti M, Del Bene F, Germani M, Fiorentini F, Poggesi I, Pesenti E, Magni P, De Nicolao G (2009) Testing additivity of anticancer agents in pre-clinical studies: a PK/PD modelling approach. Eur J Cancer 45(18):3336–3346

Sato M, Ochiai Y, Kijima S, Nagai N, Ando Y, Shikano M, Nomura Y (2017) Quantitative modeling and simulation in PMDA: a Japanese regulatory perspective. CPT Pharmacometrics Syst Pharmacol 6:413–415

Schindler E (2018) Pharmacometrics to improve clinical benefit assessment in oncology. Digital comprehensive summaries of Uppsala Dissertations from the Faculty of Pharmacy 243. ISSN 1651-6192 ISBN 978-91-513-0191-4urn:nbn:se:uu:diva-336420

Socinski MA, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, Nogami N, Rodríguez-Abreu D, Moro-Sibilot D, Thomas CA, Barlesi F, Finley G, Kelsch C, Lee A, Coleman S, Deng Y, Shen Y, Kowanetz M, Lopez-Chavez A, Sandler A, Reck M, IMpower150 Study Group (2018) Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. N Engl J Med 378:2288–2301. https://doi.org/10.1056/NEJMoa1716948

Stein S, Zhao R, Haeno H, Vivanco I, Michor F (2018) Mathematical modeling identifies optimum lapatinib dosing schedules for the treatment of glioblastoma patients. PLoS Comput Biol 14(1):e1005924. https://doi.org/10.1371/journal.pcbi.1005924

Tam CS, Anderson MA, Pott C, Agarwal R, Handunnetti S, Hicks RJ, Burbury K, Turner G, Di Iulio J, Bressel M, Westerman D, Lade S, Dreyling M, Dawson SJ, Dawson MA, Seymour JF, Roberts AW (2018) Ibrutinib plus venetoclax for the treatment of mantle-cell lymphoma. N Engl J Med 378:1211–1223. https://doi.org/10.1056/NEJMoa171$32#5519

Tan CRC, Abdul-Majeed S, Cael B, Barta SK (2018) Clinical pharmacokinetics and pharmacodynamics of bortezomib. Clin Pharmacokinet. https://doi.org/10.1007/s40262-018-0679-9. [Epub ahead of print]

Tham LS, Wang L, Soo RA, Lee SC, Lee HS, Yong WP, Goh BC, Holford NH (2008) A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. Clin Cancer Res 14(13):4213–4218

Torres-Ayuso P, Sahoo S, Ashton G, An E, Simms N, Galvin M, Leong HS, Frese KK, Simpson K, Cook N, Hughes A, Miller CJ, Marais R, Dive C, Krebs MG, Brognard J (2018) Signaling pathway screening platforms are an efficient approach to identify therapeutic targets in cancers that lack known driver mutations: a case report for a cancer of unknown primary origin. NPJ Genom Med 3:15. https://doi.org/10.1038/s41525-018-0055-6. eCollection 2018

Tsimberidou AM, Tsimberidou AM, Hong DS, Ye Y, Cartwright C, Wheler JJ, Falchook GS, Naing A, Fu S, Piha-Paul S, Janku F, Meric-Bernstam F, Hwu P, Kee B, Kies MS, Broaddus R, Mendelsohn J, Hess KR, Kurzrock R (2017) Initiative for molecular profiling and advanced cancer therapy (IMPACT): an MD Anderson Precision Medicine Study. JCO Precis Oncol. https://doi.org/10.1200/PO.17.00002

Turan T, Kannan D, Patel M, Matthew Barnes J, Tanlimco SG, Lu R, Halliwill K, Kongpachith S, Kline DE, Hendrickx W, Cesano A, Butterfield LH, Kaufman HL, Hudson TJ, Bedognetti D, Marincola F, Samayoa J (2018) Immune oncology, immune responsiveness and the theory of everything. J Immunother Cancer 6(1):50. https://doi.org/10.1186/s40425-018-0355-5

Turner D, Kondic AG, Anderson KM, Robinson A, Garon EB, Riess JW, Jain L, Mayawala K, Kang J, Ebbinghaus SW, Sinha V, de Alwis DP, Stone JA (2018) Pembrolizumab exposure-response assessments challenged by association of cancer cachexia and

catabolic clearance. Clin Cancer Res. https://doi.org/10.1158/1078-0432.CCR-18-0415. pii: clincanres.0415.2018. [Epub ahead of print]

Uckun FM, Qazi S (2010) Bruton's tyrosine kinase as a molecular target in treatment of leukemias and lymphomas as well as inflammatory disorders and autoimmunity. Expert Opin Ther Patents 20:1457–1470

Uckun FM, Zheng Y, Cetkovic-Cvrlje M, Vassilev A, Lisowski E, Waurzyniak B, Chen H, Carpenter R, Chen CL (2002) In vivo pharmacokinetic features, toxicity profile, and chemosensitizing activity of alpha-cyano-beta-hydroxy-beta-methyl-*N*-(2,5-dibromophenyl)propenamide (LFM-A13), a novel antileukemic agent targeting Bruton's tyrosine kinase. Clin Cancer Res 8:1224–1233

Uckun FM, Vassilev A, Dibirdik I, Tibbles HE (2007) Targeting JAK3 kinase-linked signal transduction pathways with rationally designed inhibitors [Review]. Anti Cancer Agents Med Chem 7:612–623

Uckun FM, Qazi S, Ma H, Tuel-Ahlgren L, Ozer Z (2010) STAT3 is a substrate of SYK tyrosine kinase in B-lineage leukemia/lymphoma cells exposed to oxidative stress. Proc Natl Acad Sci U S A 107:2902–2907

Valla K, Flowers CR, Koff JL (2018) Targeting the B cell receptor pathway in non-Hodgkin lymphoma. Expert Opin Investig Drugs 27:513–522. https://doi.org/10.1080/13543784.2018.1482273

Waller CF, Tiessen RG, Lawrence TE, Shaw A, Liu MS, Sharma R, Baczkowski M, Kothekar MA, Micales CE, Barve A, Ranganna GM, Pennella EJ (2018) A pharmacokinetics and pharmacodynamics equivalence trial of the proposed pegfilgrastim biosimilar, MYL-1401H, versus reference pegfilgrastim. J Cancer Res Clin Oncol 144:1087–1095. https://doi.org/10.1007/s00432#2-018-2643-3

Wicki A, Brown N, Xyrafas A, Bize V, Hawle H, Berardi S, Cmiljanović N, Cmiljanović V, Stumm M, Dimitrijević S, Herrmann R, Prêtre V, Ritschard R, Tzankov A, Hess V, Childs A, Hierro C, Rodon J, Hess D, Joerger M, von Moos R, Sessa C, Kristeleit R (2018) First-in human, phase 1, dose-escalation pharmacokinetic and pharmacodynamic study of the oral dual PI3K and mTORC1/2 inhibitor PQR309 in patients with advanced solid tumors (SAKK 67/13). Eur J Cancer 96:6–16. https://doi.org/10.1016/j.ejca.2018.03.012

Zamboni WC, D'Argenio DZ, Stewart CF, MacVittie T, Delauter BJ, Farese AM, Potter DM, Kubat NM, Tubergen D, Egorin MJ (2001) Pharmacodynamic model of topotecan-induced time course of neutropenia. Clin Cancer Res 7(8):2301–2308

Zhou Q, Gallo JM (2011) The pharmacokinetic/pharmacodynamic pipeline: translating anticancer drug pharmacology to the clinic. AAPS J 13(1):111–120. https://doi.org/10.1208/s12248-011-9253-1

Pharmacodynamic Evaluation: Gene Therapy

17

Nicolas Grandchamp

Contents

Abstract

Gene therapy is based on the genome modifications by introducing nucleic acids in the cells of the patient to treat diseases. There are different gene therapy strategies, the most common of which is to add a healthy copy of the disease-inducing gene. Other strategies consist of replacing the gene which is responsible for the disease for a healthy copy, or to inactivate a gene that is functioning improperly. Gene therapy is designed for untreatable diseases or with a heavy treatment.

N. Grandchamp (✉)
Biosource / GEG Tech, Paris, France
e-mail: nicolas.grandchamp@geg-tech.com

To modify the genome of patients, different gene vectors are used. There are two major classes of them, synthetic vectors which are complexes between nucleic acids (DNA, RNA, or protein) and synthetic molecules and viral vectors which are derived from viruses. Each of them has their advantages and limits.

The vectors can be administered directly in vivo or used for ex vivo approaches which consist to genetically modify the patient cells in vitro before being reintroduced into the patient.

This chapter gives a statement of the art of gene therapy. It describes the different strategies and tools used in this area before describing the seminal clinical trials and the range of diseases for which gene therapy could be

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_51

relevant. This chapter concludes by considering the issues and challenges of the gene therapy.

Gene Therapy: Concept and Strategies

There are several ways and contexts to genetically modify cells with different implications in terms of technology, biosafety, and ethics. Firstly, the origin of the cells used somatic or germline cells. In the first case, the genome modifications of patient cells cannot be passed on to subsequent generations of children, which is not the case with germline cells. Therefore, although the genetic modification of germline cells can be attractive in the field of transgenesis, this way is not explored to design gene therapy given the ethical and societal implications.

Somatic gene therapy can be broadly split into two categories: in vivo and ex vivo gene therapy (Fig. 1). Ex vivo gene therapy consists of modifying the patient's cell outside his body. The cells are cultured and genetically modified in vitro in labs, and then they are selected and reintroduced into the patient to treat the disease. Ex vivo gene therapy can be applied only to some cell types or selected tissues and is frequently used for diseases involving bone marrow cells. Compared to the in vivo gene therapy, ex vivo gene therapy is less likely to induce adverse immunological reactions in the patient's body since the genetic correction is done in vitro. The success depends on stable genic modification after reintroduction into the patient's body and the severity rate of side effects which could be induced by the genic modification.

In vivo gene therapy consists of directly modifying the genome of the cells inside the patient's body to treat genetic diseases. It can be administered directly in the blood by systemic injection or applied specifically in a tissue such as the liver, muscle, skin, lung, spleen, eye, or brain. The success depends on several factors, such as efficient delivery of the vector to the target cells, extracellular and intracellular degradation of the vector, the stability of the genic modification, and the level of toxicity or severity rate of side effects induced by the vector or the genic modification. The side effects which may be observed directly depend on the vector type used and the genic modification induced.

There are different types of gene modifications, the choice being guided by the diseases and the experimental constraints. The common way is to add a therapeutic gene to restore a function of the body that is working improperly. The strategy is to add a transgene, which is integrated in the genome of the patient, to be expressed over a long period. However, side effects can be observed such as genotoxicity due to the locus of integration which can interfere with the expression of other genes such as oncogenes (Hacein-Bey-Abina

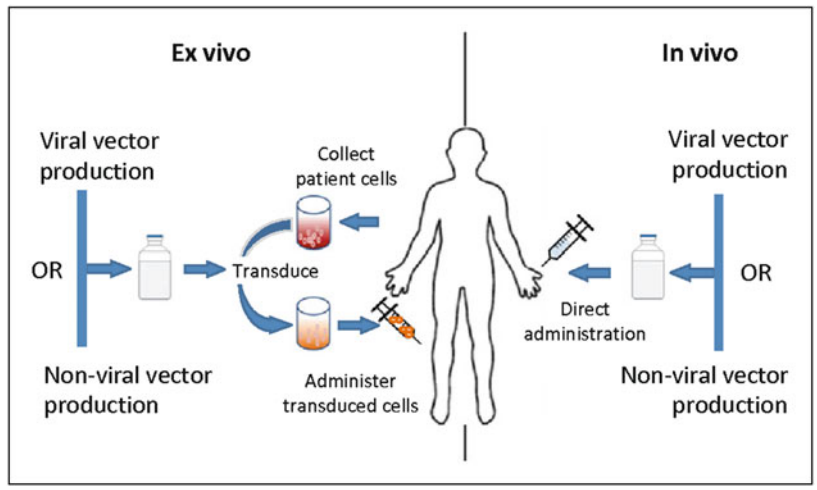


Fig. 1 Strategies of in vivo gene therapy and ex vivo gene therapy

et al. 2003). To avoid that, strategies allowing the targeting of the integration of the therapeutic gene in a safe locus are developed.

Another strategy consists of replacing the defective gene for a healthy copy thanks to the homologous recombination (HR) process of the cells. For this purpose, genome editing tools are used to induce a DNA double-strand break which will promote RH mechanisms of the cell (Fig. 2). This strategy allows to avoid side effects induced by the genome modification. Indeed, replacing, and not adding, a defective gene by a healthy copy avoids genotoxic effects induced by an unsafe locus of integration. Furthermore, this strategy allows to have a physiologic expression of the gene which will regulate as the wild type, restoring the phenotype in recessive and dominant genetic disease. The issue is to very accurately design genome editing tools able to precisely and efficiently perform genomic recombination into the desired locus.

Another way for some diseases is to inhibit or inactivate the defective gene when that provides a therapeutic effect. This type of strategy has been firstly explored by introducing repressive RNA specific sequences, such as miRNA or shRNA, to inhibit the expression of the target gene (Moroni et al. 1992). However, the efficiency of this approach is often too low to observe a therapeutic effect. Another way has been explored developing the strategy of exon skipping (Benchaouir et al. 2007). This smart approach consists of introducing small DNA probes which interfere in splicing to eliminate the mutated part of the protein. However, this strategy can be used only in very specific contexts, when truncated proteins can have a therapeutic effect. A third way to inactivate a gene, which is very attractive due to its high efficiency, is to cut the gene sequence to induce nonhomologous end joining (NHEJ) DNA repair mechanisms which have the potential to introduce mutations inactivating the gene (Fig. 2).

All those strategies are very attractive, but integration of the vector and/or the use of genome editing tools poses the principal issue of avoiding side effects by controlling the locus of genomic modification. As described in the next chapter, several tools are available;

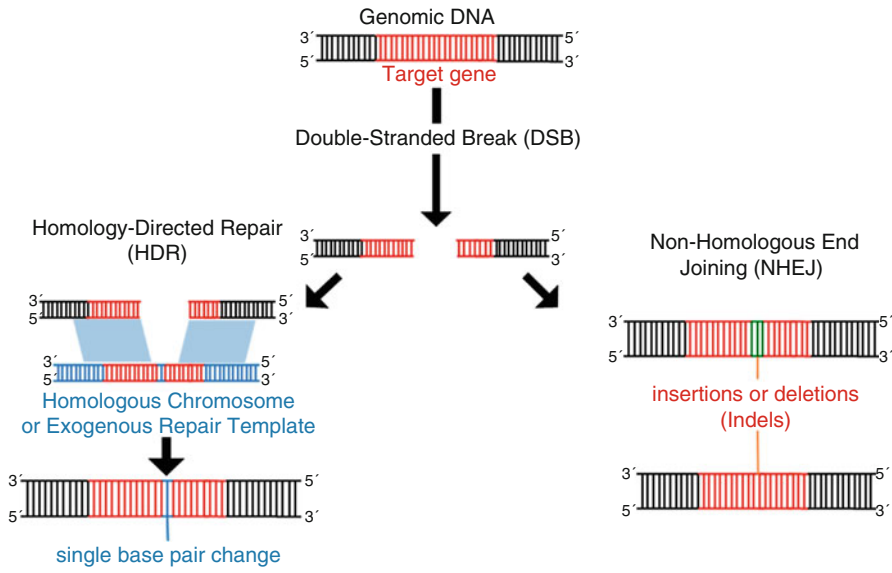


Fig. 2 Double-stranded breaks (DSBs) induce endogenous DNA repair mechanisms. DSBs can be repaired by nonhomologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ often leads to deleterious insertions or deletions (indels), while HDR leads to high-fidelity DNA repair using the homologous chromosome or exogenously introduced DNA as a template

however, none have shown any off-target effects which could depend on the genomic context and the target (Pattanayak et al. 2011; Wang et al. 2015). For this reason, the research is ongoing to find and design new generations of genome editing tools to outperform those existing. Awaiting the next generations, to design gene therapy strategies based on existing genome editing tools, the biosafety level must be carefully evaluated, and an accurate risk/benefit balance must be established.

Another strategy avoiding potential side effects caused by the genomic integration of the vector is to use non-integrating vectors. In this case, the gene encoding the therapeutic factor is not inserted in the genome of the patient but is in a DNA molecule which is extrachromosomic. This type of strategy has a higher level of biosafety but has a restricted scope because the therapeutic factor is, in most of the cases, expressed transiently. Indeed, extrachromosomic DNA molecules are lost during the cellular divisions. Consequently, these vectors are used to design gene therapy strategies needing transient expression of therapeutic factors such as in vaccination (Negri et al. 2007; Tatsis and Ertl 2004) or to transiently promote or inhibit physiologic mechanisms such as angiogenesis (Gounis et al. 2005).

In some cases where the target cells are nondividing cells, there is no dilution of the extrachromosomic DNA molecules, and the expression of the transferred gene can be observed for several years. Non-integrating vectors can be used to induce long-term expression in the muscle (Xiao et al. 1996), brain (Ahmed et al. 2018; Gray et al. 2010; Yáñez-Muñoz et al. 2006), or eyes (Li et al. 1994; Philippe et al. 2006) and address genetic diseases involving these types of structures.

To implement the discussed gene therapy strategies, the key tool is the vector combined or not with a genome editing tool. In this area, the perfect vector does not exist. Each of them has their advantages and limits in terms of targets, efficiency, or level of biosafety, and the choice of the vector is done according the therapeutic context.

Genome Editing and Genic Vectors

In the field of gene therapy, vectors are vehicles for delivering foreign DNA or RNA into patient cells. The vectors used in gene therapy are divided into two distinct categories: synthetic and viral vectors.

In the first case, the nucleic acid of interest is complexed with synthetic molecules to be able to enter in cells. In the second case, the viruses are modified to be non-replicative and non-pathogen. The nucleic acid of interest is included in a virus, replacing its viral genome and allowing its delivery into the cells. These two classes of vectors can be non-integrating, inducing transient modifications in most of the cells or long-term modifications in nondividing cells, or integrating, allowing to permanently modify the genome of patients. The viral vectors integrate, either they naturally have a system of integration or they are combined with an ectopic genome editing system, while synthetic vectors must be combined with a genome editing system to be integrating.

Genome Editing Systems

Genome editing systems allow to permanently modify the genome of cells. There are nonspecific systems such as transposons (Ding et al. 2005; Ivics et al. 1997), which randomly integrate the therapeutic DNA, and specific systems allowing to target or to choose a genomic locus of integration. The most suitable editing systems are systems which allow the choosing of a targeted genome locus to inactive a gene or to insert a therapeutic DNA.

In human cells, the first system used, allowing to efficiently choose the target genomic locus for genetic engineering, is based on ZFN (Urnov et al. 2005). This system combines zinc finger domains able to recognize a DNA sequence, combining with the nuclease FokI, which can induce DNA double-strand breaks (Fig. 3). This system can be combining with synthetic and viral vectors and is very attractive to design gene

therapy. For example, Sangamo Biosciences, which is an American firm, has used ZFNs to provide a therapeutic solution to X-linked severe immunodeficiency syndrome (Urnov et al. 2005). During their investigations, the authors demonstrate the possibility of making gene repair with a ZFN specific mutated il2rγ gene through the stimulation of RH mechanisms. In this way, it is possible to replace the mutated sequence with the wild-type sequence of the gene. However, engineering of ZFN is complex, and Sangamo has locked the technology by filing several major patents in the field. Consequently, research laboratories have a very restrictive access to this technology, avoiding the development of its potential.

At the beginning of 2009, an alternative technology has emerged, the transcription activator-like effector nucleases (TALENs, Fig. 3) (Boch et al. 2009). This genome editing system is easier to design than ZFN, allowing research laboratories themselves to design the tools to develop gene therapy strategies. However, this technology is limited by the very large size of the proteins and the many repeat sequences making its DNA synthesis and its combination with vector systems complex.

Recently a new genome editing tool has emerged which is revolutionizing the field of genome editing and thus gene therapy, the CRISPR system (clustered regularly interspaced short palindromic repeats) (Cong et al. 2013). This system is based on adaptable immune mechanisms used by many bacteria to protect themselves from foreign nucleic acids, such as viruses or plasmids. It is made of two components: a guide RNA (gRNA) which targets a DNA sequence and the Cas9 nuclease allowing a DNA double-strand break (Fig. 3). This system is very easy to design and cheap to produce; any molecular laboratory has the capabilities to produce their own CRISPR/Cas9 systems, democratizing genome editing areas and accelerating research in many fields. Now, a wide range of gene therapy strategies are based on CRISPR/Cas9 system, and several companies have emerged in this area, such as CRISPR Therapeutics or Editas.

All genome editing systems, as efficient and as accurate as they are, require a vector to be delivered into a cell to be able to edit a genome.

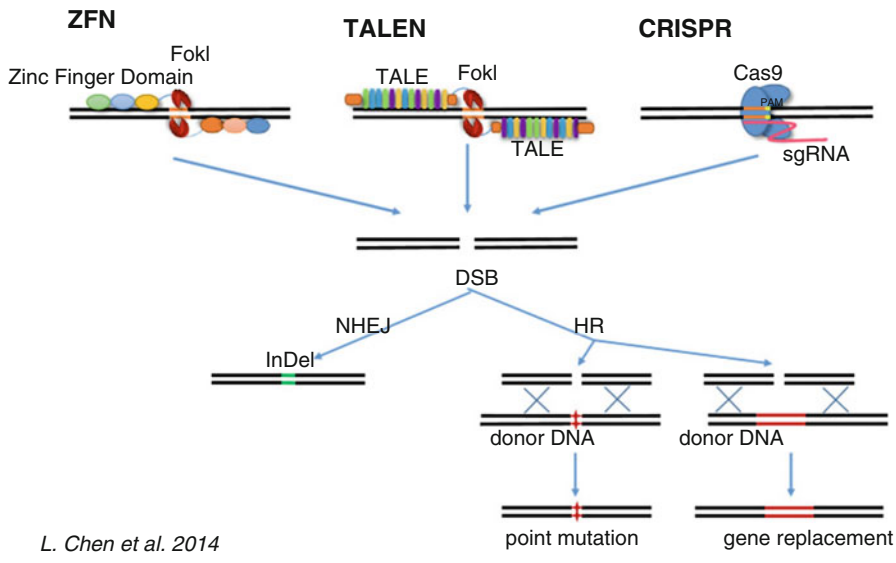


Fig. 3 Mechanism of ZFN, TALEN, and CRISPR/Cas9. ZFN, TALEN, and CRISPR/Cas9 achieve genome modification by inducing targeted DNA double-stranded breaks (DSBs), which would be corrected by NHEJ and HR repair mechanisms. NHEJ-mediated repair leads to the introduction of variable length insertion or deletion. HR-mediated repair could lead to point mutation and gene replacement, in the present of donor DNA

Vector Systems

Synthetic Vectors

Synthetic vectors have the advantage of being easy and inexpensive to produce. In addition, they are not immunogenic which is a definite asset for the development of gene therapies. To allow an optimal transport system, a synthetic vector must have several characteristics and be able to:

- Form a complex with the nucleic acid to vectorize and compact it
- Protect the nucleic acid from various sources of degradation such as nucleases
- Obtain homogeneous synthetic vector/nucleic acid particles of small size (<500 nm)
- Facilitate the fixation of the nucleic acid on the cell because the latter, being negatively charged, cannot adhere spontaneously to the polyanionic plasma membrane of cells
- Promote the passage of the hydrophilic nucleic acid (through the hydrophobic medium constituted by the plasma and/or endosomal membrane
- Allow intracellular trafficking and penetration of the nucleic acid into the nucleus
- Be transfected efficiently in vivo

Currently, three main classes of synthetic vectors are described: peptide vectors, cationic polymer vectors, and lipid vectors.

Peptide Vectors

Cell-penetrating peptides (CPP) are short natural or synthetic peptides that are capable of complexing certain molecules of interest (NA, drugs, proteins) in order to allow or facilitate their access in vitro or in vivo to cells. In most of the cases, the entry into the cell is by endocytosis; the nucleic acids brought then reach the nucleus either by passive diffusion or after cellular mitosis. Currently, the peptide most used as a vector is the HIV transcriptional activation factor, TAT (Green and Loewenstein 1988).

Peptide vectors can be an attractive gene transfer tool due to their ease of production and safety after injection into an organism. However, this technology currently has efficiencies which are far too low to claim ex vivo or in vivo use in therapy.

Cationic Polymer Vectors

The electrostatic attraction between the cationic charges of the polymers and the negative charges of the DNA forms a particulate complex called polyplexes, where the nucleic acids are very strongly condensed, which is able to transfect the cells. Among the natural polymers are HMG1 and chitosan. The first cationic polymer that was found to be effective for gene transfer was Poly-L-lysine. However, it has the disadvantage of requiring the addition of adjuvant to function. Different polymers developed more recently such as nanoparticles, PEIs (polyethylenimines), PAMAMs, or histidylated Poly-L-lysine overcome this limitation. PEIs are currently the most widely used polymers because they are the most effective both in vitro and in vivo. There are many PEIs that differ in their molecular weight and geometry (branched or linear). The most efficient among them are capable of efficiently condensing DNA to lead to very stable polyplexes whose average size is very homogeneous (Campeau et al. 2001). In addition, the large number of primary, secondary, and tertiary amine functions gives them buffering power over a wide pH range. This buffer capacity of PEIs makes it possible to oppose the acidification of endosomes which causes the breakdown of the organelle, thus promoting the release into the cytosol of polyplexes that have been endocytosed by the cells.

Despite good efficacy in vitro and in vivo, the lack of cell targeting produces accumulation in the filter organs (lung, liver, spleen) after systemic injection. Furthermore, excess positive charges of PEI favor nonspecific interactions with the cell membrane, a source of toxicity on which several teams are working to reduce this adverse effect. Another type of synthetic vector has the advantage of being low in toxicity and based on the use of lipids.

Lipid Vectors

Liposomes or lipoplexes are artificial phospholipid vesicles, a size of between 20 nm and 2 μm which can contain numerous soluble molecules water in their aqueous core or insoluble in the hydrophobic compartment of the phospholipid bilayer. The pioneering work of gene delivery using lipids dates to 1987 (Felgner et al. 1987). Early versions of these vectors had limited efficacy in vivo. Indeed, as for the cationic vectors, an accumulation at the level of the filter organs, after systemic injection, was observed. In addition, lipid vectors interacted nonspecifically with different serum components such as albumin, fibronectins, or immunoglobulins (Ogris et al. 1999); aggregates then formed were removed by macrophages (Plank et al. 1996). Meanwhile, a de-complexing phenomenon could also occur by interaction of the complex with anionic macromolecules such as heparin, exposing DNA degradation by nucleases (Xu and Szoka 1996). To solve these problems, the development of liposomes called "stealth" was performed by molecules of polyethylene glycol (PEG) grafted either to lipids or DNA (Monfardini and Veronese 1998; Ogris et al. 1999). In fact, PEGs are hydrophilic and non-immunogenic molecules; therefore covering the surface of the lipoplexes through PEG creates a dense hydrophilic network around the particles which renders them undetectable by the immune system. The morphology adopted by these polymers is variable, and the resulting mesh creates a steric barrier between the particle and the surrounding environment. This steric and hydrophilic masking prevents all hydrophobic or electrostatic interactions resulting in a decrease in the adsorption phenomena or adhesion of blood components on the particles. This specificity limits their capture or degradation and thus increases their bioavailability. As a result, the efficacy of these lipoplexes surpasses that of peptide vectors or cationic polymers.

On the other hand, recent works highlight an interesting possibility offered by this type of vector. The developed strategy exposes an original system based on "PEGylated" immuno-liposomes (PILS). This system developed by the WM. Pardridge's team offers a unique way to reach the brain by simple systemic injection, without any alteration of the blood-brain barrier (BBB). This strategy is based on the encapsulation of plasmid DNA in PEGylated liposomes complexed with monoclonal antibodies that recognize the BBB. The injected particles are then able to recognize the transferrin receptors and are transported to the brain. Much work has been done by the WM. Pardridge's team has demonstrated the interest and effectiveness of this system, particularly in animal models of Parkinson's disease (Pardridge 2005) and Alzheimer's disease (Boado et al. 2007). In addition to the ability to target the brain, this system is an undeniable advance because it allows one to effectively consider targeting a given cell population with the appropriate antibody. Indeed, several studies have established that it is possible to target glioma-forming cells by coupling one PILS with two antibodies, the first to pass the BBB and the second to specifically target retinal cells (Zhang et al. 2003). However, this approach remains limited by expensive production methods and by the great instability of the synthesized particles which lose 50% of their efficiency 24 h after their production.

The major assets of these synthetic vectors reside in their ease, their cost of production (except for the PILS), as well as in their biosecurity. However, these benefits can only be enhanced if these systems achieve the required efficiency for a given application. In most strategies requiring nucleic acids transfer, these vectors do not yet show satisfactory efficacy to make them candidates. In addition, it is complicated to target only the desired cells which decreases their potential as a gene vector. Vectors derived from viruses, however, exhibit a high efficiency in delivering a DNA sequence to cells both in vitro and in vivo. In addition, some of them have pseudotyping flexibility that makes them specific to a given cell population.

Viral Vectors

The cycle of a virus consists of bringing its genetic material into a cell and diverting its machinery to produce its own proteins and thus replicate itself.

For many years, various works have aimed to modify these viruses to make them vectors of gene transfer in vitro but also in vivo. More or less free of initial virulence factors, viral vectors are generally unable to replicate but allow the delivery of nucleic acids of interest in cells. These viral vectorization methods are generally more efficient than conventional methods of transfection of naked or complexed DNA (certain molecules, in particular in vivo). Indeed, they take advantage of all strategies developed by viruses to penetrate the most effective cell and advantageously divert its machinery. Among the viruses used for gene transfer are adeno-associated virus (AAV), herpes simplex virus (HSV), adenovirus, and retrovirus.

Adenoviral Vectors

Adenoviruses are 36kb linear double-stranded DNA viruses and devoid of envelopes but surrounded by an icosahedral capsid. The adenoviral vectors used as gene transfer vectors are generally derived from group C serotype 2 or 5 adenoviruses and can be produced easily. The genome of this virus remains extrachromosomal in linear form, making it a platform for genome editing tools or a good candidate to induce transient factor expressions (Alonso-Padilla et al. 2016). In order to take advantage of the high transduction efficiency offered by this type of vector, different teams have worked to develop strategies for integration into the genome of the target cell, and they showed promising results in terms of efficiency and specificity of genome engineering. While the effectiveness and specificity of these systems are undeniable assets for gene therapy, there are other aspects to consider before imagining clinic applications. Indeed, there are two main limitations related to the adenoviral vector itself. The first is the fact that its use may be compromised by the presence of neutralizing antibodies produced following a primary infection with a wild-type adenovirus or a first therapeutic adenoviral vector injection. The second is about the vector production techniques that require the use of an auxiliary vector to provide in trans the functions necessary for the replication of the virus. It is very difficult not to observe contamination of the vector batch produced by this auxiliary vector (of the order of $0.01 \pm 0.1\%$).

Adeno-Associated Vectors (AAV)

AAV is a small (20–25 nm), non-enveloped, single-stranded, linear parvovirus of 4.7 kb. Its genome encodes three capsid proteins (Cap) and four replication proteins (Rep). These regions are framed by sequences called ITRs (inverted terminal repeat), which are needed in cis for the replication and assembly of viral particles. This virus requires the help of a helper virus such as adenovirus, herpesvirus, or vaccinia virus to replicate. In the absence of a helper virus, wild-type AAV preferentially integrates at a specific locus of human chromosome 19 at the q13.3-qter position through two proteins encoded by the rep gene.

The first generation of vectors has retained the ability to integrate into chromosome 19; however its weak cloning ability has limited its use. The suppression of all the viral sequences, except for the ITRs, made it possible to increase this cloning capacity (up to 4.6 kb) but also to overcome the integrative and inhibitory effect on the gene expression caused by rep. However, the question of the integration of this vector is always controversial. In some cell types such as epithelial or muscle cells, it seems to remain mainly in the form of episomal concatemers. However, it has been shown that these vectors may integrate randomly into the genome of some proliferating cells such as Hela, 293T, and CD34 + or quiescent such as neuronal cells.

Several strategies have been tested to control the integration of this vector, exploiting the linear phenotype of the virus to induce passive RH. Some of them have been tested in vitro to restore the function of the gene or to modify the human gene. Nevertheless, the proportion of nonspecific integrations observed is still high (near 10%). Although this tool has a specific integration efficiency that may be satisfactory for certain applications, the proportion of nonspecific events is much too high to make it a secure gene therapy tool.

In order to increase the potential in gene editing of the AAV, strategies have been designed to produce a double-strand break within the genome of the cells to be modified in order to stimulate the HR mechanisms or to inactive a gene. In this context, the specific cleavage was initiated by genome editing tools, without DNA template to inactivate a gene or with a DNA template to induce RH and integrate a DNA sequence. However, the use of AAV in this context is sorely limited by its cargo capacity given that genome editing tools are composed of large cDNA.

On the other hand, AAV is largely used for stable expression in nondividing cells to design strategy for neurodegenerative diseases, retinal degeneration, or muscular atrophies. The results of several clinical studies are encouraging (Wang et al. 2019b); however, the biosafety and the balance risk benefits need to be well evaluated because the studies highlight genotoxic effects of AAV injection, inducing in some cases tumor cells (Donsante et al. 2007). This data, although still contested today, requires the utmost caution when designing therapy with this type of vector. Furthermore, its use may be compromised by the presence of neutralizing antibodies produced by a first therapeutic AAV vector injection. Consequently, therapeutic strategy has to take into consideration that the patient could probably have a unique injection of the therapeutic vector.

Retroviral Vectors

Retroviruses are a family of enveloped viruses with negative RNA. Their common feature is the use of an enzyme, the reverse transcriptase that allows them to convert their RNA genome into DNA in the cells they infect. In DNA form, the viral genome is stably integrated into the genome of the host cell by a second viral enzyme, integrase. Once integrated, the viral genome can then be transmitted to the offspring of the cell. Retroviruses can be divided into three groups: oncogenic retroviruses, lentiviruses, and spumaviruses. Oncogenic retroviruses, such as murine leukemia virus (MLV), have a so-called simple genome that includes three open reading frames, gag, pol, and env, encoding the structural proteins and enzymes required for the replication cycle. On the other hand, lentiviruses such as human immunodeficiency virus (HIV) as well as spumaviruses like human foamy virus (HFV) are called "complex" because they possess, in addition to the three elements mentioned, genes for regulation and accessory genes such as tat, rev, nef, vif, and vpr for the first and vpu for the second. In addition to the coding sequences, the retrovirus genome also contains several sequences that act in cis. Among them, the LTRs, located at both ends of the viral genome, contain the signals of initiation and termination of transcription; the psi sequence corresponds to the encapsidation signal of the RNA genome. Some of these sequences occur during reverse transcription, such as primer binding site (PBS) and polypurine tract (PPT). Lentiviruses also contain a central polypurine tract (cPPT) and a central termination sequence (CTS), both of which allow a unique mode of lentivirus-specific reverse transcription and appear to be involved in the nuclear import of the lentivirus pre-integration complex.

The knowledge built up over years of the biology of these viruses has made them powerful tools for gene transfer. Indeed, the vectors that derive from retroviruses have a high transduction efficiency as well as easy production and which is free from the use of vector helper source of contamination. In addition, their integrative phenotype offers the possibility to edit the genomes without ectopic system, preserving the modification of the genome over the cellular divisions, making them the tools of choice for gene therapy targeting dividing cells (Maetzig et al. 2011).

The first retrovirus use in gene therapy has been derived from MLV. Although this type of vector has shown a high efficiency to modify patient dividing cells, the MLV vector has some limitations such as their ineffectiveness to transduce nondividing cells. Furthermore, their uncontrolled integration is largely in the regions of initiation of transcription, thereby deregulating the transcriptome of the cell. This feature can have dramatic consequences when setting up gene therapy based on the use of this vector. Indeed,

the clinical trial of the team of A. Fischer to treat children with severe X-linked immunodeficiency syndrome has been qualified by severe side effects induced by the genotoxic effects of the vector; this point is described in the part III (Hacein-Bey-Abina et al. 2003). In addition, the sensitivity of this vector to the methylation mechanisms described in several studies is also problematic in gene therapy. Indeed, this sensitivity exposes the transgene to silencing effects that can cause its partial extinction, or sometimes even total, and thus compromise the therapeutic strategy implemented. Consequently, MLV is less and less used to design gene therapy strategies in favor of another vector also derived from retrovirus and showing almost all advantages comparing to MLV: the lentiviral vectors derived from HIV-1.

Lentiviral Vectors Derived from HIV-1

Lentiviral vectors appear to be more attractive because they can transduce the cells into division arrest (Naldini et al. 1996). Moreover, although they integrate preferentially into active areas for transcription, integration events appear to be less subject to position effects and induce less genotoxic effects than MLV. Then lentiviral vectors can be producing with a very high concentration and efficiency (>1E9 TU/mL), and they have a large cargo capacity allowing to deliver a wide range therapeutic transgene and, in some strategies, several of them.

However, the main limitation of these vectors is, as for the MLV, their uncontrolled integration even the lentiviral vector genotoxic potential is lower.

In order to overcome this problem, advances in the biology of the HIV virus have made it possible to consider modifying the viral integrase to make it specific to a desired site. Several strategies have been tested but, for now, none has given acceptable results to go ahead in therapy. Another way explored is to inactivate the wild-type integrase of the vector either to design a therapeutic strategy based on non-integrating vector with the advantages of lentiviruses or to replace it by an ectopic genome editing tools such as ZFN or CRISPR. This generation of vector is very attractive because, as the integrating version, it can be producing with very high concentration to be very efficient; it has a large cargo capacity and offers a wide range of design to target specific cell populations without the side effects due to the uncontrollable integration. However, for the moment, all the types of integrase mutant tested have shown a background of integration, in the range of 0.1–0.5%. This limit has to be taken into consideration in the designing of gene therapy strategies.

Recently, a new generation of lentiviral vectors has emerged with the particularity of being non-integrating without background and having an unprecedented level of biosafety (Sarkis et al. 2014). This new generation of lentiviral vector is mutated for the reverse transcription, and then the RNA genome of the vector cannot be converted in DNA and is taken charge by the cells as a mRNA. This type of vector induces only transient expression in all type of cells (dividing and non-dividing) and is perfectly suitable to express genome editing tools or to design a strategy needing a transient expression of a therapeutic factor, such as in the field of vaccination, vascularization, or oncology. Although RNA lentiviral vector keeps several advantages of the integrating version, such as large cargo capacity or high flexibility of design, its use is limited by its efficiency which is too low for some applications. Now, several ways are being explored to overcome this limit, and it is probable that this new generation of vectors offering new standards in terms of biosafety will be the basis for a wide range of the next gene therapy strategies.

Clinical Trials in Gene Therapy

Seminal Trials

The concept of gene therapy is already old, since it was born in the early 1970s when the scientists Rogers and then Friedmann and Roblin raised the possibility of using exogenous DNA to replace defective DNA in people affected by genetic defects (Friedmann and Roblin 1972). This idea materialized experimentally in 1989 by French

Anderson. He conducted the first human gene therapy safety test where a harmless marker was injected into a 52-year-old man. The test succeeded, and the following year he treated a 4-year-old girl, Ashanti DeSilva, who was suffering from an immune disorder caused by a defective ADA gene (Blaese et al. 1995). At the time the only solution was to treat the patient with an artificial PEG-ADA supplement, though F. Anderson set out to make a more permanent cure. The strategy has been to take T cells out the body and to genetically modify them by the introduction of a healthy functioning copy of the ADA gene. Currently DeSilva is alive and well but still taking her PEG-ADA supplements; the issue is still not clear as to whether it was the gene therapy treatment carried out by F. Anderson or the supplement itself that has been keeping her going all these years. The trial received a mixed reception, but nonetheless it was a shot in the arm for gene therapy enthusiasts all over the world. However, in 1999, the death of a patient, Jesse Gelsinger, 18 years old and participating in a gene therapy trial, triggered a huge reaction (Science 2000). Since then, the sector has alternated between enthusiasm and reserve for gene therapy.

Jesse Gelsinger joined a clinical trial which included 18 patients who suffered from ornithine transcarbamylase (OTC) deficiency, an X-linked genetic disease of the liver induced by a dysfunctional OTC gene, the symptoms of which include an inability to metabolize ammonia. This clinical trial has been derived by the Dr. James M. Wilson's team of the Pennsylvania University. The protocol validated by the FDA consisted of a systemic injection of an adenoviral vector containing a healthy version of the OTC gene. Before the vector injection, the control results of Jesse Gelsinger indicated that he had neutralizing antibodies to adenovirus, and thus CD4 T cells activated against it as well. These results were satisfactory for continuing the study on him. Twelve hours after injection, Jesse Gelsinger had the same symptoms as all subjects: fever, myalgia, and biochemical consequences such as anemia and thrombocytopenia. However, 18 h after the injection, he began to show signs of mental deterioration and jaundice that the other 17 subjects did not have. As the hours went by, his condition worsened, doctors had to place him under artificial respiration to control his air since he found himself in compensated respiratory alkalosis caused by a high level of ammonia in the blood. Similarly, to lower his ammonia, they did a hemodialysis; however, all these organs worsened, starting with the lungs. Finally, 98 h after the injection, Jesse Gelsinger died. The findings from his autopsy reveal that Jesse died as a result of systemic activation of his innate immune system. A very high concentration of the vectors has been detected in the liver but also in organs such as lymph nodes, bone marrow, and spleen. Pennsylvania University highlighted the need for further research on gene therapy to understand and promote the success of future trials to treat this disease.

It took until 2000 to observe therapeutic effects in the field of gene therapy with Alain Fischer at the Necker Hospital for Sick Children in Paris who has led the SCID-X (severe combined immunodeficiency X-linked*)* gene therapy trial (Cavazzana-Calvo et al. 2000). This clinical trial involved children who presented X-linked SCID which is a faulty copy of a gene on the X chromosome that makes the immune protein interleukin-2. As a result, they have no resistance to infection, and they must live in a sterile environment to not die. A total of 15 patients, have been treated so far – 11 in Paris and 4 in London. The gene therapy strategy has been to modify ex vivo the HSC of the children with an MLV vector carrying the healthy version of the SCID gene and to reinject them. For the first time, the trial was a success as all the children showed a significant health improvement. However, the second time five children developed leukemia and one of them died. The clinical trial was stopped in an emergency, and the analyses showed that leukemia was induced by the insertion of the MLV vector near the LIM domain-only 2 (LMO2) proto-oncogene inducing its deregulation (Hacein-Bey-Abina et al. 2003). This dramatic side effects show that gene therapy needs to be optimized, involving vector design, and the balance risk/benefit must be accurately

evaluated. Since this failure, other trials for SCID-X have started taking these points into consideration.

This proof of concept has opened the way for several clinical trials all around the world (Fig. 4) using different vectors (Fig. 5), to treat different diseases (Fig. 6) and in different phases (Fig. 7). Recently, for some of them, a gene therapy product has obtained a marketing authorization (Table 1).

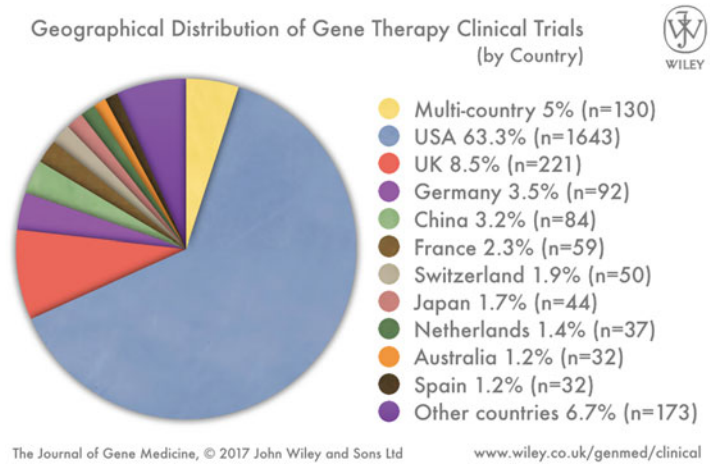


Fig. 4 Geographical distribution of gene therapy clinical trials (by country)

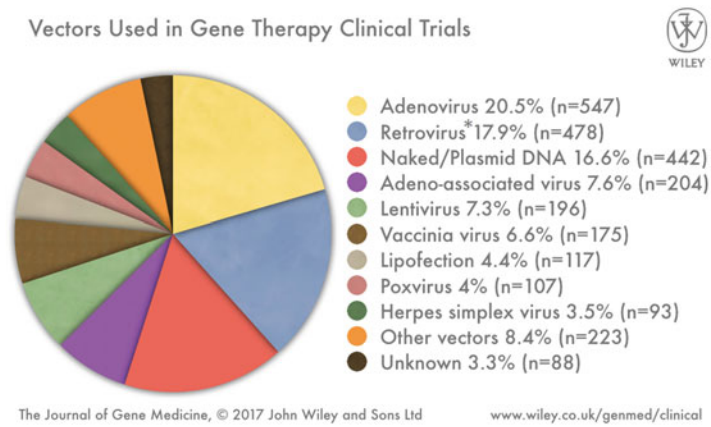


Fig. 5 Vectors used in gene therapy clinical trials. ∗ Other than lentivirus

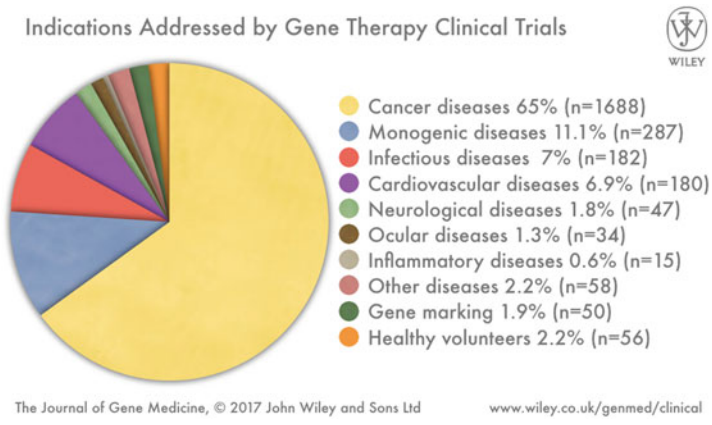


Fig. 6 Indications addressed by gene therapy clinical trials

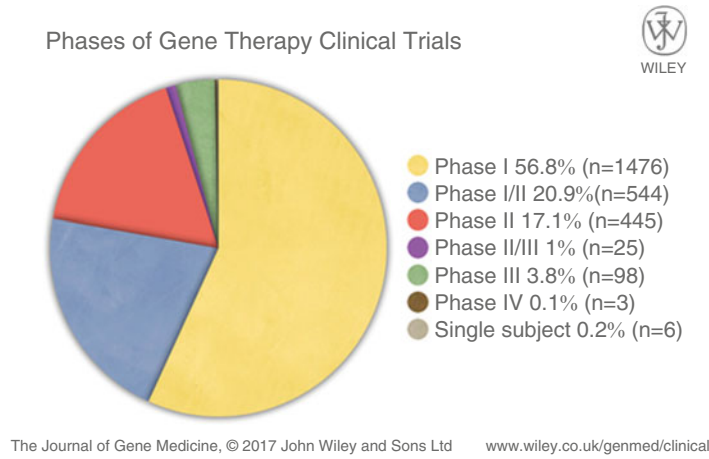


Fig. 7 Phases of gene therapy clinical trials

Table 1 Approved gene therapy products

Name	Date of approval	First Approving agency	Disease	Type of vector used	Manufacturer
Gendicine	October 2003	State Food and Drug Administration of China	Head and neck squamous cell carcinoma	Adenoviral Vector	Shenzhen SiBiono GeneTech (Shenzhen, China)
Glybera®	November 2012	European Marketing Authorization (EMA)	Lipoprotein lipase deficiency	AAV1	UniQure (Amsterdam, Netherlands)
Strimvelis™	June 2016	EMA	Adenosine deaminase deficiency (ADA-SCID)	Lentiviral Vector	GlaxoSmithKline (Middlesex, United Kingdom)
Kymriah™	August 2017	FDA	Acute lymphoblastic leukaemia	Lentiviral Vector	Novartis Pharmaceuticals (Basel, Switzerland)
Yescarta™	October 2017	FDA	B-cell lymphoma	Lentiviral Vector	Kite Pharma, Incorporated (Santa Monica, California, USA)
Luxturna™	December 2017	FDA	Retinal dystrophy (biallelic RPE65 mutation)	AAV2	Spark Therapeutics, Inc. (Philadelphia, Pennsylvania, USA)
Zolgensma™	May 2019	FDA	Spinal muscular atrophy	AAV9	Novartis Pharmaceuticals (Basel, Switzerland)

Type of Diseases Treated by Gene Therapy

Neuromuscular Diseases

Inherited neuromuscular disorders encompass a broad group of genetic conditions, and the discovery of these underlying genes has expanded greatly in the past three decades. The discovery of such genes has enabled more precise diagnosis of these disorders and the development of specific gene therapeutic approaches. Duchenne muscular dystrophy and spinal muscular atrophy are the most common debilitating neuromuscular disorders affecting children. The development of such

therapeutic approaches can be categorized into two broad strategies. The first strategy, initiated in the late 2000s, relates to the correction of mutant RNA processing, using either antisense oligonucleotides or small molecules that can modify mutant RNA splicing (Lu et al. 2003; Magee et al. 2006; Benchaouir et al. 2007; Kinali et al. 2009). The second strategy involves the use of AAVs to deliver a functional or partially functional gene copy to the affected cells and tissues. Such translational research has led to the approval of two genetic therapies by the US Food and Drug Administration: eteplirsen for Duchenne muscular dystrophy (Railroading at the FDA 2016) and nusinersen for spinal muscular atrophy, which are both antisense oligonucleotides that modify pre-mRNA splicing (Khorkova and Wahlestedt 2017). Very recently, in May 2019, Novartis received US approval for its spinal muscular atrophy gene therapy Zolgensma which provides through AAV infusion a normal copy of the SMN1 gene to babies born with a defective gene. The one-time treatment costs $2.1M

Ocular Diseases

The eye is an easily accessible, highly compartmentalized, and immune-privileged organ that offers unique advantages as a gene therapy target. Proof-of-concept for vector-based gene therapies has also been established in several experimental models of human ocular diseases involving photoreceptor, ganglion cell, and optic nerve degeneration. Gene therapy has also been explored as a potential therapy for red-green color blindness, corneal neovascularization, and a variety of other corneal diseases, including allograft rejection, optic nerve trauma, autoimmune uveitis, and melanoma (Williams et al. 2017; Cavalieri et al. 2018; Sahel and Dalkara 2019).

In 2017, the US Food and Drug Administration allowed the first gene therapy to be placed on the market, Luxturna, which consists of direct injection of the vector into the patient's body to treat Leber congenital amaurosis which is an inherited retinal disease (Darrow 2019). The vector used derives from the AAV and contains a healthy copy of the RPE65 gene. The cost of the treatment is $425K per eyes. Following this first authorization, this approach is also tested in several other monogenic ophthalmic diseases such as Leber's optic neuropathy or Stargardt's disease.

Blood Diseases

Blood disorders comprise of a wide range of diseases involving red blood cell and iron disorders, white blood cell disorders, bone marrow failure syndromes, thrombosis and anticoagulation disorders, bleeding disorders, and autoimmune blood cell disorders. Many of them are inherited disorders involving genetic dysfunction. Gene therapy holds great potential as a cure for blood, which has limited treatment options (Kohn 2019).

As discussed previously, gene therapy strategies are designed to address several immune diseases SCID-X, but there are also other immune disorders for which gene therapy strategies have been designed. There is, for example, ADA-DICS, a severe immunodeficiency characterized by the absence of the ADA protein, necessary to produce lymphocytes (Aiuti et al. 2002, 2009). Results are also encouraging for the treatment of Wiskott-Aldrich syndrome (WAS) (Aiuti et al. 2013). This disease, resulting from a mutation on the WAS gene, is characterized by dysfunctions of the blood cells and the absence of platelets. Trials conducted in Europe and the USA confirm the effectiveness of the approach to improve the health status of those treated, including adults. Approaches targeting other immunodeficiencies are currently being tested, including X-linked chronic septic granulomatous disease. In 2016, Strimvelis™ was the first ex vivo gene therapy approved for the treatment of ADA-SCID (Stirnadel-Farrant et al. 2018). This therapy was evaluated in 18 patients who are still alive with no major side effects, but the price of the treatment is about $700K.

Hemoglobin disorders are inherited diseases, such as thalassemia and sickle cell anemia, and are the most common monogenic disorders worldwide, causing significant morbidity, mortality, and healthcare expenditures. Gene therapy strategy for β-globinopathies is to add a healthy copy of the β- or γ-globin

genes and elements of the locus control region in hematopoietic stem cells (HSCs) with a lentiviral vector (Perumbeti and Malik 2010). In contrast to immune disease where a relatively small proportion of HSCs and a relatively small amount of transgene protein expression are adequate to produce a phenotypic correction, β-globinopathies requires high amounts of gene expression in individual red blood cells. Furthermore, engraftment of a large proportion of gene-corrected HSCs that repopulate the bone marrow is necessary in order to affect significant clinical effects. Therefore, the clinical trials in this field use the lentiviral vector which currently offers potential for delivering high expression and transducing a high proportion of HSCs, allowing genetic therapy approaches for β-globinopathies.

In France, a 13-year-old boy became the first sickle cell disease patient to be treated with gene therapy. In October 2014, the boy received LentiGlobin BB305 gene therapy (Thompson et al. 2018). After receiving the modified stem cells, the boy kept getting red blood cell transfusions until researchers could measure adequate levels of the modified hemoglobin. The patient received the last transfusion on day 88 after the cell transfer. Six months after the procedure, his total hemoglobin levels were stable. He had some side effects from the chemotherapy used to get rid of his blood cells, but all issues seem resolved, and the gene therapy itself has not, as of yet, caused any side effects. However, caution should be observed because several past clinical trials have already shown late side effects.

Hemophilia A and B are X-linked bleeding disorders, respectively, caused by mutations in the gene coding for FVIII or gene coding for FIX, resulting in deficient and/or defective coagulation, leading to spontaneous or traumatic bleeding into joints, muscles, or body cavities (Perrin et al. 2019). Gene therapy strategies have the aim of bringing the healthy version of the gene to hepatocytes which are the cells producing FVIII and FIX. At the beginning, for most clinical trials, AAV vector was chosen for its high efficiency to modify hepatocytes and non-integrating feature which makes a safer vector than those derived from retroviruses. However, some studies have shown that in some cases, AAV vectors can be integrated, and, in the meantime, several studies have upgraded the design of lentiviral vectors to be safer. Furthermore, pre-existing neutralizing antibodies to AAV following natural exposure to the wild-type virus may inhibit gene transfer with AAV vectors restricting its use. For these reasons, more and more gene therapy strategies based on lentiviral vectors are being designed.

Fanconi anemia is an inherited bone marrow failure disorder characterized by aplastic anemia and an enhanced risk for the development of leukemia. The syndrome may occur as a consequence of a defect in 1 out of at least 15 genes. The development of gene therapy approaches for this disorder has focused on the protein encoded by the FANCA gene, 95,96 which is most commonly mutated. An international working group has been established to chart the path and facilitate the development of gene therapy for Fanconi anemia. 95,96 gene therapy for patients with Fanconi anemia is particularly challenging because of low numbers of hematopoietic stem cells and sensitivity to myelosuppressive regimens.

Metachromatic leukodystrophy (MLD) is an inherited autosomal recessive disorder secondary to a deficiency of the lysosomal enzyme arylsulfatase A. Massive accumulation of non-metabolite sulfatides damages both the central and peripheral nervous systems. A mouse knockout model of MLD has been developed and used for the exploration of gene therapy approaches (Hess et al. 1996). Stem cell-targeted gene transfer followed by autologous transplantation is one approach (Matzner et al. 2002). An alternative that is also being explored in the mouse model is the direct injection of AAV vectors encoding ARSA into the CNS (Sevin et al. 2006). Early correction is essential because the most common form of MLD develops in the second year of life with rapid, progressive CNS dysfunction. Thus, the direct introduction of AAV vectors into the brain seems preferable than stem cell-targeted gene transfer in that the correction of the phenotype with the latter approach requires several

months, during which the patients continue to deteriorate.

Neurodegenerative Diseases

Neurodegenerative disease processes often involve the progressive accumulation of dysfunctional proteins within cells, leading to cell death. For now, there are no existing therapies that correct underlying neurodegenerative disease processes; current therapies provide merely symptomatic relief. Gene therapy is a promising strategy by altering or inducing the expression of specific proteins for neuroprotection, neurorestoration, and, ultimately, correction of the underlying pathogenic mechanism (Deverman et al. 2018).

For example, in adrenoleukodystrophy, a demyelinating disease of the central nervous system, patients' blood stem cells are corrected ex vivo with a lentivirus and then reinjected. Twenty nine patients were treated, and the results show a stabilization or improvement of their condition in most cases (Duncan et al. 2019). Promising results have also been obtained with a similar approach in children with metachromatic leukodystrophy (Sessa et al. 2016), and other work is underway in Sanfilippo's disease. A French trial involving four children with this disease is ongoing. The gene therapy strategy was to inject AAV vector into different areas of the patient's brain to induce the production of the missing enzyme by the brain cells (Tardieu et al. 2017). No noticeable side effects were noted during the 30 months following treatment, and an improvement in intellectual and behavioral development was observed, paving the way for a phase III trial.

Gene therapy approaches are also designed to address other more common neurological diseases such as Parkinson's disease or Alzheimer's disease.

Skin Diseases

There is a wide range of skin diseases involving gene malfunctioning such as epidermolysis bullosa (EB), pachyonychia congenita, melanoma, ichthyosis, xeroderma pigmentosum, wound healing, or Netherton syndrome. Junctional epidermolysis bullosa (JEB) is an EB disease for which there are the most advanced results. In 2017, a 7-year-old child suffering from this rare and serious disease received several autologous transplants of skin cells genetically modified with a retrovirus vector to correct the mutation causing the disease, affecting the LAMB3 gene (Hirsch et al. 2017). This mutation prevents the junction between the dermis and the epidermis. Faced with a vital emergency, German and Italian doctors performed this procedure successfully on 80% of the body surface. In another form of the disease, a clinical trial is underway in the USA to produce genetically corrected epidermal leaflets to treat the skin lesions that appear in dystrophic epidermal necrolysis.

Cardiovascular Diseases

In the cardiovascular field, researchers are developing gene therapies to promote the regeneration of vascular tissues in the case of arterial ischemia. Other strategies aim to fight against restenosis (narrowing of an artery occurring after stent implantation), using genes coding for proteins that slow down or, on the contrary, stimulate these processes. No significant results have been reported in the treatment of chronic heart failure with an adeno-associated vector encoding Serca2; however, efforts are underway to improve the cardiac tropism of the vectors (Greenberg et al. 2016).

Cancer

Research in gene therapy for cancer is currently focused in multiple areas, including genetically engineered viruses that directly kill cancer cells, gene transfer to offset the abnormal functioning of cancer cells, and immunotherapy, which stimulate the immune system to recognize and kill tumor cells.

Oncolytic viruses are defined as genetically engineered (or in some cases naturally) occurring viruses that selectively replicate in and kill cancer cells without harming the normal tissues (Bommareddy et al. 2018). In addition to the direct lysis of tumor cells, viral infection can induce acute devascularization of the tumor,

causing the formation of large necrotic regions, such as those created by antiangiogenic agents. In addition, tumor infection by oncolytic viruses stimulates the antitumor immune response, which in turn contributes to the elimination of even uninfected tumor cells. Among the most currently studied oncolytic viruses are herpes simplex virus 1 (HSV-1), vaccinia virus (VV), serotype 5 adenovirus, serotype 3 reovirus, Newcastle disease virus (NDV), measles virus, and vesicular stomatitis virus (VSV).

Directing the virus to the tumor in sufficient quantity remains a major challenge for effective oncolytic virotherapy in a clinical setting. In addition, exploiting the antitumor immune response induced by oncolytic viruses while limiting the antiviral response remains another obstacle to overcome. New generations of oncolytic viruses including transgene are developed to be more aggressive and to induce better anticancer activity. Linked to these innovative developments, the marriage of virotherapy and chemotherapy leads to better control of the efficacy of these biotherapeutic agents. These developments point to a promising future for this multimodal therapeutic platform, with remarkable flexibility and specificity. As proof, several oncolytic viruses are in the advanced phase of clinical evaluation and could be approved soon for the treatment of certain tumors (Ribas et al. 2017; Hirooka et al. 2018).

In the same time, other strategies based on gene transfer to offset the abnormal functioning of cancer cells are explored. Although most cancers harbor multiple oncogenic mutations, accumulating preclinical and clinical data now supports the fact that many cancers are sensitive to inhibition of single oncogenes, a concept named oncogene addiction. Consequently, one way to restore the functions of tumor suppressors is to reintroduce wild-type tumor suppressor gene into target cancer cells for expression such as TP53, PTEN, or BRCA1 with a vector.

Given the high-mutation frequency of the tumor suppressor p53 in human cancers, restoration of wild-type p53 function is the most explored strategy in this approach of cancer therapy. The first vector used in this aim has been the adenoviral vector (Adp53). Preclinical studies have shown that Adp53 induces tumor regression in various cancers, including head and neck cancer, colorectal cancer, lung cancer, ovarian cancer, bladder cancer, and prostate cancer (Roth et al. 1998; Zhang et al. 2005). In 2003, the outcome of these promising results was the first commercialization of the world's first gene therapy product, the Gendicine, approved by the State Food and Drug Administration of China for the treatment of head and neck squamous cell carcinoma (Wilson 2005). Despite its use for 15 years in more than 30,000 patients and several studies showing a good safety and a significantly better responses compared to standard therapies, SCH-58500 and Advexin, which are the versions developed in the USA, have not been approved by the FDA, owing to their ineffectiveness (Guo and Xin 2006). No Adp53 therapies have been approved in the USA and Europe for now.

Although the strategies based on oncogene addiction are attractive, the principal weakness of using these approaches is that they do not address the problem of cancer progression as selected by the recessive phenotypes of genetic instability and apoptotic resistance that arise from loss-of-function defects of tumor suppressors.

Other strategies that could answer these issues are based on the immunotherapy. The immune system interacts closely with tumors during the disease development and progression to metastasis. The complex communication between the immune system and the tumor cells can prevent or promote tumor growth. Based on this observation, several strategies of immunotherapy cancer are born. The principle is to enable the immune system to recognize tumor cells to kill them. Initial strategies did not involve gene therapy factor; they used the T-cell growth factor interleukin 2 (IL-2) or monoclonal antibodies to empower the immune system in killing cancer cells. Although encouraging results have been observed is some cases, these strategies have several limitations and specifically the accessibility to the tumor.

Cancer immunotherapy has been greatly upgraded by its combination with gene therapy allowing to genetically modify immune cells such

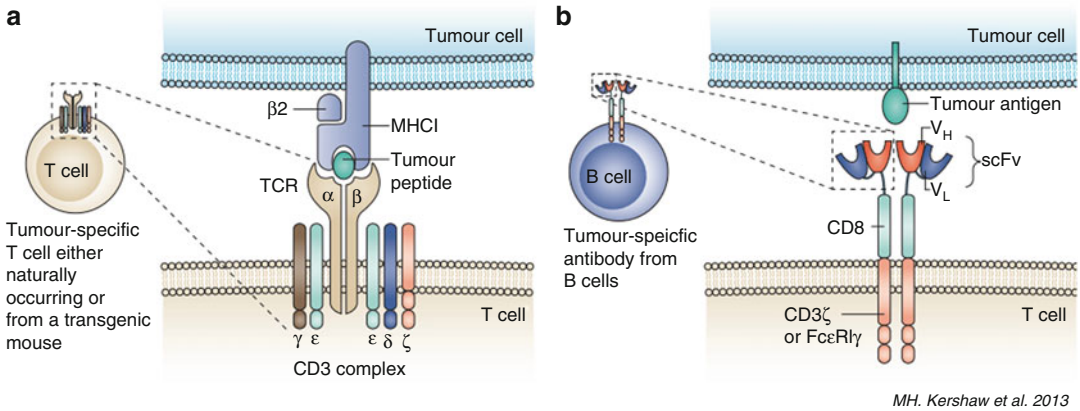


Fig. 8 Derivation of TCRs and CARs for the genetic modification of T cells. (**a**) T-cell receptor (TCR) genes, made up of α- and β-chains, can be derived from tumor-specific T cells, which can naturally occur in humans, or from the immunization of human leukocyte antigen (HLA)-transgenic mice. Alternatively, they can be derived from screening bacteriophage libraries of antibodies. The α- and β-chains associate with the γ-, δ-, ε-, and ζ-chains of the CD3 complex. When the TCR encounters a processed tumor antigen peptide fragment displayed on the major histocompatibility complex (MHC) of the tumor cell, phosphorylation of immunoreceptor tyrosine-base activation motifs (ITAMs) occurs, leading to a cascade of intracellular signaling that results in the release of cytokines and cytotoxic compounds from T cells. (**b**) Chimeric antigen receptors (CARs) are composed of a single-chain antibody variable fragment (scFv) extracellular domain linked through hinge and transmembrane domains to a cytoplasmic signaling region. Genes encoding the scFv are derived from a B cell that produces a tumor-specific antibody. An scFv is shown linked by a CD8 hinge to transmembrane cytoplasmic signaling regions derived from CD3ζ. CARs usually exist as a dimer, and they recognize tumor antigen directly (with no requirement for MHC) on the surface of a tumor cell. *MHCI* MHC class I

as T cells. First strategies have been based on ex vivo gene transfer of T-cell receptor (TCR, Fig. 8). After identification of a TCR specific to the tumor patient, their T cells are collected, modified with a lentiviral vector containing the TCR, and reinjected into the patient. In this way, modified T cells are able to recognize tumor cells to kill them. This type of strategy appeared very efficient in some cases but was often limited by the fact that TCR is constrained by the HLA restriction. Consequently, TCR immunotherapy fits only with patients who express the particular HLA type (similar to organ or bone marrow transplantation). In addition, tumors can lose their antigen by downregulation of HLA.

Recently a new strategy based on the same principle as TCR but non-HLA restricted is revolutionizing the field of the cancer immunotherapy. This type of strategy used chimeric antigen receptors (CARs) which consist of a tumor antigen-binding domain of a single-chain antibody (scFv) fused to intracellular signaling domains capable of activating T cells upon antigen stimulation (Fig. 8) (Curran et al. 2012). In the same way as the strategy used TCR, T cells are removed from the patient, genetically modify with a lentiviral vector containing the appropriate CAR, and reinjected to the patient. Thus, the modified T cells recognize cancer cells and kill them, independently of the HLA restriction. CAR T-cell strategy rapidly showed very promising results for lymphoma with moderate side effects such as cytokine release syndrome (CRS), neurologic events (encephalopathy, confusion, aphasia, or agitation), and low white or red blood cell count. Most of the side effects can be managed with drugs or resolved on their own without the need for treatment. CRS is the most common and severe side effect, but the improvement of the clinical protocol highly reduces it. Consequently, two products, Kymriah and Yescarta, have been rapidly approved for commercialization in the USA in 2017 and in Europe in 2018, but they are expensive, about $400K (Sheridan

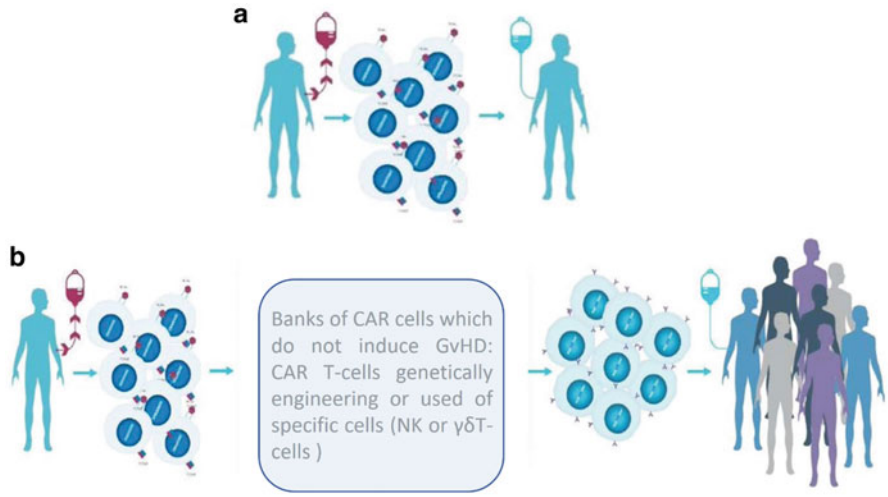


Fig. 9 Autologous versus allogenic CAR cells. (**a**) Autologous strategy: Cells are collected from the patient, genetically modified to express CAR, and reinjected into the patient. Donor and recipient are the same person. (**b**) Allogenic strategy: Cells are collected from donors, genetically modified to express CAR, and used to establish off-the-shelf CAR products, usable to treat patients. Donor and recipient are different

2017, FDA 2018). This is due to the production process which is complex and needs lentiviral vectors produced in GMP grade and requires the use of the patient cells (autologous context, Fig. 9a), avoiding the generation of T-cell universal banks. Indeed, the T cells distinguish between what "belongs" and "does not belong" to the body and induce immune response as soon as they detect a foreign antigen. Consequently, it is not possible to use CAR T cells in an allogenic context without inducing graft versus host disease (GVHD).

Two principal strategies are being explored to address this issue and generate universal chimeric antigen receptor T cells which will be "off-the-shelf" allogeneic treatments (Fig. 9b). The first consists of using immune cells which does not induce immune response if they interact with a foreign antigen such as natural killer (Hu et al. 2018) or gamma delta T cells which is a particular subpopulation of T cells (Fisher and Anderson 2018). The principal issues are the efficacy of collecting the cells and genetically modifying them.

The second uses CAR T cells genetically modified with genome editing tools such as CRISPR or TALEN to inactivate genes responsible for the foreign antigen recognizing. UCAR19 is one of the most advanced therapies in the field, it is a genetically modified CAR T-cell product (anti-CD19 scFv-41BB-CD3ζ) manufactured from healthy donor cells, in which TRAC and CD52 genes have been disrupted to allow administration in non-HLA-matched patients (Qasim et al. 2017). UCAR19 is designed for relapsed/refractory acute lymphoblastic leukemia. In June 2015 at Great Ormond Street Hospital in London, the first baby, an 11-month-old girl called Layla, benefited from compassionate care using UCAR19 although it was a phase I trial. Another compassionate care has been realized with the same treatment in December 2015 with a 16-month-old girl. For now, they are healthy. However, themselves and other patients receiving CAR treatments or involving in CAR clinical trials must be followed up extremely closely. Indeed, this type of therapeutic strategy is very new and dramatic recent events impose great prudence. In 2018, University of Pennsylvania's Abramson Cancer Centre announced the death of one patient participating in the clinical trial which had allowed the validation of Kymriah. The researchers reported an unexpected mechanism, the accidental genetic engineering of a single leukemic cell during the early manufacturing of CAR T-cell therapy which induced a relapse

and the death of the patient (Ruella et al. 2018). This unanticipated event shows clearly that we have the necessary background knowledge to be able to evaluate precisely the safety of this type of therapy. The balance risk/benefit has to be carefully evaluated. On the other hand, the design of the therapeutic strategy has to be upgraded at both the clinical protocol and the design of the lentiviral vector. In this regard, the new generation of vector including switch systems allowing to inactivate modify cells reinjected is a major issue of the moment.

Infectious Diseases

A cure by gene therapy could be considered for a case of HIV infection. Several approaches are being studied. One of these is to make the CD4-T, the HIV-targeted cells, of patients resistant to the virus: HSC are removed from the patient and genetically modified with a genome editing tool to inactivate the gene coding for a CCR5 surface receptor which is the entry receptor of the HIV. Consequently, the absence of this receptor prevents the entry of the virus into the cells. Reinjected into the patient, these modified cells multiply and differentiate into immune cells resistant to the virus, making it possible to restore the subject's immune system (Gupta et al. 2019). A phase I/II trial is underway in the USA.

Recently, there was a tremendous upheaval in the scientific community in this area. Jiankui He, a scientist from a university in Shenzhen (China) claims he has succeeded in helping create the world's first genetically edited babies. At the Second International Summit on Human Genome Editing, that took place in November 2018, he presented his results about twin girls who were born earlier this month after he edited their embryos using CRISPR technology to remove the CCR5 gene which would provide them protection against HIV. Most of researchers condemned the irresponsible approach of Jiankui He, invoking a serious violation of laws, regulations, and ethical standards (Krimsky 2019; Wang et al. 2019a). Furthermore, for many of them, in addition to regulatory and ethical transgressions, this experience exposes children to risks without any real benefit; there are already safe and effective ways to prevent the transmission of the AIDS virus. Furthermore CCR5 has been also identified as a suppressor for cortical plasticity in brain regions involved in learning and memory processes; its inactivation in the whole body could induce cognitive side effects (Zhou et al. 2016). The summit's president, biologist David Baltimore, Nobel Laureate, denounced for his part "a lack of self-regulation by the scientific community due to a lack of transparency." The World Health Organization is currently forming an expert committee on human genome editing, which will meet soon. It will be tasked with examining "scientific, ethical, social and legal challenges" in this area.

Challenges

Biosafety

Despite the encouraging results obtained through several clinical trials and the approving of more and more products in gene therapy in both the USA and Europe, some severe side effects have been observed even going as far to the death of the patient in some cases. Consequently, researchers have to remain cautious about the use of gene therapy and the possible occurrence of adverse effects over time. Follow-ups of treated patients over several years will provide more information on the safety and efficacy of these drugs. The multiplication of clinical trials in various fields should allow us to learn a lot more in the years to come to further improve the processes. Research must go ahead to design new vectors with higher levels of biosafety. The next generations should address the problem of genotoxicity with controlled integration in safe locus without off targets or, when it is possible, induce a therapeutic effect from non-integrating forms. Regarding this, a new generation of lentiviral vectors remaining in RNA could meet these challenges, either to safely bring genome editing tools or to efficiently express a therapeutic factor. Another challenge in vectorology is to

answer the problem of the immune response which can develop in patients, in particular with AAV vectors, and the impossibility of reinjecting the treatment a second time. The development of new AAV capsids could answer this challenge.

Bioproduction

Another field of action is the bioproduction. Indeed, the production of vectors or genetically modified cells on an industrial scale remains a major obstacle for the development of innovative gene therapy drugs. The processes derive from research labs at an early stage and are not always suitable for large-scale deployment according to good manufacturing practices in pharmaceutical production plans. Technological and industrial innovations are still needed to improve production yields. Indeed, the doses required to treat a patient are often very large and the validation of clinical trials need high numbers of people, showing an important step to take for going to the market.

Pricing

The price of these drugs is also a new topic of public health thinking. The Glybera costed about one million dollars causing its commercial failure, the Strimvelis™ more than $700K per treatment, and the Spinraza is announced at several hundred thousand euros per year, for life. If these pricings can be understood for the actual benefit and the reduction of the costs of continuing care given to individuals suffering from rare genetic diseases, full-cost economic studies remain to be carried out in order to discuss emerging issues. Who could afford these types of therapies? How will the health insurance or the health systems in some countries manage these costs? How do you make gene therapy drugs available to disadvantaged populations is also one of the questions to start thinking about?

Legislation

Another issue is about the legislation. Indeed, the development of these innovative therapies raises new questions and requires adaptation of the laws. Some precautionary principles remain unclear when a patient is treated by gene therapy; what about the risk of dissemination of the vector? The solution in transgenesis is clear; genetically modified organisms must remain confined. Nevertheless, such a measure would have dramatic consequences in gene therapy. In this case, which regulations apply if extracellular detection of the vector persists in the treated patient? It seems that the legislation differs from one continent to another; in Japan, for example, the legislation provides for the holding of a patient treated by gene therapy in quarantine until the demonstration of the absence of dissemination of the vector in biological fluids is demonstrated. Knowing that in some clinical trials, AAV vector dissemination was found several years after vector injection, it is easy to imagine the consequences for the patient under Japanese law. It remains up for debate as to the relevance of European legislation compared to that of other countries.

References

Ahmed SG, Waddington SN, Boza-Morán MG, Yáñez-Muñoz RJ (2018) High-efficiency transduction of spinal cord motor neurons by intrauterine delivery of integration-deficient lentiviral vectors. J Control Release 273:99–107

Aiuti A, Slavin S, Aker M, Ficara F, Deola S, Mortellaro A, Morecki S, Andolfi G, Tabucchi A, Carlucci F et al (2002) Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science 296:2410–2413

Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L, Scaramuzza S, Andolfi G, Mirolo M, Brigida I et al (2009) Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med 360:447

Aiuti A, Biasco L, Scaramuzza S, Ferrua F, Cicalese MP, Baricordi C, Dionisio F, Calabria A, Giannelli S, Castiello MC et al (2013) Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science 341:1233151

Alonso-Padilla J, Papp T, Kaján GL, Benkő M, Havenga M, Lemckert A, Harrach B, Baker AH (2016) Development of novel adenoviral vectors to overcome challenges observed with HAdV-5–based constructs. Mol Ther 24:6–16

Benchaouir R, Meregalli M, Farini A, D'Antona G, Belicchi M, Goyenvalle A, Battistelli M, Bresolin N, Bottinelli R, Garcia L et al (2007) Restoration of human dystrophin following transplantation of exon-skipping-engineered DMD patient stem cells into dystrophic mice. Cell Stem Cell 1:646–657

Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, Clerici M, Shearer G, Chang L, Chiang Y, Tolstoshev P et al (1995) T lymphocyte-directed gene therapy for ADA−SCID: initial trial results after 4 years. Science 270:475–480

Boado RJ, Zhang Y, Zhang Y, Xia C-F, Pardridge WM (2007) Fusion antibody for Alzheimer's disease with bidirectional transport across the blood-brain barrier and abeta fibril disaggregation. Bioconjug Chem 18:447–455

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding sof TAL-type III effectors. Science 326:1509–1512

Bommareddy PK, Shettigar M, Kaufman HL (2018) Integrating oncolytic viruses in combination cancer immunotherapy. Nat Rev Immunol 18:498

Campeau P, Chapdelaine P, Seigneurin-Venin S, Massie B, Tremblay JP (2001) Transfection of large plasmids in primary human myoblasts. Gene Ther 8:1387–1394

Cavalieri V, Baiamonte E, Lo Iacono M (2018) Non-primate lentiviral vectors and their applications in gene therapy for ocular disorders. Viruses 10:316

Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, Gross F, Yvon E, Nusbaum P, Selz F, Hue C, Certain S, Casanova JL et al (2000) Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science 288:669–672

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823

Curran KJ, Pegram HJ, Brentjens RJ (2012) Chimeric antigen receptors for T cell immunotherapy: current understanding and future directions. J Gene Med 14:405–415

Darrow JJ (2019) Luxturna: FDA documents reveal the value of a costly gene therapy. Drug Discov Today 24:949

Deverman BE, Ravina BM, Bankiewicz KS, Paul SM, Sah DWY (2018) Gene therapy for neurological disorders: progress and prospects. Nat Rev Drug Discov 17:641–659

Ding S, Wu X, Li G, Han M, Zhuang Y, and Xu T (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122 473–483

Donsante A, Miller DG, Li Y, Vogler C, Brunt EM, Russell DW, Sands MS (2007) AAV vector integration sites in mouse hepatocellular carcinoma. Science 317:477–477

Duncan C, Eichler F, Thrasher AJ, Orchard PJ, De Oliveira SN, Lund TC, Amartino H, Smith NJC, McNeil E, Chin W et al (2019) Updated safety and efficacy outcomes from an ongoing phase 2/3 trial of lenti-D autologous hematopoietic stem cell gene therapy for the treatment of cerebral adrenoleukodystrophy. Biol Blood Marrow Transplant 25:S89–S90

Felgner PL, Gadek TR, Holm M, Roman R, Chan HW, Wenz M, Northrop JP, Ringold GM, Danielsen M (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc Natl Acad Sci U S A 84:7413–7417

Fisher J, Anderson J (2018) Engineering approaches in human gamma delta T cells for cancer immunotherapy. Front Immunol 9:1409

Friedmann T, Roblin R (1972) Gene therapy for human genetic disease? Science 175:949–955

Gounis MJ, Spiga M-G, Graham RM, Wilson A, Haliko S, Lieber BB, Wakhloo AK, Webster KA (2005) Angiogenesis is confined to the transient period of VEGF expression that follows adenoviral gene delivery to ischemic muscle. Gene Ther 12:762–771

Gray SJ, Woodard KT, Samulski RJ (2010) Viral vectors and delivery strategies for CNS gene therapy. Ther Deliv 1:517–534

Green M, Loewenstein PM (1988) Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. Cell 55:1179–1188

Greenberg B, Butler J, Felker GM, Ponikowski P, Voors AA, Desai AS, Barnard D, Bouchard A, Jaski B, Lyon AR et al (2016) Calcium upregulation by percutaneous administration of gene therapy in patients with cardiac disease (CUPID 2): a randomised, multinational, double-blind, placebo-controlled, phase 2b trial. Lancet Lond Engl 387:1178–1186

Guo J, Xin H (2006) Splicing out the west? Science 314:1232–1235

Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, Salgado M, Martinez-Picado J, Nijhuis M, Wensing AMJ, Lee H et al (2019) HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. Nature 568:244–248

Hacein-Bey-Abina S, Kalle CV, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E et al (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302:415–419

Hess B, Saftig P, Hartmann D, Coenen R, Lüllmann-Rauch R, Goebel HH, Evers M, von Figura K, D'Hooge R, Nagels G et al (1996) Phenotype of arylsulfatase A-deficient mice: relationship to human metachromatic leukodystrophy. Proc Natl Acad Sci U S A 93:14821–14826

Hirooka Y, Kasuya H, Ishikawa T, Kawashima H, Ohno E, Villalobos IB, Naoe Y, Ichinose T, Koyama N, Tanaka M et al (2018) A Phase I clinical trial of EUS-guided intratumoral injection of the oncolytic virus, HF10 for unresectable locally advanced pancreatic cancer. BMC Cancer 18:596

Hirsch T, Rothoeft T, Teig N, Bauer JW, Pellegrini G, De Rosa L, Scaglione D, Reichelt J, Klausegger A, Kneisz D et al (2017) Regeneration of the entire human epidermis using transgenic stem cells. Nature 551:327–332

Hu Y, Tian Z, Zhang C (2018) Chimeric antigen receptor (CAR)-transduced natural killer cells in tumor immunotherapy. Acta Pharmacol Sin 39:167–176

Ivics Z, Hackett PB, Plasterk RH, and Izsvák Z (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91 501–510

Jordan Calmes-Miller (2016) Railroading at the FDA. Nat Biotechnol 34:1078

Jordan Calmes-Miller (2018) FDA approves second CAR T-cell therapy. Cancer Discov 8:5–6

Khorkova O, Wahlestedt C (2017) Oligonucleotide therapies for disorders of the nervous system. Nat Biotechnol 35:249–263

Kinali M, Arechavala-Gomeza V, Feng L, Cirak S, Hunt D, Adkin C, Guglieri M, Ashton E, Abbs S, Nihoyannopoulos P et al (2009) Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. Lancet Neurol 8:918–928

Kohn DB (2019) Gene therapy for blood diseases. Curr Opin Biotechnol 60:39–45

Krimsky S (2019) Ten ways in which He Jiankui violated ethics. Nat Biotechnol 37:19–20

Li T, Adamian M, Roof DJ, Berson EL, Dryja TP, Roessler BJ, Davidson BL (1994) In vivo transfer of a reporter gene to the retina mediated by an adenoviral vector. Invest Ophthalmol Vis Sci 35:2543–2549

Lu QL, Mann CJ, Lou F, Bou-Gharios G, Morris GE, Xue S, Fletcher S, Partridge TA, Wilton SD (2003) Functional amounts of dystrophin produced by skipping the mutated exon in the mdx dystrophic mouse. Nat Med 9:1009–1014

Maetzig T, Galla M, Baum C, Schambach A (2011) Gammaretroviral vectors: biology, technology and application. Viruses 3:677–713

Magee TR, Artaza JN, Ferrini MG, Vernet D, Zuniga FI, Cantini L, Reisz-Porszasz S, Rajfer J, Gonzalez-Cadavid NF (2006) Myostatin short interfering hairpin RNA gene transfer increases skeletal muscle mass. J. Gene Med. 8:1171–1181

Matzner U, Hartmann D, Lüllmann-Rauch R, Coenen R, Rothert F, Månsson J-E, Fredman P, D'Hooge R, De Deyn PP, Gieselmann V (2002) Bone marrow stem cell-based gene transfer in a mouse model for metachromatic leukodystrophy: effects on visceral and nervous system disease manifestations. Gene Ther 9:53–63

Moroni MC, Willingham MC, Beguinot L (1992) EGF-R antisense RNA blocks expression of the epidermal growth factor receptor and suppresses the transforming phenotype of a human carcinoma cell line. J Biol Chem 267:2714–2722

Monfardini C, Veronese FM (1998) Stabilization of substances in circulation. Bioconjug. Chem. 9 418–450

Naldini L, Blömer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D (1996) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 272:263–267

Negri DR, Michelini Z, Baroncelli S, Spada M, Vendetti S, Buffa V, Bona R, Leone P, Klotman ME, Cara A (2007) Successful immunization with a single injection of non-integrating lentiviral vector. Mol Ther 15:1716–1723

Ogris M, Brunner S, Schüller S, Kircheis R, Wagner E (1999) PEGylated DNA/transferrin-PEI complexes: reduced interaction with blood components, extended circulation in blood and potential for systemic gene delivery. Gene Ther. 6 595–605

Pardridge WM (2005) Tyrosine hydroxylase replacement in experimental Parkinson's disease with transvascular gene therapy. NeuroRx 2:129–138

Pattanayak V, Ramirez CL, Joung JK, Liu DR (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods 8:765–770

Perrin GQ, Herzog RW, Markusic DM (2019) Update on clinical gene therapy for hemophilia. Blood 133:407–414

Perumbeti A, Malik P (2010) Therapy for β-globinopathies: a brief review and determinants for successful and safe correction. Ann N Y Acad Sci 1202:36–44

Philippe S, Sarkis C, Barkats M, Mammeri H, Ladroue C, Petit C, Mallet J, Serguera C (2006) Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo. Proc Natl Acad Sci U S A 103:17684–17689

Plank C, Mechtler K, Szoka FC, Wagner E (1996) Activation of the complement system by synthetic DNA complexes: a potential barrier for intravenous gene delivery. Hum. Gene Ther. 7 1437–1446

Qasim W, Zhan H, Samarasinghe S, Adams S, Amrolia P, Stafford S, Butler K, Rivat C, Wright G, Somana K et al. (2017) Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. Sci. Transl. Med. 9

Ribas A, Dummer R, Puzanov I, Vander Walde A, Andtbacka RHI, Michielin O, Olszanski AJ, Malvehy J, Cebon J, Fernandez E et al (2017) Oncolytic virotherapy promotes intratumoral T cell infiltration and improves anti-PD-1 immunotherapy. Cell 170:1109–1119.e10

Roth JA, Swisher SG, Merritt JA, Lawrence DD, Kemp BL, Carrasco CH, El-Naggar AK, Fossella FV, Glisson BS, Hong WK et al (1998) Gene therapy for non-small cell lung cancer: a preliminary report of a phase I trial of adenoviral p 53 gene replacement. Semin Oncol 25:33–37

Ruella M, Xu J, Barrett DM, Fraietta JA, Reich TJ, Ambrose DE, Klichinsky M, Shestova O, Patel PR, Kulikovskaya I et al (2018) Induction of resistance to chimeric antigen receptor T cell therapy by transduction of a single leukemic B cell. Nat Med 24:1499

Sahel JA, Dalkara D (2019) Gene therapy for retinal dystrophy. Nat Med 25:198

Sarkis C, Altemir D, Grandchamp N, Philippe S (2014) Transient expression vectors, preparation and uses thereof. France

Science AA for the A. of (2000) Biomedical ethics on the front burner. Science 290:2225–2225

Sessa M, Lorioli L, Fumagalli F, Acquati S, Redaelli D, Baldoli C, Canale S, Lopez ID, Morena F, Calabria A et al (2016) Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. Lancet 388:476–487

Sevin C, Benraiss A, Van Dam D, Bonnin D, Nagels G, Verot L, Laurendeau I, Vidaud M, Gieselmann V, Vanier M et al (2006) Intracerebral adeno-associated virus-mediated gene transfer in rapidly progressive forms of metachromatic leukodystrophy. Hum Mol Genet 15:53–64

Sheridan C (2017) First approval in sight for Novartis' CAR-T therapy after panel vote. Nat Biotechnol 35:691–693

Stirnadel-Farrant H, Kudari M, Garman N, Imrie J, Chopra B, Giannelli S, Gabaldo M, Corti A, Zancan S, Aiuti A et al (2018) Gene therapy in rare diseases: the benefits and challenges of developing a patient-centric registry for Strimvelis in ADA-SCID. Orphanet J Rare Dis 13:49

Tardieu M, Zérah M, Gougeon M-L, Ausseil J, de Bournonville S, Husson B, Zafeiriou D, Parenti G, Bourget P, Poirier B et al (2017) Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: an uncontrolled phase 1/2 clinical trial. Lancet Neurol 16:712–720

Tatsis N, Ertl HCJ (2004) Adenoviruses as vaccine vectors. Mol Ther 10:616–629

Thompson AA, Walters MC, Kwiatkowski J, Rasko JEJ, Ribeil J-A, Hongeng S, Magrin E, Schiller GJ, Payen E, Semeraro M et al (2018) Gene therapy in patients with transfusion-dependent β-Thalassemia. N Engl J Med 378:1479

Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435:646–651

Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, Chang T, Huang H, Lin R-J, Yee J-K (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. Nat Biotechnol 33:175–178

Wang C, Zhai X, Zhang X, Li L, Wang J, Liu D (2019a) Gene-edited babies: Chinese Academy of Medical Sciences' response and action. Lancet 393:25–26

Wang D, Tai PWL, Gao G (2019b) Adeno-associated virus vector as a platform for gene therapy delivery. Nat Rev Drug Discov 18:358–378

Williams DS, Chadha A, Hazim R, Gibbs D (2017) Gene therapy approaches for prevention of retinal degeneration in Usher syndrome. Gene Ther 24:68–71

Wilson JM (2005) Gendicine: the first commercial gene therapy product. Hum Gene Ther 16:1014–1015

Xiao X, Li J, Samulski RJ (1996) Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. J Virol 70:8098–8108

Xu Y, Szoka FC (1996) Mechanism of DNA release from cationic liposome/DNA complexes used in cell transfection. Biochemistry (Mosc.) 35 5616–5623

Yáñez-Muñoz RJ, Balaggan KS, MacNeil A, Howe SJ, Schmidt M, Smith AJ, Buch P, MacLaren RE, Anderson PN, Barker SE et al (2006) Effective gene therapy with nonintegrating lentiviral vectors. Nat Med 12:348–353

Zhang Y, Schlachetzki F, Li JY, Boado RJ, Pardridge WM (2003) Organ-specific gene expression in the rhesus monkey eye following intravenous non-viral gene transfer. Mol Vis 9:465–472

Zhang S, Xu G, Liu C, Xiao S, Sun Y, Su X, Cai Y, Li D, Xu B (2005) Clinical study of recombinant adenovirus-p53 (Adp53) combined with hyperthermia in advanced cancer (a report of 15 cases). Int J Hyperth 21:631–636

Zhou M, Greenhill S, Huang S, Silva TK, Sano Y, Wu S, Cai Y, Nagaoka Y, Sehgal M, Cai DJ et al (2016) CCR5 is a suppressor for cortical plasticity and hippocampal learning and memory. elife 5:e20985

Pharmacological Therapy in Inborn Errors of Metabolism

18

Anibh M. Das and Sabine Illsinger

Contents

Abstract

Different therapeutic principles can improve clinical outcome in patients suffering from inborn errors of metabolism (IEM). IEMs are inherited disorders which are based on a primary enzyme defect or deficiency of a cellular transporter. Substrate reduction by diet (exogenous substrate), pharmacological substrate reduction (endogenous substrate), supplementation of a missing cofactor/vitamin, activation of alternative pathways for the elimination of toxic compounds, augmentation of enzyme activity by chaperones, and enzyme replacement therapy (in selected diseases like lysosomal storage diseases) are therapeutic options for alimentary and pharmacological treatment in IEM. In this chapter we will discuss options for pharmacological and/or dietary therapy in selected prototype IEM.

Introduction

Inborn errors of metabolism (IEMs) are inherited disorders which are either due to a primary enzyme defect or deficiency of a transporter. In enzyme deficiencies, the physiological substrate of the enzyme reaction accumulates, while the product of this reaction (distal to the reaction) is

A. M. Das (✉)
Department of Paediatrics, Hannover Medical School, Hannover, Germany
e-mail: das.anibh@mh-hannover.de

S. Illsinger
Department of Neuropaediatrics, Children's Hospital Oldenburg, Oldenburg, Germany
e-mail: Illsinger.Sabine@klinikum-oldenburg.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_62

decreased. Accumulation of substrates can occur quickly, especially in catabolic states or during dietary indiscretions, thus intoxicating the organism. Therefore, this group of IEM is called "intoxication-type" IEM. If an intracellular transporter is compromised, the substrate to be transported accumulates in one cell compartment and is low in others (storage disease, e.g., lysosomal storage disorders (LSD)) (Fig. 1). As it takes some time until substantial levels of substrate accumulate, these disorders mostly have a more chronic course, with the exception of M. Niemann-Pick type C where neonatal liver failure occurs in some patients.

All IEMs are rare diseases, so-called orphan diseases. In Europe, these are defined as diseases affecting fewer than 5 in 10,000 inhabitants, in the USA a disease affecting less than 200,000 inhabitants (<1:1,500), and in Japan a condition present in less than 50,000 inhabitants (<1:2,500) is called an orphan disease.

It was Sir Archibald Garrod who first described an IEM, namely, alkaptonuria, which is a disorder in the degradative pathway of the amino acid tyrosine, and coined the notion "Inborn Error of Metabolism" more than 100 years ago (Garrod 1902). Today, several hundred IEMs are known; each year new IEMs are described.

Based on the rarity of IEM, it is difficult to perform randomized controlled studies; mostly a multinational, multicenter approach is required.

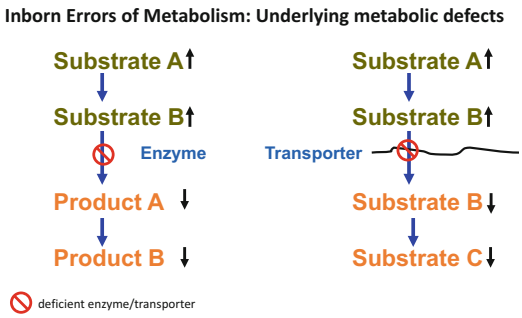


Fig. 1 Principles of inborn errors of metabolism (IEMs): IEMs are either due to an enzyme deficiency of compromised transporter function. This leads to the accumulation of the substrate of the deficient enzyme and reduced levels of its product or accumulation of the substrate of a deficient transporter in one cell compartment

The development of new therapies is challenging for pharmaceutical companies as the small patient number renders the studies necessary for approval by regulatory bodies difficult; at the same time, the economical return is limited after approval is granted resulting in a high financial burden per patient for the health system. Politically, the challenges for treating patients with orphan diseases have been recognized, and measures have been taken to support pharmaceutical companies in marketing orphan drugs. For example, the European Medicines Agency (EMA) offers protocol assistance free of charge to companies intending to market orphan drugs; on the other hand, marketing exclusivity for these compounds is longer (10 years), and a reduced fee for approval is levied for orphan drugs. This has led to the approval of 128 new orphan drugs from 2000 to 2016; 1805 drugs got orphan drug designation (http://www.ema.europa.eu/ema/). Nineteen percent of orphan drugs were approved for the indication "alimentation and metabolism."

For many IEM, a good, meaningful outcome parameter is missing. Obviously, the clinical outcome can be judged; sometimes the substrate and the product of the missing enzyme can be measured (e.g., phenylalanine and tyrosine in phenylketonuria (PKU)) to monitor the therapy.

Therapeutic Principles of Treatment in IEM

Understanding the pathophysiology of an IEM is essential for designing a therapy. If the disease is due to an enzyme defect, reducing the substrate and supplementing the product of the compromised enzymatic reaction are useful. In diseases where the enzyme substrate is administered via food, dietary measures are promising. The prototype of these disorders is classical phenylketonuria (PKU) where a low-protein diet has to be followed, while tyrosine, the product of the deficient phenylalanine hydroxylase, is supplemented by a special amino acid mixture. If an endogenous substance is the substrate of the deficient enzyme reaction, pharmacological inhibition of the synthesis of the substrate is an

option (substrate reduction therapy). In some patients with an enzyme deficiency, the enzyme is not completely absent; chaperones can stabilize the enzyme resulting in an increased residual activity. Some enzymes depend on a cofactor which may be deficient due to an IEM. In these diseases, supplementation of the cofactor is helpful (e.g., in biotinidase deficiency). In other IEMs, alternative excretion pathways may be activated by pharmacological treatment leading to the excretion/elimination of accumulating toxic substances by alternative pathways (e.g., in urea cycle defects, benzoate and butyrate can eliminate nitrogen atoms independently of the compromised urea cycle). Chaperones are able to stabilize the structure of the deficient enzyme; only mutations where the enzyme is still present, hence milder mutations, are amenable to enzyme augmentation by chaperones. For a small number of IEMs, it is possible to produce the missing enzyme in a bioreactor using cell cultures, couple the enzyme to mannose 6-phosphate or expose mannose 6-phosphate, and apply the drug to the patients via intravenous infusion every week or every other week depending on the underlying disease. Via mannose 6-phosphate receptors on the cell surface, the enzyme is taken up and is trafficked to the lysosome to replace the missing enzyme. Such therapeutic options exist for some of the so-called lysosomal storage disorders.

In some disorders where the enzyme is localized exclusively or predominantly in a single organ, solid organ or cell transplantation can be offered to replace the missing enzyme. These procedures are not without risk and lead to chronic disease burden (immunosuppression, risk of infection, transplantation-associated malignancies); therefore alternative treatment options should be carefully considered. Transplantation will not be discussed further in this article which is focused on dietary and pharmacological treatment strategies.

In summary, we have the following options for alimentary and pharmacological treatment in IEM:

1. Substrate reduction by diet (exogenous substrate)
2. Pharmacological substrate reduction (endogenous substrate)
3. Supplementation of a missing cofactor/vitamin
4. Activation of alternative pathways for the elimination of toxic compounds
5. Augmentation of enzyme activity by chaperones
6. Enzyme replacement therapy (in selected diseases like lysosomal storage diseases)

Obviously, it will not be possible to discuss options of pharmacological therapy in all IEMs; examples for these therapeutic options will be presented in the following sections of this chapter. Some pharmacological compounds used for the treatment of IEM were originally developed for other purposes; during toxicity testing, it was revealed that these substances have effects on metabolic pathways. This will be exemplified in the section on pharmacological substrate reduction in patients suffering from tyrosinemia type 1.

Substrate Reduction by Diet

Classical phenylketonuria (PKU) due to phenylalanine hydroxylase (PAH) deficiency is the prototype of this group of IEM amenable to dietary treatment.

The pathophysiology of the disease was discovered in 1934 by A. Folling (Christ 2003) which was the basis for treating the condition. In PKU, the metabolism of phenylalanine is impaired due to variants of the gene encoding PAH. Usually, PAH converts phenylalanine to tyrosine requiring tetrahydrobiopterin (BH4) as cofactor. PAH deficiency leads to the accumulation of phenylalanine in the blood and brain. If left untreated, this condition will lead to irreversible brain damage presenting with intellectual disability, severe developmental delay, psychiatric abnormalities, epilepsy, and microcephaly (Blau et al. 2010).

In their seminal work, H. Bickel and coworkers in 1953 showed the positive clinical effect of a low-protein diet (without meat and fish) supplemented with an amino acid formula devoid of phenylalanine and rich in tyrosine, essential

amino acids, vitamins, and minerals (Bickel et al. 1954). Today, low-protein diet is still the cornerstone of treatment in classical PKU; new age-adjusted amino acid formulas devoid of phenylalanine and enriched in tyrosine with improved palatability have been developed in the last years. These formulas should be given in three daily doses to avoid amino acid overload as well as to minimize losses of L-amino acids by oxidative processes and to minimize fluctuations in phenylalanine concentrations (MacDonald et al. 1996).

The most recent achievement in terms of palatability are glycomacropeptides, proteins derived from cheese whey which is low in phenylalanine and rich in branched-chain amino acids and threonine. Glycomacropeptides have to be supplemented with essential amino acids like tyrosine, tryptophan, arginine, cysteine, and histidine (for an overview, see, e.g., Al Hafid and Christodoulou (2015)).

Possible side effects of phenylalanine-free L-amino acid supplements are gastrointestinal symptoms, proteinuria, and dental caries in some patients (van Wegberg et al. 2017a).

In many countries, PKU is a target disease of newborn mass screening programs; thus treatment can be started just after birth. The amount of phenylalanine in the diet has to be individually titrated; the amount of amino acid formula is calculated based on the recommended daily protein intake. Phenylalanine is essential for protein synthesis and must be provided in an amount that supports growth and tissue repair while keeping plasma phenylalanine concentrations within recommended age-specific ranges for PKU (Macleod and Ney 2010).

Modifications of phenylalanine target levels have recently been proposed (van Spronsen and Derks 2014). Diet for life secures relatively normal life for PKU patients allowing academic careers; however, minor neurological, psychiatric, or behavioral sequelae may be observed in some patients (Koch et al. 2002); long-term outcome in the aging PKU population is still unclear.

In pregnancies of females with PKU ("maternal PKU"), separate, strict treatment recommendations have to be followed to avoid severe fetal and perinatal compromise (Lenke and Levy 1980).

Similar therapeutic principles of low-protein diet supplemented with an amino acid formula devoid of the amino acid(s) which cannot be metabolized due to a missing enzyme are available for maple syrup urine disease (MSUD), hepatorenal tyrosinemia, and organic acidurias (e.g., methylmalonic aciduria, glutaric aciduria, etc.). Some of them are target diseases of the newborn screening in some countries as well.

In recent years, sapropterin, a synthetic analogue of tetrahydrobiopterin BH4, was approved in Europe as a pharmacological chaperone treatment option in PKU patients harboring amenable *PAH* mutations with relatively high residual enzyme activity (see section "Enzyme Augmentation by Chaperones").

An enzyme 'replacement' therapy is under development (polyethylene glycol phenylalanine ammonia-lyase, PEG-PAL) replacing not the missing enzyme but an enzyme activating an alternative pathway of elimination (see section "Enzyme Replacement Therapy (ERT)").

Pharmacological Substrate Reduction

Treatment of Hereditary Tyrosinemia Type 1 (Hepatorenal Tyrosinemia, HT-1) by Nitisinone

Hereditary tyrosinemia type 1 (hepatorenal tyrosinemia, HT-1) is a rare inborn error of metabolism in the catabolism of the amino acid tyrosine (Fig. 2).

Biochemically, there is deficiency of fumarylacetoacetase leading to the accumulation not only of tyrosine but of different metabolites like maleylacetoacetate, fumarylacetoacetate, and succinylacetone (SA); the latter is commonly used as a surrogate parameter of toxicity in blood and/or urine (Fig. 2).

Therapy consists of nitisinone (Orfadin®, 2-(2-**n**itro-4**t**rifluoromethyl**b**enzoyl)**c**yclohexane 1-,3-dione (NTBC)) in order to decrease levels of toxic compounds (Fig. 2). Nitisinone blocks the degradation of tyrosine proximal to the enzyme fumarylacetoacetase and prevents the production

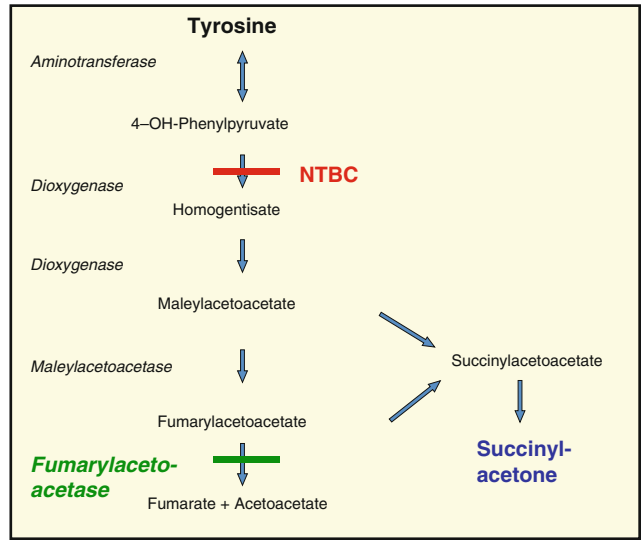


Fig 2 Tyrosinemia is due to deficiency of the enzyme fumarylacetoacetase in the degradative pathway of tyrosine. It results in the accumulation of toxic compounds; succinylacetone is used as a surrogate parameter of toxicity. NTBC (nitisinone) inhibits the enzyme homogentisate dioxygenase with subsequent reduction of toxic metabolites and increased tyrosine levels

of toxic metabolites. A low-protein diet combined with special amino acid mixtures devoid of tyrosine and its precursor phenylalanine is required to prevent excessive tyrosine levels. Nitisinone was approved by the European Medicines Agency (EMA) under exceptional circumstances in 2005. The first clinical use of nitisinone in HT-1 dates back to 1991. Originally, nitisinone was developed as a weed killer by Zeneca Agrochemicals (for a review, see Das AM The Application of Clin Genet 2017). It was epidemiologically observed that growth of plants and weeds was inhibited under the bottlebrush plant (*Callistemon citrinus*). It became clear that neither the shade nor the litterfall of these plants was responsible for suppression of plant and weed growth. Rather a substance, which was identified as leptospermone, in the soil under the bottlebrush plant was shown to have bleaching activity on the emerging plants. The allelochemical leptospermone was extracted from the bottlebrush plant and chemically characterized. It belongs to the triketone family which inhibits chloroplast development due to lack of plastoquinone secondary to hepatic 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibition and served as a blueprint for the synthesis of nitisinone.

Toxicology testing of nitisinone revealed that it was not acutely toxic, but eye lesions (keratopathy) could be observed in animals after longer treatment. Lesions were reversible upon withdrawal of the compound. Elevated tyrosine levels were found in blood and urine after nitisinone exposure. It was established that nitisinone is a potent inhibitor of rat hepatic 4-hydroxyphenylpyruvate dioxygenase (HPPD) at the Zeneca Central Toxicity Laboratories (Fig. 2). Inhibition of human HPPD was demonstrated in the human liver by Sven Lindstedt and his group at Gothenburg University (Sweden). Already for some time, they were looking for an inhibitor of HPPD to treat patients with the lethal disease HT-1. The Zeneca Pharmaceuticals took over the compound as a potential drug though they were reluctant to do so as only a very limited number of patients with this rare disease would possibly benefit from this drug. The Swedish Medical Agency approved a clinical trial with nitisinone in HT-1 patients, and in February 1991, a critically ill 2-month-old baby was the first HT-1 patient treated with nitisinone in Gothenburg. Succinylacetone quickly disappeared from the urine and the clinical state gradually improved. Subsequently, several HT-1 patients were successfully treated with nitisinone on a compassionate use basis. In the following years, many patients were successfully treated with nitisinone culminating in the approval of nitisinone by the US Food and Drug Administration in 2002 and the European Medicines Agency (EMA) in 2005 under "exceptional

circumstances" as large field trials and real-world data in HT-1 patients were lacking. The first sublicense-holder Swedish Orphan International was obliged to perform post-marketing studies which are now carried out by the present sublicense-holder Swedish Orphan Biovitrum (Sobi) in the framework of the "OPAL" study.

In the rat liver, enzyme kinetic studies revealed inhibition of HPPD by nitisinone in a dose- and time-dependent manner with a rate constant of 9.9×10^{-5} s^{-1} $(nmol/L)^{-1}$ (Ellis et al. 1995). Binding of nitisinone to HPPD is not irreversible; the enzyme-inhibitor complex dissociates with a half-life of 63 h in rats at a temperature of 25 °C. Tests in human adult volunteers revealed that the half-life of nitisinone was 54 h (Hall et al. 2001). It is recommended to take nitisinone in two daily doses. However, based on the long half-life of nitisinone, once-daily dosing was advocated (Schlune et al. 2012) which seems to be adequate to maintain metabolic control. Once-daily dosing may improve adherence to pharmacological therapy. However, metabolic control with a once daily dosing regime has been recently questioned (Kienstra et al. 2018).

The recommended dose of nitisinone is 1–2 mg/kg per day given in two doses. In a recent survey (Mayorandan et al. 2014a), we found several HT-1 patients in whom nitisinone was titrated down without hampering metabolic control as judged by the absence of succinylacetone; doses as low as 0.3 mg/kg per day of nitisinone have been shown to be sufficient.

While it is clear that nitisinone can reduce the levels of toxic compounds as judged by the surrogate parameter succinylacetone and secure survival, the impact on long-term clinical outcome is less clear. This is mainly due to the rarity of the disease. In the literature, there is only one multinational cross-sectional study including 168 patients from 21 centers (Mayorandan et al. 2014a). This study showed that there is a clear benefit of nitisinone treatment in combination with low-protein diet supplemented by amino acid mixtures devoid of tyrosine and phenylalanine. If present, liver dysfunction/failure, renal dysfunction, tubulopathy, and rickets could be reversed by nitisinone treatment. Also the rate of HCC could be reduced by nitisinone treatment. Long-term complications, especially HCC formation, critically depend on early initiation of treatment, hence early diagnosis. If treatment is started beyond the first year of life, the risk to develop HCC is 13 times higher compared to a treatment start in the neonatal period, risk for liver cirrhosis is 40-fold, rickets is 19-fold, and tubular dysfunction is 4.3-fold increased (Mayorandan et al. 2014b). Thus, early diagnosis in the newborn period is essential to secure a good long-term outcome.

Neurocognitive deficits are a problem in long-term management (Bendadi et al. 2014; Garcia et al. 2017). It is not clear whether these are direct side effects of nitisinone treatment, are a result of high tyrosine or low phenylalanine levels, or are part of the natural course of the disease. It seems reasonable to reduce nitisinone levels without compromising metabolic control (as judged by succinylacetone in urine and/or blood). Therefore, nitisinone levels should be regularly monitored in dried blood spots, together with succinylacetone (Sander et al. 2011). In most cases, it is possible to titrate down the daily dose of nitisinone from initially 1 mg/kg per day without hampering metabolic control. The tentative therapeutic range of random nitisinone concentration in dried blood is 20–40 μM (Sander et al. 2011).

Gaucher Disease and Eliglustat and Miglustat

Gaucher disease is a rare autosomal recessive storage disorder based on a deficiency of acid ß-glucosidase (glucocerebrosidase). This leads to the storage of glucosylceramide in different organs of the body. Hepatosplenomegaly, bone marrow infiltration with subsequent pancytopenia, bone destruction, and lung disease are common symptoms (Bennett and Mohan 2013). While enzyme replacement therapy (ERT) for M. Gaucher type 1 is available since decades (see section "Enzyme Replacement Therapy (ERT)"), eliglustat (Cerdelga®) has recently been approved by the US Food and Drug Administration and EMA for pharmacological substrate reduction. Eliglustat is given orally twice daily and as a ceramide analogue inhibits glucosylceramide

synthase (IC 50, 24 nM) (Lukina et al. 2010a). Glucosylceramide synthase is the first enzyme that catalyzes the biosynthesis of glycosphingolipids. Eliglustat is metabolized via cytochrome 450 2D6 and is dosed according to the activity of this enzyme. Favorable safety and efficacy data were observed in several studies before approval was granted. The half-life of eliglustat was 6.8 h (Lukina et al. 2010b), hemoglobin and platelets quickly increased, spleen and liver volumes decreased, and bone crises were less frequent. A specific and sensitive biomarker to evaluate the activity of Gaucher disease is still missing. For M. Gaucher types 2 and 3 (progressive neuropathic forms), no approved drugs are presently available.

Miglustat (Zavesca®) is also approved for the treatment of non-neuronopathic Gaucher disease acting via substrate reduction. Based on the high rate of gastrointestinal side effects of miglustat, eliglustat is preferred by most patients.

Niemann-Pick Type C Disease and Miglustat

M. Niemann-Pick type C (NPC) is an autosomal recessive disorder characterized by impaired trafficking of endocytosed cholesterol with sequestration of unesterified cholesterol in lysosomes and late endosomes (Vanier 2010). Other lipids accumulate as well with variation in different tissues. The disease is due to a defect of the NPC1 (95% of cases) or NPC2 gene. Clinically, visceral organs like liver, spleen, and lungs may be affected; neurological symptoms (clumsiness, learning difficulties, ataxia, dysphagia, dysarthria, vertical gaze palsy) with neurodegeneration are common CNS symptoms.

Substrate reduction using miglustat (Zavesca®) is a disease-modifying therapeutic option, not a cure, probably due to the complexity of the disease pathophysiology with accumulation of many substrates.

Miglustat (deoxynojirimycin) is an iminosugar which reversibly inhibits glycosphingolipid synthesis thus reducing accumulation of glycosphingolipids in NPC (Williamson 2014). This drug can stabilize the clinical course of the disease and may reduce disease progression in some patients. Moderate to severe gastrointestinal side effects (abdominal discomfort, flatulence, and diarrhea) are common and can be ameliorated by modifying the diet (elimination of disaccharides).

Supplementation of a Missing Cofactor/Vitamin

Biotinidase Deficiency

Biotinidase plays a critical role in the uptake of biotin from dietary sources and in biotin recycling. Free biotin enters the biotin cycle from dietary sources or from the cleavage of biocytin or biotinyl peptides by the action of biotinidase. In biotinidase deficiency, an autosomal recessive disease caused by mutations in the *BTD* gene, both deficiencies of biotin and in biotin-dependent carboxylases occur. Treatment with oral administration of free biotin (up to 10 mg per day) can prevent symptoms like cutaneous rash, hair loss, seizures, inner ear hearing loss, and developmental delay. It should ideally be started during the neonatal period to secure complete absence of clinical symptoms after establishing the diagnosis via newborn mass screening. For patients with biotinidase deficiency, newborn screening is a very efficient preventive measure. All individuals with profound biotinidase deficiency, even those who have residual enzyme activity, should be treated with biotin for life. Raw eggs should be avoided because they contain avidin, an egg-white protein that binds biotin and decreases the bioavailability of the vitamin (Wolf 1993).

Metabolic ketolactic acidosis accompanied by organic aciduria (3-hydroxyisovaleric acid may be the only metabolite present) and hyperammonemia are suggestive of biotinidase deficiency; however, urinary organic acids can be normal in symptomatic individuals with biotinidase deficiency.

Cobalamin (Vitamin B12) in Transcobalamin Deficiency and Methylmalonic Aciduria

Cobalamin, also known as vitamin B12, is a water-soluble vitamin that is crucial for the normal function of the gastrointestinal, dermatologic,

immunologic, neuropsychiatric, and hematopoietic systems. The vitamin is essential and must be supplied by diet. Transcobalamin (TCN2) is required to internalize vitamin B12 into the cells through membrane receptor-mediated endocytosis. Vitamin B12 is then processed in the cytoplasm and mitochondria by complementation factors leading to its active metabolites methylcobalamin and 5-deoxyadenosyl-cobalamin.

Deficiency of TCN2 results in an elevation of methylmalonic acid in urine and homocysteine in plasma. Patients usually present with macrocytic anemia, pancytopenia, failure to thrive, gastrointestinal symptoms, and neurological dysfunction. Early detection and early initiation of parenteral vitamin B12 treatment are associated with a better prognosis and disease control (Chao et al. 2017).

Vitamin B12 deficiency due to malnutrition can result in irreversible developmental sequelae. Breast-fed infants from vegan/vegetarian mothers are at risk for vitamin B12 deficiency; this vitamin should be supplemented.

Methylmalonic aciduria is a genetic defect of cobalamin metabolism; several subtypes are known. Some subtypes respond to pharmacological doses of hydroxycobalamin, given intravenously or via intramuscular injection (typically 2–5 mg per day). The vitamin-responsive subtypes mostly have a better long-term outcome (Horster et al. 2007).

N-Carbamylglutamate in N-Acetylglutamatesynthase Deficiency

N-Acetylglutamatesynthase deficiency (estimated incidence, less than 1:2,000,000) is inherited as an autosomal recessive trait and leads to a lack of N-acetylglutamate which serves as a cofactor and allosteric activator of carbamoyl phosphate synthase (CPS1), the first enzymatic step in the urea cycle. N-Carbamylglutamate (carglumic acid, Carbaglu®, Orphan Europe, Paris France, 100–300 mg/kg per day) is a commercially available pharmacological formulation approved by EMA (Summar et al. 2013; Haberle et al. 2012; Haberle 2011).

In some organic acidurias (e.g., methylmalonic aciduria), secondary inhibition of N-acetylglutamatesynthase can occur under poor metabolic control. N-Carbamylglutamate (Carbaglu®) can (partially) compensate this deficiency.

Miscellaneous

Several inborn errors of metabolism respond to vitamins as cofactors of the compromised enzyme. For example, some subtypes of homocystinuria benefit from vitamin B6 and folate, and patients with thiamine-responsive megaloblastic anemia improve under pharmacological therapy with vitamin B1.

Activation of Alternative Pathways for the Elimination of Toxic Compounds

Butyrate and Benzoate in Urea Cycle Defects

Urea cycle defects (UCD) are potentially life-threatening inborn errors of metabolism affecting one of the enzymes involved in the urea cycle. They lead to severe hyperammonemia resulting in encephalopathy, especially during catabolic spells.

Benzoate is conjugated with glycine forming hippurate with one N-atom which, in contrast to ammonia, is water-soluble. Similarly, phenylbutyrate is metabolized to phenylacetate in the liver which then binds glutamine containing two N-atoms. Phenylacetylglutamine is water-soluble; thus two N-atoms can be excreted (Fig. 3). In Europe, no pharmacological benzoate formulation is available; only off-label use of the chemical compound is possible. Sodium phenylbutyrate is available as an oral commercial pharmacological formulation (Ammonaps®); recommended dose is 250–500 mg/kg per day divided in 3–6 doses. Palatability has been improved by carbohydrate coating (Pheburane®). Recently, an alternative, more palatable liquid compound, glycerol phenylbutyrate (Ravicti®), was approved by EMA (Diaz et al. 2013). If required, phenylacetate can be given intravenously; if liver function is intact, this compound is metabolized to phenylbutyrate.

Benzoate in Nonketotic Hyperglycinemia

Nonketotic hyperglycinemia (NKH) is an inherited defect of the glycine cleavage system leading to the accumulation of glycine with a

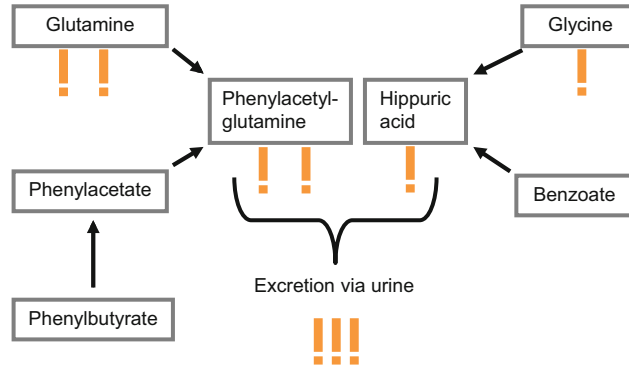


Fig. 3 Use of scavengers in urea cycle defects: Phenylacetate binds glutamine with two N-atoms (orange) with the formation of phenylacetylglutamate which is water-soluble. Benzoate bind glycine (one N-atom) resulting in the production of hippurate which is water-soluble and can be excreted via urine

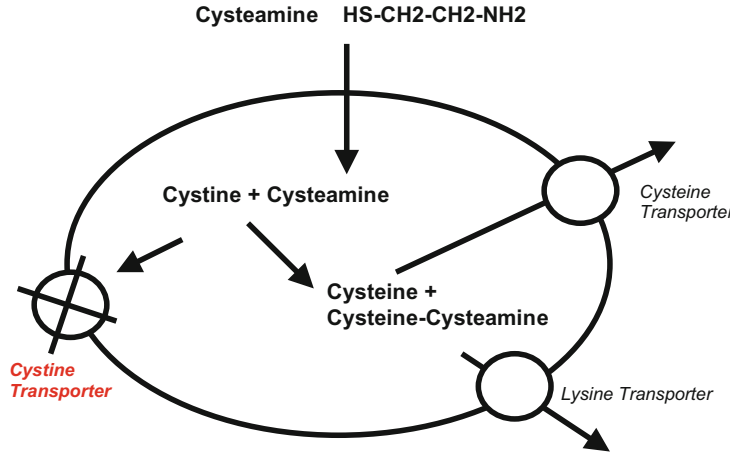


Fig. 4 Cystinosis is due to deficiency of the cystine transporter. Cysteamine binds cystine leading to cysteine and cysteine-cysteamine which are transported out of the lysosomes via the cysteine transporter and the lysine transporter, respectively

local production of glycine in the CNS. The CNS/plasma ratio is elevated. Clinically, the patients suffer from encephalopathy with severe epilepsy and psychomotor retardation. As mentioned above (section "Butyrate and Benzoate in Urea Cycle Defects"), benzoate can bind glycine leading to water-soluble hippurate which can be eliminated via urine. Clinically, this therapeutic option has only a minor effect in most patients (Boneh et al. 2008).

Cysteamine in Cystinosis

Cystinosis is an autosomal recessive storage disease due to mutations of the cystinosin gene which results in the dysfunction of the lysosomal transporter cystinosin. As a result of this dysfunction, cystine accumulates in lysosomes leading to clinical symptoms like nephropathy (predominantly tubulopathy) mostly culminating in renal failure, growth retardation, corneal crystals, endocrinological abnormalities, sometimes hepatosplenomegaly, severe myopathy, and CNS symptoms in adulthood (for a recent review, see Veys et al. (2017)).

Treatment consists of activating alternative pathways for the transport of cystine out of the lysosomes (Fig. 4). Since the mid-1990s, immediate-release cysteamine bitartrate (Cystagon®) at a dose of 1.3 g/m^2 per day (divided in four doses per day) was the only therapeutic option for many years delaying the progress of disease. A serious side effect of this therapy is halitosis and poor taste which hampers compliance. In the last decade, a new extended-release formulation of cysteamine bitartrate (PROCYSBI®) was developed which shall be taken at 80% of the daily Cystagon® dose, divided in two servings. Immediate- and delayed-release preparations were compared in a crossover study (Langman et al. 2012). Pharmacokinetic data have been measured in the

crossover study, $t_{1/2}$ was 254 min in the delayed-release preparation; Cmax 3.7 mg/l and $AUC_{0\text{-}12h}$ were 739 min x mg/l compared to 90 min, 2.7 mg/l and 357 min x mg/l in the immediate-release preparation, respectively. The new formulation has improved compliance; halitosis seems to be reduced (Franke et al. 2017); however, this formulation is much more expensive than the immediate-release product.

The cystine-depleting therapy is monitored by analyzing cystine content in leukocytes using high-performance liquid chromatography.

Oral cystine-depleting medication does not affect formation of corneal cystine crystals in the eyes. Therefore, topical cysteamine hydrochloride preparations have to be administered to control ophthalmological symptoms. Cystaran® eye drops containing 0.44% cysteamine hydrochloride are approved in the USA, while Cystadrops® (0.55% cysteamine hydrochloride) was approved in 2016 in Europe by the EMA (Labbe et al. 2014).

Enzyme Augmentation by Chaperones

Sapropterin Dihydrochloride and Phenylketonuria

Tetrahydrobiopterin (6R-tetrahydrobiopterin, BH4) is the natural cofactor of phenylalanine hydroxylase, the enzyme deficient in classical phenylketonuria (PKU). Some mutants in mild PKU can be stabilized by BH4 leading to an increase in the residual activity of phenylalanine hydroxylase, hence an increased protein tolerance. This can have a positive impact on the quality of life in these patients (Burton et al. 2010). BH4 responsiveness should be determined individually with respect to improvement in biochemical control and increase in natural protein intake. Thus, BH4 acts as a chaperone. Sapropterin dihydrochloride is now available as a pharmaceutical formulation (Kuvan®, marketed by BioMarin) of the natural cofactor BH4; approval of the drug for patients younger than 4 years has only recently been granted by EMA. In young children, elimination half-life was calculated at about 1 h, absorption half-life was 3 h, absorption was rate-limiting for disposition, and once-daily dosing was reported to be adequate (Muntau et al. 2017). The recommended dose is 10 mg/kg once daily. While blood phenylalanine levels can be reduced by about 30% in amenable mutations, about 20% of PKU patients respond to this therapeutic option (Heintz et al. 2013). This compound is tolerated well; there are no serious short-term side effects (Longo et al. 2015); however, long-term safety has still to be assessed as BH4 is also a cofactor for other enzymes (e.g., NO synthase).

The cost-effectiveness of Kuvan® is not established, especially when dietary treatment and phenylalanine-free L-amino acid supplements are still required (Longo et al. 2015; Shintaku and Ohura 2014); a detailed health technology assessment is required.

Migalastat and Fabry Disease

Fabry disease (also known as Anderson disease, OMIM 301500) is an X-linked lysosomal storage disorder caused by a deficiency in the activity of alpha-galactosidase A. Symptoms are usually more severe in males compared to females based on X-chromosomal inheritance. Clinically, M. Fabry is a multisystemic disease affecting the skin, heart, kidney, brain, peripheral nervous system, gastrointestinal tract, ear, eye, endocrinological system, etc. Globotriaosylceramide (Gb3) is the accumulating substance and can be used as a biomarker, together with lyso-Gb3 (globotriaosylsphingosine) which can be measured in blood and/or urine. Typical clinical symptoms are angiokeratoma, acroparesthesia, hypohydrosis, cornea verticillata, hearing difficulties, vertigo, progressive vascular disease of the heart and kidney, hypothyroidism, diarrhea, and psychiatric symptoms; however, in some patients, clinical manifestations are limited to the heart or kidney.

Since 2001, enzyme replacement therapy (Fabrazyme® and Replagal®) is available (see section "Enzyme Replacement Therapy (ERT)") which leads to improvement of symptoms in many, but not all, patients; however, symptoms do not disappear completely. The brain is not accessible to this treatment option based on the blood-brain barrier (e.g., Lidove et al. 2016). This

therapeutic intervention is challenging, intravenous therapy has to be given every other week over several hours, and immunological processes may lead to allergic or even anaphylactic reactions. In some patients, neutralizing antibodies may form which renders therapy ineffective. Based on a short half-live of the infused enzyme, fluctuation of clinical symptoms is often reported by the patients.

Some mutations are amenable to oral pharmacological chaperone therapy by migalastat (1-deoxygalactonojirimycin, Galafold®) by stabilizing protein folding. The mutated, misfolded enzyme is retained in the endoplasmic reticulum and prematurely degraded. Intracellular processing can be restored by migalastat in some mutations (Ishii et al. 2007; Ishii 2012) where proper folding can be induced. Preliminary studies and pharmacokinetic and safety studies were performed in healthy volunteers showing that migalastat is rapidly absorbed and up to 60% of migalastat can be recovered unchanged in urine, the half-life was approximately 4 h in plasma, and no serious side effects were observed (Johnson et al. 2013).

Subsequent phase 2 studies on safety and pharmacodynamic effects of migalastat were performed in adult male patients demonstrating safety and efficacy in terms of increased alpha-galactosidase activity and reduced Gb3 levels not only in plasma and urine but also in the skin and kidney (Germain et al. 2012). Based on a lysosomal half-life of 110–120 h (Benjamin et al. 2009), migalastat was given every other day. Another phase 2 study was performed in females with amenable and non-amenable mutations showing that migalastat is well tolerated and shows efficacy in patients with amenable mutations (Giugliani et al. 2013).

Galafold® has been approved by the European Medicines Agency (EMA) in 2016 for the use in patients with Fabry disease older than 16 years carrying an amenable mutation. Amenability of mutations is based on in vitro testing; a continuously updated list on amenable mutations can be found at www.GalafoldAmenabilityTable.com.

Despite these advances in therapy of Fabry patients, there are limitations of chemical chaperones in clinical practice. They are mutation specific and therefore only selected patients are responsive. Furthermore, chemical chaperones bind to the catalytic site of the enzymes with the risk that they impair activity instead of enhancing it if given at high doses and/or more often than required.

Enzyme Replacement Therapy (ERT)

Enzyme Replacement Therapy (ERT) in Selected Lysosomal Storage Diseases

Deficiency of lysosomal enzymes will lead to the accumulation of complex lysosomal molecules like glycogen in patients suffering from Pompe disease, glycosaminoglycans in cases of mucopolysaccharidoses (for instance, Morbus Hurler, MPS 1), glycoproteins in patients suffering from oligosaccharidoses (like Morbus Fabry), and sphingolipids in cases of Niemann-Pick disease types A and B, Gaucher disease, as well as cerebral lysosomal diseases like Tay-Sachs disease, Krabbe disease, and metachromatic leukodystrophy (Ferreira and Gahl 2017).

Sebelipase alfa (Kanuma®) is a recombinant human lysosomal acid lipase (LAL) approved for the treatment of LAL deficiency (Wolman disease) (Frampton 2016).

Asfotase alfa (Strensiq®) is a human recombinant enzyme replacement therapy for patients suffering from hypophosphatasia (Whyte 2017).

The goal of ERT is to compensate metabolic defects in patients with lysosomal storage disease by regular intravenous infusions of recombinant enzymes (every week or every other week depending on the underlying disorder). Classically, the enzyme is produced in bioreactors using human fibroblasts or Chinese hamster ovary cells as starter cultures. Biotechnologically produced enzymes are coupled to mannose 6-phosphate or mannose 6-phosphate is exposed as a signal substance. By virtue of the mannose-6-phosphate receptor at the cell surface, the intravenously applied enzymes can be taken up by cells and transported to the lysosomes where they take over catalytic functions. Storage material will be degraded by this treatment leading to clinical

improvement of some but not all organ functions. The central nervous system does not benefit from intravenous ERT due to the blood-brain barrier; bradytrophic tissues poorly respond to ERT. In addition, the therapeutic effect may be hampered by the development of antibodies against recombinant enzymes leading to complete inefficacy of ERT (Banugaria et al. 2011).

To deliver recombinant enzyme to the brain, intrathecal infusion was performed in a phase 1/phase 2 study in 12 patients with MPS IIIA (Sanfilippo disease A) (Jones et al. 2016); results were not convincing. Alternatively to this invasive procedure, enzymes have been modified to facilitate crossing of the blood-brain barrier by fusion to a monoclonal antibody against the human insulin receptor, called HIRMAb (Boado et al. 2014).

In the last years, enzyme replacement therapy has become available for a number of lysosomal storage diseases: imiglucerase (Cerezyme®), taliglucerase (Elelyso®, not approved in Europe), and velaglucerase (VIPRIV®) for Gaucher disease type 1, laronidase (Aldurazyme®) for Hurler disease (MPS type 1), idursulfase (Elaprase®) for Hunter disease (MPS type 2), elosulfase alfa (Vimizim®) for Morquio disease (MPS type 4), galsulfase (Naglazyme®) for Maroteaux-Lamy disease (MPS type 6), alglucosidase alfa (Myozyme®) for Pompe disease (glycogen storage disease type 2), and agalsidase alfa (Replagal®) and beta (Fabrazyme®) for Fabry disease. For patients suffering from Niemann-Pick disease types A and B, a recombinant human acid sphingomyelinase (olipudase alpha) is currently tested in clinical trials as an intravenous enzyme replacement therapy (Wasserstein et al. 2015).

ERT in lysosomal storage disorders is costly and requires a multi-professional approach which makes this therapeutic option logistically and economically challenging (Das et al. 2017).

Patients with lysosomal storage disorders are generally normal at birth, with symptoms developing in the first year(s) of life.

Many different cell types and tissues are affected by lysosomal storage disorders, with involvement at different stages in the disease process. ERT cannot cure patients but can stabilize organ function or slow progression.

A new biotechnology for large-scale production of enzymes that can be used for ERT is moss-based. Alpha-galactosidase A produced by the moss *Physcomitrella patens* was successfully used as an ERT in Fabry mice (Reski et al. 2015). Cellular uptake of the deficient enzyme occurs via mannose receptors after mannose-terminated enzymes bind to these receptors. Enzyme concentrations of 10 μg/ml were used; dose-dependent uptake of the enzyme up to 40 μg/ml was observed. One advantage of moss-based ERT is a higher homogeneity and reproducibility of enzyme glycosylation compared to mammalian cell-based ERT (Shen et al. 2016).

Compared with ERT, substrate reduction therapy or chaperone therapy in lysosomal storage diseases (see above) has some advantages as these small molecules can be given orally, they do not generate immune reactions, and they have the potential to cross the blood-brain barrier. However, this therapeutic option is currently only available for M. Gaucher type 1 (eliglustat) and Morbus Niemann-Pick type C (miglustat) and as a chaperone therapy in M. Fabry (migalastat).

Enzyme Replacement Therapy by Phenylalanine Ammonia Lyase in Phenylketonuria

As mentioned above, patients with classical PKU have a compromised activity of phenylalanine hydroxylase, the enzyme that catalyzes the irreversible conversion of phenylalanine to tyrosine. In the absence of treatment, systemic phenylalanine concentrations can increase to neurotoxic levels and impair cognitive development. The treatment of PKU requires lifelong selective reduction of phenylalanine intake and an adequate dietary supply of tyrosine by a protein-restricted diet supplemented with special amino acid mixtures devoid of phenylalanine and enriched with tyrosine. Adherence to this restricted diet is often challenging.

A possible enzyme 'replacement' therapy using PEG phenylalanine ammonia-lyase (PEG-PAL) or Pegvaliase® subcutaneously is under investigation.

This enzyme converts phenylalanine to ammonia and trans-cinnamic acid. PEG-PAL clinical phase 2 trials have proven short-term reduction in the blood phenylalanine concentrations in adult PKU patients, but further studies are required to assess long-term efficacy and safety. Results of a phase 3 extension study (NCT01819727) are awaited. Therefore, phenylalanine ammonia lyase as a non-mammalian protein providing alternative phenylalanine metabolism has the potential to relief dietary restrictions and secure a better quality of life in the long run; however, tyrosine remains an essential amino acid which has to be supplemented (Longo et al. 2014; van Wegberg et al. 2017b).

Outlook

Taken together, different therapeutic principles can improve clinical outcome in patients suffering from IEM. These options interfere with the pathophysiology of disease based on the underlying biochemical defect. By better understanding the pathobiochemistry of IEMs, more therapeutic options will become available. Individualized medicine in the field of IEM will improve efficacy and reduce side effects.

As IEMs are monogenetic diseases, genetic correction will be a further milestone beyond pharmacological therapy.

References and Further Reading

Al Hafid N, Christodoulou J (2015) Phenylketonuria: a review of current and future treatments. Transl Pediatr 4(4):304–317. https://doi.org/10.3978/j.issn.2224-4336.2015.10.07

Banugaria SG, Prater SN, Ng YK, Kobori JA, Finkel RS, Ladda RL, Chen YT, Rosenberg AS, Kishnani PS (2011) The impact of antibodies on clinical outcomes in diseases treated with therapeutic protein: lessons learned from infantile Pompe disease. Genet Med 13(8):729–736. https://doi.org/10.1097/GIM.0b013e3182174703

Bendadi F, de Koning TJ, Visser G, Prinsen HC, de Sain MG, Verhoeven-Duif N, Sinnema G, van Spronsen FJ, van Hasselt PM (2014) Impaired cognitive functioning in patients with tyrosinemia type I receiving nitisinone. J Pediatr 164(2):398–401. https://doi.org/10.1016/j.jpeds.2013.10.001

Benjamin ER, Flanagan JJ, Schilling A, Chang HH, Agarwal L, Katz E, Wu X, Pine C, Wustman B, Desnick RJ, Lockhart DJ, Valenzano KJ (2009) The pharmacological chaperone 1-deoxygalactonojirimycin increases alpha-galactosidase A levels in Fabry patient cell lines. J Inherit Metab Dis 32(3):424–440. https://doi.org/10.1007/s10545-009-1077-0

Bennett LL, Mohan D (2013) Gaucher disease and its treatment options. Ann Pharmacother 47(9):1182–1193. https://doi.org/10.1177/1060028013500469

Bickel H, Gerrard J, Hickmans EM (1954) The influence of phenylalanine intake on the chemistry and behaviour of a phenyl-ketonuric child. Acta Paediatr 43(1):64–77

Blau N, van Spronsen FJ, Levy HL (2010) Phenylketonuria. Lancet 376(9750):1417–1427. https://doi.org/10.1016/S0140-6736(10)60961-0

Boado BJ, Ka-Wai Hui E, Lu JZ, Pardridge WM, (2014) Insulin receptor antibody-iduronate 2-sulfatase fusion protein: Pharmacokinetics, anti-drug antibody, and safety pharmacology in Rhesus monkeys. Biotechnol Bioeng 111(11):2317–2325

Boneh A, Allan S, Mendelson D, Spriggs M, Gillam LH, Korman SH (2008) Clinical, ethical and legal considerations in the treatment of newborns with non-ketotic hyperglycinaemia. Mol Genet Metab 94(2):143–147. https://doi.org/10.1016/j.ymgme.2008.02.010

Burton BK, Bausell H, Katz R, Laduca H, Sullivan C (2010) Sapropterin therapy increases stability of blood phenylalanine levels in patients with BH4-responsive phenylketonuria (PKU). Mol Genet Metab 101(2–3):110–114. https://doi.org/10.1016/j.ymgme.2010.06.015

Chao MM, Illsinger S, Yoshimi A, Das AM, Kratz CP (2017) Congenital Transcobalamin II deficiency: a rare entity with a broad differential. Klin Padiatr 229(6):335–357. https://doi.org/10.1055/s-0043-120266

Christ SE (2003) Asbjorn Folling and the discovery of phenylketonuria. J Hist Neurosci 12(1):44–54

Das AM, Lagler F, Beck M, Scarpa M, Lampe C (2017) Lysosomal storage diseases: challenges in multiprofessional patient care with enzyme replacement therapy. Klin Padiatr 229(3):168–174. https://doi.org/10.1055/s-0043-103088

Diaz GA, Krivitzky LS, Mokhtarani M, Rhead W, Bartley J, Feigenbaum A, Longo N, Berquist W, Berry SA, Gallagher R, Lichter-Konecki U, Bartholomew D, Harding CO, Cederbaum S, McCandless SE, Smith W, Vockley G, Bart SA, Korson MS, Kronn D, Zori R, Merritt JL 2nd, C S Nagamani S, Mauney J, Lemons C, Dickinson K, Moors TL, Coakley DF, Scharschmidt BF, Lee B (2013) Ammonia control and neurocognitive outcome among urea cycle disorder patients treated with glycerol phenylbutyrate. Hepatology 57(6):2171–2179. https://doi.org/10.1002/hep.26058

Ellis MK, Whitfield AC, Gowans LA, Auton TR, Provan WM, Lock EA, Smith LL (1995) Inhibition of 4-hydroxyphenylpyruvate dioxygenase by 2-(2-nitro-4-trifluoromethylbenzoyl)-cyclohexane-1,3-dione and 2-(2-chloro-4-methanesulfonylbenzoyl)-cyclohexane-

1,3-dione. Toxicol Appl Pharmacol 133(1):12–19. https://doi.org/10.1006/taap.1995.1121

Ferreira CR, Gahl WA (2017) Lysosomal storage diseases. Transl Sci Rare Dis 2(1–2):1–71. https://doi.org/10.3233/TRD-160005

Frampton JE (2016) Sebelipase alfa: a review in lysosomal acid lipase deficiency. Am J Cardiovasc Drugs 16(6): 461–468. https://doi.org/10.1007/s40256-016-0203-2

Franke D, Steffens R, Thomas L, Pavicic L, Ahlenstiel T, Pape L, Gellermann J, Müller D, Querfeld U, Haffner D, Zivicnjak M (2017) Kidney transplantation fails to provide adequate growth in children with chronic kidney disease born small for gestational age. Pediatr Nephrol 32(3):511–519. https://doi.org/10.1007/s00467-016-3503-5

Garcia MI, de la Parra A, Arias C, Arredondo M, Cabello JF (2017) Long-term cognitive functioning in individuals with tyrosinemia type 1 treated with nitisinone and protein-restricted diet. Mol Genet Metab Rep 11:12–16. https://doi.org/10.1016/j.ymgmr.2017.01.016

Garrod AE (1902) About Alkaptonuria. Med Chir Trans 85:69–78

Germain DP, Giugliani R, Hughes DA, Mehta A, Nicholls K, Barisoni L, Jennette CJ, Bragat A, Castelli J, Sitaraman S, Lockhart DJ, Boudes PF (2012) Safety and pharmacodynamic effects of a pharmacological chaperone on alpha-galactosidase a activity and globotriaosylceramide clearance in Fabry disease: report from two phase 2 clinical studies. Orphanet J Rare Dis 7:91. https://doi.org/10.1186/1750-1172-7-91

Giugliani R, Waldek S, Germain DP, Nicholls K, Bichet DG, Simosky JK, Bragat AC, Castelli JP, Benjamin ER, Boudes PF (2013) A phase 2 study of migalastat hydrochloride in females with Fabry disease: selection of population, safety and pharmacodynamic effects. Mol Genet Metab 109(1):86–92. https://doi.org/10.1016/j.ymgme.2013.01.009

Haberle J (2011) Role of carglumic acid in the treatment of acute hyperammonemia due to N-acetylglutamate synthase deficiency. Ther Clin Risk Manag 7:327–332. https://doi.org/10.2147/TCRM.S12703

Haberle J, Boddaert N, Burlina A, Chakrapani A, Dixon M, Huemer M, Karall D, Martinell D, Crespo PS, Santer R, Servais A, Valayannopoulos V, Linder M, Rubio V, Dionisi-Vici C (2012) Suggested guidelines for the diagnosis and management of urea cycle disorders. Orphanet J Rare Dis 7:32. https://doi.org/10.1186/1750-1172-7-32

Hall MG, Wilks MF, Provan WM, Eksborg S, Lumholtz B (2001) Pharmacokinetics and pharmacodynamics of NTBC (2-(2-nitro-4-fluoromethylbenzoyl)-1,3-cyclohexanedione) and mesotrione, inhibitors of 4-hydroxyphenyl pyruvate dioxygenase (HPPD) following a single dose to healthy male volunteers. Br J Clin Pharmacol 52(2):169–177

Heintz C, Cotton RG, Blau N (2013) Tetrahydrobiopterin, its mode of action on phenylalanine hydroxylase, and importance of genotypes for pharmacological therapy of phenylketonuria. Hum Mutat 34(7):927–936. https://doi.org/10.1002/humu.22320

Horster F, Baumgartner MR, Viardot C, Suormala T, Burgard P, Fowler B, Hoffmann GF, Garbade SF, Kölker S, Baumgartner ER (2007) Long-term outcome in methylmalonic acidurias is influenced by the underlying defect (mut0, Mut-, cblA, cblB). Pediatr Res 62(2):225–230. https://doi.org/10.1203/PDR.0b013e3180a0325f

Ishii S (2012) Pharmacological chaperone therapy for Fabry disease. Proc Jpn Acad Ser B Phys Biol Sci 88 (1):18–30

Ishii S, Chang HH, Kawasaki K, Yasuda K, Wu HL, Garman SC, Fan JQ (2007) Mutant alpha-galactosidase a enzymes identified in Fabry disease patients with residual enzyme activity: biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin. Biochem J 406(2):285–295. https://doi.org/10.1042/BJ20070479

Johnson FK, Mudd PN Jr, Bragat A, Adera M, Boudes P (2013) Pharmacokinetics and safety of Migalastat HCl and effects on Agalsidase activity in healthy volunteers. Clin Pharmacol Drug Dev 2(2):120–132. https://doi.org/10.1002/cpdd.1

Jones SA, Breen C, Heap F, Rust S, de Ruijter J, Tump E, Marchal JP, Qiu Y, Chung JK, Nair N, Haslett PA, Wijburg FA (2016) A phase 1/2 study of intrathecal heparan-N-sulfatase in patients with mucopolysaccharidosis IIIA. Mol Genet Metab 118(3):198–205. https://doi.org/10.1016/j.ymgme.2016.05.006

Kienstra NS, van Reemst HE, van Ginkel VG, Daly A, van Dam E, MacDonald A, Burgerhof JGM, de Blaauw P, McKiernan PJ, Rebecca Heiner-Fokkema M, van Spronsen FJ (2018) Daily variation of NTBC and its relation to succinylacetone in tyrosinemia type 1 patients comparing a single dose to two doses a day. J Inherit Metab Dis 41(2):181–186

Koch R, Burton B, Hoganson G, Peterson R, Rhead W, Rouse B, Scott R, Wolff J, Stern AM, Guttler F, Nelson M, de la Cruz F, Coldwell J, Erbe R, Geraghty MT, Shear C, Thomas J, Azen C (2002) Phenylketonuria in adulthood: a collaborative study. J Inherit Metab Dis 25(5):333–346

Labbe A, Baudouin C, Deschenes G, Loirat C, Charbit M, Guest G, Niaudet P (2014) A new gel formulation of topical cysteamine for the treatment of corneal cystine crystals in cystinosis: the Cystadrops OCT-1 study. Mol Genet Metab 111(3):314–320. https://doi.org/10.1016/j.ymgme.2013.12.298

Langman CB, Greenbaum LA, Sarwal M, Grimm P, Niaudet P, Deschênes G, Cornelissen E, Morin D, Cochat P, Matossian D, Gaillard S, Bagger MJ, Rioux P (2012) A randomized controlled crossover trial with delayed-release cysteamine bitartrate in nephropathic cystinosis: effectiveness on white blood cell cystine levels and comparison of safety. Clin J Am Soc Nephrol 7(7):1112–1120. https://doi.org/10.2215/CJN.12321211

Lenke RR, Levy HL (1980) Maternal phenylketonuria and hyperphenylalaninemia. An international survey of the outcome of untreated and treated pregnancies. N Engl J Med 303(21):1202–1208. https://doi.org/10.1056/NEJM198011203032104

Lidove O, Barbey F, Joly D (2016) Treatment of Fabry disease: successes, failures, and expectations. Nephrol Ther 12(Suppl 1):105–113. https://doi.org/10.1016/j.nephro.2016.02.003

Longo N, Harding CO, Burton BK, Grange DK, Vockley J, Wasserstein M, Rice GM, Dorenbaum A, Neuenburg JK, Musson DG, Gu Z, Sile S (2014) Single-dose, subcutaneous recombinant phenylalanine ammonia lyase conjugated with polyethylene glycol in adult patients with phenylketonuria: an open-label, multicentre, phase 1 dose-escalation trial. Lancet 384 (9937):37–44. https://doi.org/10.1016/S0140-6736(13)61841-3

Longo, N., Arnold, G. L., Pridjian, G., Enns, G. M., Ficicioglu, C., Parker, S.,. .. Safety, R. (2015). Long-term safety and efficacy of sapropterin: the PKUDOS registry experience. Mol Genet Metab, 114(4), 557–563. https://doi.org/10.1016/j.ymgme.2015.02.003

Lukina E, Watman N, Arreguin EA, Dragosky M, Iastrebner M, Rosenbaum H, Phillips M, Pastores GM, Kamath RS, Rosenthal DI, Kaper M, Singh T, Puga AC, Peterschmitt MJ (2010a) Improvement in hematological, visceral, and skeletal manifestations of Gaucher disease type 1 with oral eliglustat tartrate (Genz-112638) treatment: 2-year results of a phase 2 study. Blood 116(20):4095–4098. https://doi.org/10.1182/blood-2010-06-293902

Lukina E, Watman N, Arreguin EA, Banikazemi M, Dragosky M, Iastrebner M, Rosenbaum H, Phillips M, Pastores GM, Rosenthal DI, Kaper M, Singh T, Puga AC, Bonate PL, Peterschmitt MJ (2010b) Improvement in hematological, visceral, and skeletal manifestations of Gaucher disease type 1 with oral eliglustat tartrate (Genz-112638) treatment: 2-year results of a phase 2 study. Blood 116(20):4095–4098. https://doi.org/10.1182/blood-2010-06-293902

MacDonald A, Rylance G, Hall SK, Asplin D, Booth IW (1996) Factors affecting the variation in plasma phenylalanine in patients with phenylketonuria on diet. Arch Dis Child 74(5):412–417

Macleod EL, Ney DM (2010) Nutritional Management of Phenylketonuria. Ann Nestle Eng 68(2):58–69. https://doi.org/10.1159/000312813

Mayorandan, S., Meyer, U., Gokcay, G., Segarra, N. G., de Baulny, H. O., van Spronsen, F.,. .. Das, A. M. (2014a). Cross-sectional study of 168 patients with hepatorenal tyrosinaemia and implications for clinical practice. Orphanet J Rare Dis, 9, 107. https://doi.org/10.1186/s13023-014-0107-7

Mayorandan S, Meyer U, Hartmann H, Das AM (2014b) Glycogen storage disease type III: modified Atkins diet improves myopathy. Orphanet J Rare Dis 9:196. https://doi.org/10.1186/s13023-014-0196-3

Muntau AC, Burlina A, Eyskens F, Freisinger P, De Laet C, Leuzzi V, Rutsch F, Serap Sivri H, Vijay S, Bal MO, Gramer G, Pazdírková R, Cleary M, Lotz-Havla AS, Munafo A, Mould DR, Moreau-Stucker F, Rogoff D (2017) Efficacy, safety and population pharmacokinetics of sapropterin in PKU patients <4 years: results from the SPARK open-label, multicentre, randomized phase IIIb trial. Orphanet J Rare Dis 12(1):47. https://doi.org/10.1186/s13023-017-0600-x

Reski R, Parsons J, Decker EL (2015) Moss-made pharmaceuticals: from bench to bedside. Plant Biotechnol J 13(8):1191–1198. https://doi.org/10.1111/pbi.12401

Sander J, Janzen N, Terhardt M, Sander S, Gokcay G, Demirkol M, Ozer I, Peter M, Das AM (2011) Monitoring tyrosinaemia type I: blood spot test for nitisinone (NTBC). Clin Chim Acta 412(1–2):134–138. https://doi.org/10.1016/j.cca.2010.09.027

Schlune A, Thimm E, Herebian D, Spiekerkoetter U (2012) Single dose NTBC-treatment of hereditary tyrosinemia type I. J Inherit Metab Dis 35(5):831–836. https://doi.org/10.1007/s10545-012-9450-9

Shen J-S, Busch A, Day TS, Meng X-L, Chun IY, Dabrowska-Schlepp P, Fode B, Niederkrüger H, Forni S, Chen S, Schiffmann R, Frischmuth T, Schaaf A (2016) Mannose receptor-mediated delivery of moss-made alpha-galactosidase a efficiently corrects enzyme deficiency in Fabry mice. J Inherit Metab Dis 39 (2):293–303. https://doi.org/10.1007/s10545-015-9886-9

Shintaku H, Ohura T (2014) Sapropterin is safe and effective in patients less than 4-years-old with BH4-responsive phenylalanine hydrolase deficiency. J Pediatr 165(6):1241–1244. https://doi.org/10.1016/j.jpeds.2014.08.003

Summar ML, Koelker S, Freedenberg D, Le Mons C, Haberle J, Lee HS, Kirmse B (2013) The incidence of urea cycle disorders. Mol Genet Metab 110 (1–2):179–180. https://doi.org/10.1016/j.ymgme.2013.07.008

van Spronsen FJ, Derks TG (2014) Recombinant phenylalanine ammonia lyase in phenylketonuria. Lancet 384 (9937):6–8. https://doi.org/10.1016/S0140-6736(13)62075-9

van Wegberg AMJ, MacDonald A, Ahring K, Belanger-Quintana A, Blau N, Bosch AM, Burlina A, Campistol J, Feillet F, Gizewska M, Huijbregts SC, Kearney S, Leuzzi V, Maillot F, Muntau AC, van Rijn M, Trefz F, Walter JH, van Spronsen FJ (2017a) The complete European guidelines on phenylketonuria: diagnosis and treatment. Orphanet J Rare Dis 12(1):162. https://doi.org/10.1186/s13023-017-0685-2

van Wegberg AMJ, MacDonald A, Ahring K, Bélanger-Quintana A, Blau N, Bosch AM, Burlina A, Campistol J, Feillet F, Giżewska M, Huijbregts SC, Kearney S, Leuzzi V, Maillot F, Muntau AC, van Rijn M, Trefz F, Walter JH, van Spronsen FJ (2017b) The complete European guidelines on phenylketonuria: diagnosis and treatment. Orphanet J Rare Dis 12(1):162. https://doi.org/10.1186/s13023-017-0685-2

Vanier MT (2010) Niemann-pick disease type C. Orphanet J Rare Dis 5:16. https://doi.org/10.1186/1750-1172-5-16

Veys KR, Elmonem MA, Arcolino FO, van den Heuvel L, Levtchenko E (2017) Nephropathic cystinosis: an update. Curr Opin Pediatr 29(2):168–178. https://doi.org/10.1097/MOP.0000000000000462

Wasserstein MP, Jones SA, Soran H, Diaz GA, Lippa N, Thurberg BL, Culm-Merdek K, Shamiyeh E, Inguilizan H, Cox GF, Puga AC (2015) Successful within-patient dose escalation of olipudase alfa in acid sphingomyelinase

deficiency. Mol Genet Metab 116(1–2):88–97. https://doi.org/10.1016/j.ymgme.2015.05.013
Whyte MP (2017) Hypophosphatasia: enzyme replacement therapy brings new opportunities and new challenges. J Bone Miner Res 32(4):667–675. https://doi.org/10.1002/jbmr.3075
Williamson L (2014) Counterfeit drugs are everyone's business. Krankenpfl Soins Infirm 107(7):72–73
Wolf B (1993) Biotinidase Deficiency. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A (eds) GeneReviews((R)). University of Washington, Seattle

Clinical Studies in Infants (Pediatric Pharmacology)

19

Karel Allegaert

Contents

Abstract

Treatment of children with effective and safe medicines is crucial to improve their outcome. Despite this relevance, it is still common practice in children to administer medicines outside their market authorization. Even if authorized, pediatric medicines may not be age-appropriate for a broad range of therapeutic areas. This has been recognized as very unsatisfactory by all stakeholders involved and makes clinical pharmacological studies in children an obvious need.

However, clinical trials of medicines in children come with their specific burdens. These burdens can be qualified as either related to the specific aspects of pediatric pharmacokinetics (PK) and pharmacodynamics (PD) or relate to the logistics of clinical trials of medicines in children. This is followed by a stakeholder's analysis, discussing specific aspects related to parents and their children (International Children's Advisory Network, iCAN), recruitment challenges, and research capacity building. We hereby tried to focus on recent evolutions, including initiatives to further develop this research capacity (Institute for Advanced Clinical Trials, iACT for children; Innovative Medicines Initiative, IMI2).

K. Allegaert (✉)
Intensive Care and Department of Pediatric Surgery, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands

Department of Development and Regeneration, KU Leuven, Leuven, Belgium
e-mail: karel.allegaert@uzleuven.be

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_44

Perhaps progress is slower than anticipated, but pediatric medicines research is evolving, and we should keep this momentum. A further structured collaboration between the different stakeholders involved (the society, parents and children, sponsors, regulatory authorities) at the international level is crucial to use the available, but limited, resources as effective as possible to further improve pharmaceutical care in children.

On the Knowledge Gap and the Initiatives to Close This Gap

Similar to adults, the treatment of children with effective and safe medicines is crucial to improve their outcome or protect their quality of life. Despite this relevance, it is still common practice in children to administer medicines outside their market authorization. Health care professionals routinely prescribe medicines to children off label or unlicensed, hereby using dosing regimens initially developed for adults, and extrapolating from indications initially validated in adults and based on adult – not necessary similar to pediatric – pathophysiology. As a first effort to raise awareness for this issue that is perhaps specific but not limited to pediatrics (e.g., pregnancy, geriatrics as other "special" populations), the term "therapeutic orphan" was introduced in the late 1960s by Harry Shirkey. He reported about the situation where many medicines were not labeled for prescription to children due to insufficient and inaccurate clinical evidence being available, covering aspects including dosage, efficacy but also safety. The term "therapeutic orphan" hereby refers to the deprivation of children to not have access to medicines because these have not been adequately tested in children.

Actually, to a large extent and especially in specific pediatric subpopulations like newborns or infants, this situation still exists. Even if authorized, pediatric medicines may not be age-appropriate for a broad range of therapeutic areas. To illustrate this, we refer to the study of van Riet-Nales et al. on the availability and age-appropriateness of medicines authorized for children in The Netherlands. Based on a systematic search of the national medicines database and the summary of product characteristics, 3542 pediatric medicines (703 active chemical entities) were identified (van Riet-Nales et al. 2011). This is about half of the medicines and chemical entities available for human use. Of these authorized products, about 35%, 48%, 64%, 80%, and 95% were *authorized* for use in the consecutive pediatric age categories (0–27 days, 1–23 months, 2–5 years, 6–11 years, and 12–17 years). When the *suitability of the dosage form* was also considered, this was further reduced to 27%, 37%, 52%, 70%, and 88%. This reflects both the relevance of the pediatric subpopulations and the need to develop age-appropriate formulations (van Riet-Nales et al. 2011).

As a consequence, health care professionals lack the crucial information and access to knowledge on these aspects to make the best possible, informed decision for their pediatric patients. This has resulted in extensive *unlicensed* or *off-label use* of medicines in children. This is not limited to the intensive care setting (neonatal or pediatric intensive care) or the general pediatric hospital environment, but also occurs in the outpatient or ambulatory setting. In an attempt to quantify the pediatric off-label medicines use in different areas, Mühlbauer et al. estimated that 90%, 30–70%, and 10–20% of the prescriptions were off label in the intensive care unit setting, pediatric inpatients, and outpatients, respectively (Mühlbauer et al. 2009). In the most recent meta-analysis on unlicensed and off-label medicine prescription practices in pediatric hospitals (2015), 829 studies published between 1994 and 2012 were evaluated. Unlicensed and off-label medicines prescriptions ranged from 0.2% to 48% and from 12% to 71%, respectively (Magalhaes et al. 2015). Similar to the formulation-related observations of van Riet-Nales et al., this meta-analysis reconfirmed that this practice is still widespread, and that the youngest age category, i.e., (pre)term neonates are exposed most commonly to unlicensed or off-label medicines (100% exposed to at least one unlicensed or off-label medicine) (Magalhaes et al. 2015). At least, this provides further support

to improve the available knowledge and the label status for medicines administered to children.

If a medicine is used in a way that is different from that described in its license, this is called "unlicensed" or "off-license" use. This includes giving a medicine in any way or approach that is different from that described. This may involve crushing a capsule or tablet in order to turn it into a liquid to enable a toddler to take the medicine by oral route, or administer a liquid formulation intended for intravenous administration by oral route. *Off-label* medicines practice is defined as prescribing in a different manner to label recommendations in the summary of product characteristics in a given country. This may relate to age group, dose, frequency, formulation, administration route, indication, or contraindication for use in children. The term off label refers to the use of a medicine that is not included in the package inset and does not imply in itself improper, illegal, or inaccurate use of a medicine in itself. Yet, the practice of using medicines in an off-label manner or unlicensed potentially may result in suboptimal pharmacotherapy. Although the level of evidence about harm induced by off-label and unlicensed medicines use in children is rather limited and based on association type of studies, there are indications that harm does occur and that it is still very likely underreported. The incidence of adverse events related to either labeled compared to unlabeled use of medicines was estimated to the 3.9–6% for hospitalized patients and 1.4–3.4% of out of hospital patients, suggesting that the incidence doubles in the setting of unlabeled medicines prescriptions (Horen et al. 2002).

This setting has been perceived and recognized as very unsatisfactory and substandard of care by all stakeholders, including health care professionals working with children, parents and their children, politicians, and the pharmaceutical industry. In order to move and improve this setting, there have been a number of initiatives, some of which involve legislation (United States), whereas others have taken the form of regulation (Europe). Following the US initiatives (e.g., network of Pediatric Pharmacology Research units, Best Pharmaceuticals for Children Act [BPCA], Pediatric Research Equity Act [PREA]) initiated in the late 1990s, the European Union (better medicines for children, EU Pediatric Regulation 2007) also took initiatives (Hoppu et al. 2012). The European Paediatric Regulation (EC No. 1901/2006) focuses on three main objectives: increasing the number of appropriate medicines for children, increasing the available information on these medicines, and stimulating high-quality ethical research with children. The European Medicines Agency (EMA) published a 10-year report on the experience acquired as a result of the application of the pediatric regulation and a public consultation on this document ongoing (European Medicines Agency 2016a). Simultaneously, the World Health Organization (WHO) also became active in providing incentives to improve health care through improving knowledge and access to better, tailored medicines for children. This includes issues like child friendly, age-appropriate formulations, including child-size drug campaigns, excipients, but also a list of essential medicines for children. The sixth version of this WHO model list of essential medicines for children has been published in early 2017 (WHO 2017). Finally, the latest ICH E11 guidance (International Conference on Harmonization) on pediatric studies goes back to 2000, but – to the best of our knowledge – a revision effort is ongoing.

As a reflection of the proportional efforts made to generate knowledge on pharmacotherapy in children compared to adults, Table 1 provides an overview of the number of studies registered on the clinicaltrials.gov website by July 2017, using either no specific search criteria (all studies) or "child," "infant," or "newborn." For three regions of the western world (United States, Europe, and Canada) considered in this search, there is a consistent proportional contribution of studies in these specific subpopulations when compared to the overall volume of studies (United States 40–45%, Europe 20–25%, Canada 8–10%). However, there is still an overall low number of studies in infants (3%) and newborns (2%) compared to the total number of studies (National Institutes of Health 2017).

Based on these observations, we suggest that the federal US legislation and similar European

Table 1 Number of studies and proportion of pediatric studies as retrieved on the clinicaltrials.gov website on 15 July 2017. We hereby used either no specific search criteria (all studies) and compared those retrieved when either "child," "infant," or "newborn" were entered as an additional search option. For all three regions of the western world highlighted in this table, there is a consistent proportional contribution of studies in these specific pediatric subpopulations, with still an overall low number of studies in infants (3%) or newborns (2%) compared to the total number of pediatric (23%) studies retrieved in the registry (National Institutes of Health 2017)

	all studies	'child'	'infant'	'newborn'
worldwide	249 566	55 693 (23 %)	8 557 (3 %)	5 414 (2 %)
United States	103 342 (41 %)	23 583 (42 %)	3 574 (42 %)	2 024 (38 %)
Europe	70 269 (28 %)	12 091 (22 %)	2 016 (25 %)	1 534 (28 %)
Canada	17 142 (8 %)	4 161 (7 %)	780 (9 %)	406 (8 %)

initiatives (Table 1) have indeed resulted in a substantial increase in studies on medicines in children, with a subsequent significant increase in knowledge on pharmacotherapy in children, in part reflected in label changes. Interestingly, Schachter and Ramoni reported that the pediatric legislation resulted in a significant reduction in the time interval between the new medicine application approval (in adults) and the submission of the supplemental pediatric data (median decrease from 7.2 to 4.2 years), without significant reduction in the proportion of pediatric data submitted within the first year after adult approval (0–7%) (Schachter and Ramoni 2007). The potential to update product information based on already existing – but still unpublished – information should neither be underestimated. Under article 45 of the (European) Regulation, sponsors should submit existing pediatric studies to regulatory authorities for review and potential update of the product information. Based on nearly 19,000 study reports on about 1000 active substances, 262 substances have been assessed (62 centrally, 200 nationally approved) and reviewed. Based on this review, 16 new pediatric indications were added for a variety of indications (e.g., hypertension or congestive heart failure, spasticity, constipation, laryngitis subglottica, sedation and analgesia, infectious or inflammatory diseases, or contrast medium for diagnostics) (Saint-Raymond et al. 2016).

Unfortunately, only a few limited number of label changes covered medicine label changes for neonates, reflecting the fact that neonates remain underserved. To further illustrate this: on 406 pediatric label changes (FDA, 1997–2010), only 23 medicines resulted in 11 labeled indications (e.g., remifentanil, linezolid, rocuronium, nevirapine, sevoflurane, stavudine) in neonates. The absence of label change was most commonly due to unproven efficacy, despite the fact that these medicines (e.g., valganciclovir, paracetamol, caspofungin) are very likely also relevant for this population (Stiers and Ward 2014). This largely reflects the issue to proof efficacy, based on sufficient robust clinical end points or valid biomarkers in neonates.

In an additional effort to boost research on medicines in this subpopulation, the FDASIA (Food and Drug Administration Safety and Innovation Act) initiative aligned the US setting to a certain extent with the European Regulation by introducing the requirement that clinical trials also have to be performed in neonates when relevant to this population. At the same time, the FDASIA initiative made BPCA and PREA permanent and facilitated the availability of neonatal expertise within the FDA (Gonzalez et al. 2014). Another valuable initiative to improve the setting in neonates is the International Neonatal Consortium (INC) (Critical Path Institute 2015). INC integrates stakeholders to promote clinical medicine development for neonates. The consortium hereby focuses on generalizable methods to use data to support claims that a medicine is safe and efficacious when used to treat a specific indication (Turner et al. 2017).

Pharmacokinetics (PK) and Pharmacodynamics (PD) in Pediatric Studies: Aiming for a Moving Target

Preterms are not just small neonates
Neonates are not just small children
Children are not small adults

By virtue of their developmental and cognitive abilities, children are a vulnerable population. This is also true when participation in a clinical trial is considered. In addition to adaptations of the clinical protocol design (e.g., medicine formulation, dosing, sampling strategy, clinical indication), recruitment and retention strategies, incentives and the process to obtain informed consent must be modified. Although obviously needed as discussed in the introduction, clinical trials of medicines in children come with their specific burdens, as illustrated in Table 2 (Reed 2011; Allegaert and van den Anker 2015). These burdens can be qualified as either related to the specific aspects of pediatric pharmacokinetics (PK) and pharmacodynamics (PD) or relate to the logistics of clinical trials of medicines in children.

Simple extrapolation of PK or PD estimates initially documented in adults to pediatrics is obsolete, since both PK as well as PD processes differ considerably throughout childhood from the findings in adults because of maturational and nonmaturational changes. Throughout the pediatric life range, there is extensive maturation driven variability. This variability is also reflected in the use of different age categories for regulatory purposes and is commonly divided into the newborn period, infancy, childhood, and adolescence. Human growth and development consist of a sequence of physiologic events that link somatic growth with maturation. Weight gain hereby displays colinearity but is not similar to maturation. Across the pediatric age, both organ size and function change as well as body composition and (patho)physiology. We should be aware that these changes are most prominent – but not limited – to early infancy (Fig. 1). Across this pediatric life span, organ size and function change as does body composition, (patho)physiology, and ultimately cellular function. If we focus on weight changes to further illustrate this, there is an initial decrease (6–12%) in birth weight, with a subsequent increase of 50% in the first 6 weeks of postnatal life. Weight doubles in the first 3–4 months to result in a threefold higher weight at the end of infancy (Allegaert and van den Anker 2015). Consequently, total energy requirements change dramatically since these requirements are the sum of energy expenditure and energy deposition for growth (Fig. 1).

A rational approach to determine a safe and effective dose in an individual child necessitates understanding the PK and PD properties of a

Table 2 Burdens that may further complicate the conduct of clinical studies in pediatrics

Pharmacokinetics/pharmacodynamics in pediatric studies: aiming of a moving target
• Sampling strategy (number of samples, volume), specific analytical techniques
• Population pharmacokinetic modeling (mechanism, physiology based) not always sufficiently validated to support study design and sampling strategy, and uncertainty about extrapolation
• Extensive variability in pharmacokinetics and pharmacodynamics. This variability is not limited to maturational changes, but also covers nonmaturational covariates like, for example, disease severity, comedication, renal or hepatic impairment, or obesity
• How to assess efficacy? Robust and relevant pharmacodynamic endpoints and validated biomarkers are needed
• The need for specific pediatric formulations
• Safety: how to assess differences in patterns of (serious) adverse reactions in the pediatric population?
Circumstances related to pediatric studies
• Pediatric drug therapy development is a perceived "must," instead of an opportunity
• Ethics, parental consent, and assent in the child
• Study facilities (investigators, research facilities) should have experience with clinical studies in children
• Recruitment challenges, the need for multicenter collaborations
• Perceived risks and fear of negative outcomes, perceptions of society
• Drug development programs initially develop for other populations, only very rarely driven by pediatric needs
• How to secure adherence and retention in the trial?
• Healthy volunteers and pediatrics
• Natural course of pediatric diseases?

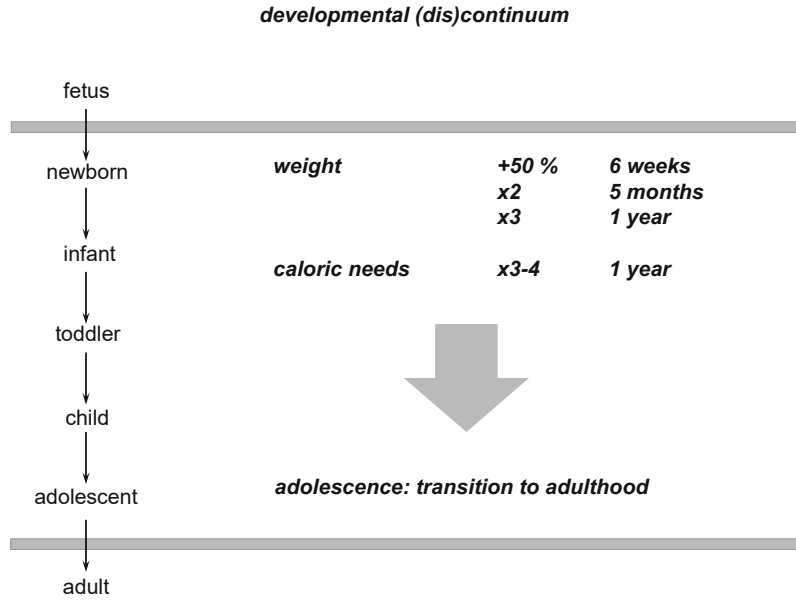


Fig. 1 Developmental discontinuum illustrated for the consecutive age subcategories in pediatrics

specific medicine, in combination with the clinical characteristics of a single, specific child. Developmental PK hereby represents the mathematical estimates of the concentration–time profile, while PD describes the relationship between a given concentration and the extent of a given response (e.g., pain relief, blood pressure, fever reduction, improvement in asthma score). Multiple differences related to maturation but also disease (obesity, renal or liver impairment) or therapeutic interventions (e.g., drug–drug interaction, diet) will result in differences in PK and probably also PD in children compared to adults. PK (***a***bsorption, ***d***istribution, and elimination by ***m***etabolism or renal ***e***limination, ***ADME***) hereby estimate the relationship between a concentration in a given compartment (e.g., plasma, central nervous system, subcutaneous tissue, bronchial tree) and time ("*what the body does to the medicine"*).

Absorption: Following oral administration, absorption displays extensive maturation because of gastro-enteral maturation (e.g., anatomy, motility, drug metabolism or transporters), but also nonenteral routes (e.g., cutaneous, muscular size, inhalation and circulation) display age-related changes. Examples of developmental absorption are illustrated in Table 3.

Distribution: Although a "theoretical volume," this depends on physical (e.g., extra- and intracellular water, lipophilic or water soluble compound, ionization and protein binding) and physiologic (protein binding, tissue uptake, permeation to deep compartments) processes. Consequently, the distribution volume is also driven by maturational changes and disease characteristics.

Metabolism: The drug metabolizing capacity is affected by multiple covariates. Besides growth and maturation, comorbidity, pharmacogenetics, and environmental issues can explain variability.

Excretion: The most relevant route of excretion is the renal route, both through glomerular filtration rate and renal tubular transport. These processes do not mature simultaneously.

PD estimates the relationship between a concentration and (side)-effects ("*what the medicine does to the body*"). Age-dependent differences in (un)anticipated effects of medicines may have a PK basis, a PD basis, or both. Differences in developmental pharmacology result in differences in medicine potency, efficacy, and/or toxicity. However, most of the variability observed relates to differences in PK, and maturational PD can

Table 3 Illustrations on the relevance of developmental changes in absorption in infants and children

Transcutaneous	The higher surface area and the higher permeability results in extensive absorption of iodine or corticosteroids following cutaneous application and subsequent endocrine disorders
	Patches are specific formulations, developed for a continuous, stable release. Manipulations (e.g., cutting) or cutaneous lesions (e.g., eczema, burned skin) may alter this disposition
Rectal	Compared to oral, rectal administration of paracetamol results in lower and less predictable absorption
Intra-vitreal	Intra-vitreal injection of bevacizumab for retinopathy of prematurity results in appearance in the systemic circulation
Inhalational	Nonbronchial steroid disposition may result in tongue hypertrophy, oral candidiasis, or systemic effects
Buccal	Buccal midazolam (Buccolam®) formulation (PUMA product), to be used in children with seizures
Swallowing	Liquids or mini-tablets are preferred formulations in the first year of life
Gastric pH and emptying	Higher bioavailability following oral penicillin administration in newborns compared to infants or children
	The peak concentration of a given compound, e.g., paracetamol is delayed and lower in infants compared to children
Intestinal enzymatic activity	First pass effect is lower and bioavailability higher following oral midazolam in (pre)term neonates because of lower intestinal (CYP3A) drug metabolism
Pancreas activity and bile	Reduced uptake of lipophilic drugs, fat-soluble vitamins, or enteral-hepatic recirculation
Comorbidity	Bioavailability after oral administration may be different in the setting of diarrhea, or critical illness

only be considered once the PK aspects have been taken into account. Table 4 provides some illustrations on developmental pharmacodynamics in neonates and infants, or children and adolescents.

Table 4 Illustrations on developmental pharmacodynamics in neonates and infants, or children and adolescents

Neonates and infants
• Dexamethasone and the risk for impaired neurodevelopmental outcome and cerebral palsy • Oxygenation saturation levels and the subsequent differences in mortality and morbidity (bronchopulmonary dysplasia, retinopathy of prematurity) in preterm neonates • The impact of ibuprofen or indomethacin on renal function • Exposure to antibiotics and its impact on weight gain and body composition • Neonatal hypo- and hyperglycemia, or hypothyroidism and neurocognitive outcome • Developmental toxicology for, for example, valproate (hepatic failure) or ifosfamide (renal tubular cell dysfunction, Fanconi) • The lymphocyte proliferation response is twofold lower in infants • Paradoxical seizures due to proportional excitatory GABA receptor overexpression
Children and adolescents
• Impact of steroids on growth • Impact of disease and treatment on normal pubertal development • Impact of either ethanol, nicotine, or other illicit drugs on neurodevelopmental and behavioral outcome • Limited to absent effects of tricyclic antidepressants because of a neurodevelopmental delay in the expression and activation of the norepinephrine system • Differences in oncological disease patterns between children compared to adolescents and young adults • The anticoagulant response to warfarin is higher in prepubertal children

The regulatory framework for pediatric medicine development in Europe and the United States provides some guidance on how this can be addressed (Manolis and Pons 2009) as illustrated in Fig. 2. Authorities usually consider information regarding: (i) how similar is disease progression between adults and children, (ii) how similar is the response to intervention between these populations, and (iii) which valid and relevant PD measurements (biomarkers, outcome variables) are available to decide on the type of product development program. If it is reasonable to assume that there is a similar concentration–response in children (similar disease progression, similar response to intervention) compared to data in adults, only PK and safety studies are needed. If one does not anticipate such

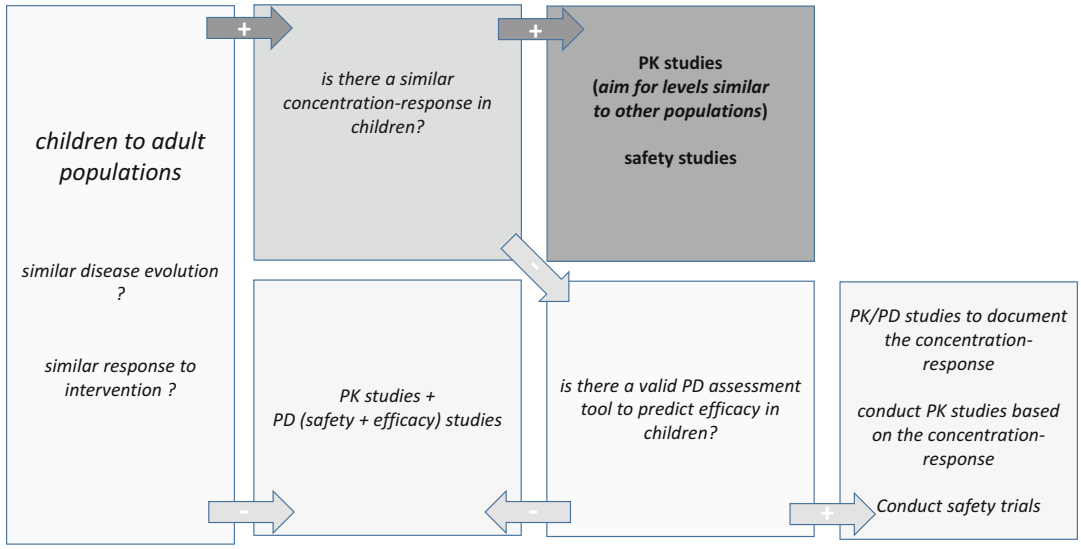


Fig. 2 Pediatric study decision tree (Manolis and Pons 2009)

a similar concentration–response relationship, PK, efficacy and safety trials are needed. In the latter scenario, the availability of a PD measurement ("biomarker") to predict efficacy will determine the final study design (e.g., conduct PK/PD studies to get a concentration–response relationship for the PD measurement) (Hampson et al. 2014). Irrespective of similarity or not in concentration–response in children versus adults, this means that PK data studies are needed in the different pediatric subpopulations to provide sufficiently robust data on the maturational aspects of this concentration–time profile (Fig. 1).

Advanced techniques (dried spot blood, other bodily fluids besides plasma or blood, microanalytical techniques) and innovative trial designs (sparse sampling, scavenged sampling, opportunistic sampling) and can be very supportive to lower the burden of such studies. Modeling and simulation is one way to circumvent several difficulties in developing medicinal products in children. It allows the quantitative use of sparse sampling, characterization and prediction of PK and PD, extrapolation from adults to children, interpolation between pediatric age subsets, optimal use of scientific literature, and in vitro/preclinical data (Manolis and Pons 2009). Both for the PK analysis and also during the study design, population modeling using either nonlinear mixed effect modeling (NONMEM) or physiology-based PK (PBPK) approaches allow analysis and interpretation based on such sparse and unbalanced datasets (European Medicines Agency 2016b). It also permits exploration of the impact of different covariates such as age, weight, disease characteristics, or comedication to explain the variability in medicine disposition or effects. Similarly, *a priori* information or information collected during a previous part of a study can be used to guide the optimal design of the study in order to obtain the maximal knowledge on the PK–PD characteristics in a given subpopulation (de Cock et al. 2011). Although population PK modeling (mechanism, physiology based) is not always yet sufficiently validated to support study design, and sampling strategy and uncertainty about extrapolation remains, they do provide a valuable tool to support study design and subsequent interpretation as a "*well educated best guess*." Guidelines on the qualification and reporting of physiologically based PK (PBPK)-modeling and simulation, including on pediatric analyses, have recently been provided by EMA (European Medicines Agency 2016b).

In contrast to the emerging knowledge about developmental PK, there is a relevant and obvious need to address the "PD gap" in pediatric medicine development. A clinical endpoint is a characteristic or variable that reflects how a patient feels, functions, or perhaps, simply survives. A biomarker is a surrogate endpoint, intended to substitute for such a clinical endpoint that predict clinical benefit, or harm, or lack of both. The US National Institute of Health (NIH) definition of a biomarker is "any characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, but also to quantify the response to a therapeutic (pharmacologic) intervention" (Regulatory Affairs Professionals Society 2016). Biomarkers can be used as surrogate endpoints, linking pathophysiological processes to clinical more relevant endpoints (e.g., blood pressure control versus long-term outcome, hemoglobin A1C versus long-term outcome of diabetes mellitus treatment). In some instances, biomarkers may serve as substitutes for clinical endpoints and may facilitate (conditional) medicine approval (Kearns 2010). Biomarkers can quantify *disease response or progression* (e.g., serum creatinine or cystatin values for renal function, exhaled nitric oxide for bronchial inflammation, 6 min walking test for cardiorespiratory or muscular strength, pulmonary function, hemoglobin A1C), to *predict a given medicine exposure or effect* (e.g., polymorphisms of medicine metabolizing enzymes or receptors), or this can be a *PD biomarker* (e.g., blood pressure, sedation or level of analgesia, neuroimaging techniques, neurocognitive outcome assessment, histamine response, pupillary diameter) (Kearns 2010).

While the concept of biomarkers can be applied to children as well as to adults, it is obvious that issues like disease severity, the natural course of a disease, and the effect of a medicine likely will be different in children, and this is even the case within the pediatric population (e.g., blood pressure, or growth velocity in preterm neonates or adolescents) (Kearns 2010). There is a huge opportunity to tailor or even to develop biomarkers for pediatric medicine development. Preferably, such biomarkers should be noninvasive, with – if any – limited discomfort, and feasible within the routine setting of clinical patient care (Kearns 2010).

Examples of PD endpoints that are somewhat specific to pediatrics are, for example, 6 min walking test in children with neuromuscular diseases, the prediction of long-term neurocognitive outcome based on the Bayley assessment tools at 18 and 24 months of life in former preterm neonates or using imaging techniques following asphyxia, body length and growth velocity assessment, or maturational and activity related differences in QT_c times to subsequently assess the impact of medicines on QT_c time intervals. Although these examples always remain disease specific, they illustrate the feasibility and relevance to validate such "surrogate" endpoints, hereby linking pathophysiological processes to clinical relevant endpoints in children.

Importantly, PD endpoints and biomarkers also cover adverse drug events and pharmacovigilance. An adverse drug reaction (ADR) has been defined by the World Health Organization (WHO) as "*any noxious or unintended medicine response at doses commonly used for prophylaxis, diagnosis, or treatment of a disease or condition*" (World Health Organization 2012). For children, it is perhaps more reasonable to consider "*an unintended and harmful effect resulting from the use of medications intended for diagnostic or therapeutic reasons (irrespective of the dose)*" as a broader definition, because of the common practice to apply off-label or unlicensed prescription of medicines, with the absence of dose guidance (Cliff-Eribo et al. 2016).

Similar to effects, medicine-related toxicity in children may also be different to that in adults. This may be due to impaired metabolism (e.g., ifosfamide-induced renal tubular cell toxicity, tolerance to paracetamol at birth, valproate toxicity in young infants) but may also be idiosyncratic or PD (e.g., more pronounced ibuprofen-related renal impairment in preterm neonates, age-related differences of cyclosporin on peripheral blood monocyte proliferation capacity, agitation after benzodiazepine administration in toddlers or young children) in its origin. It is important to recognize that the child is developing

and may be prone to different toxicities compared to adults (e.g., growth, neurodevelopment) (Cliff-Eribo et al. 2016).

Clinical trials are crucial to provide evidence-based knowledge, also on aspects of safety, since safety within clinical trials is one of the major outcome variables (Fig. 2). However, studies are commonly not sufficiently powered for safety but for efficacy. It is crucial that new but also established treatments are monitored for their effectiveness and safety under real-life conditions, including information in children (Cliff-Eribo et al. 2016). Consequently, post marketing surveillance and pharmacovigilance is essential but largely depends upon health care professionals. Sometimes, conditional approval with compulsory additional collection of data on effects or side effects is part of the product cycle with specific instructions on data collection in patients exposed to the medicine.

A Stakeholders Approach on Clinical Pharmacological Studies in Infants and Children

Pediatric Drug Therapy Is Not Yet Fully Perceived as an "Opportunity"

The success of any medicine development plan – including those in pediatrics – depends on the relevance and medical impact of the indication considered. Unfortunately, drug development programs are most commonly initially develop for other populations and "adult" indications and are not initially driven by pediatric needs. The more serious the condition and/or the lack of a satisfactory treatment will result in the perception of all stakeholders involved that a clinical trial is urgently needed and warranted. This also holds true for the investigational research boards or ethical committees (approval) and for the clinical investigators. Investigators are more likely to be recruited to study a medicine which is likely to result in significant clinical benefit to children instead of a setting one where there is already satisfactory treatment. The recent progress and ongoing studies in the field of spinal muscular atrophy (splice-switching therapy), cystic fibrosis (e.g., lumacaftor, ivacaftor), or muscular dystrophy (e.g., antisense oligonucleotides) serve as illustrations of these perceived clinical needs, but also the potential to make progress.

Parents and Their Children, Beyond Consent and Assent

The most effective way to improve consent and assent is to pay attention to the motivating and discouraging factors for research participation of parents and children. Such knowledge should enable professionals to improve and adapt the process of recruitment and informed consent to their perspectives. Using a systematic review approach, Tromp et al. documented that the most mentioned motivating factors for parents were health benefit for child, altruism, trust in research, and relation to the clinical researcher. For children, the most mentioned motivating factors for children were personal health benefit, altruism, and increasing comfort (Tromp et al. 2016). Fear of risks, distrust in research, logistical aspects, and disruption of daily life were mentioned most by parents as discouraging factors. Burden and disruption of daily life, feeling like a "guinea pig," and fear of risks were most mentioned as discouraging by children (Tromp et al. 2016). Parental drivers for consent include the perceived trust and the potential access to newer treatments. Altruism might be a soft value, but matters: the major motivation in parents to consent for their previously well child to participate in a RCT of therapy for an acute respiratory illness was to increase medical knowledge (benefit to all children 32%; improving knowledge 27%) (Sammons et al. 2007).

The legal age for consent varies in Europe (consent from the age of 14, 16, or 18 years onwards and the age at which assent is compulsory), the concept of giving assent before reaching the age of majority exist (Lepola et al. 2016). The term "informed assent" hereby describes the process whereby minors are structurally involved in the decision process about (dis)agreeing with voluntary participation in clinical trials (Lepola et al.

2016). We should not underestimate children's capacities and should try to involve them as soon as feasible in this decision process, and perhaps even consult them for advice on how to perform clinical studies. This is because the child is not just an innocent bystander in a pediatric trial as also reflected by International Children's Advisory Network (*i*CAN) (International Children's Advisory Network 2016). This is an expanding network of local groups of children that advocates for and with children, specifically those involved in clinical trials with the aim to improve pediatric health care by providing children and families a voice in health, research, medicine, and innovation. Healthy children understand the issues related to a randomized controlled trial and clinical studies; they understand the potential risks and mainly focus on the absence of interferences in their everyday life. Children with preexisting comorbidities had better understanding of study conduct and were more likely to support children's involvement in clinical trials (Cherrill et al. 2010).

Recruitment Challenges

A sufficient number of pediatric subjects should be recruited for a given trial to draw conclusions. The variability within the pediatric age range makes it only more difficult to draw these conclusions and will further raise the number of patients needed in the different subpopulation. The limited number of children with a specific disease that will qualify indeed for recruitment, the need for consent and assent, stringent inclusion and exclusion criteria, and the limited research capacity to conduct these studies will further jeopardizes effective recruitment.

All stakeholders involved should be aware that effective recruitment and retention is not only to the benefit of the sponsor but likely will also generate new knowledge on more effective pharmacotherapy for children as soon as possible, or may document the need for new approaches. When a significant number of clinical trials are recruiting at any one time, an individual patient may meet enrollment criteria for more than one clinical trial, and co-enrollment can be considered. For specific populations with relatively small numbers, like preterm neonates, critically ill children, or children with rare diseases, co-enrollment can be considered, although this raises pertinent scientific, ethical, regulatory, and industry issues that should be taken into account when considering reenrolling into multiple clinical studies (Randolph 2009).

Recruitment and subsequent retention are perceived to be different for pediatric compared to adult trials. Besides study-related issues, the reasons for these differences are multifactorial, but also relate to the child, the parents, and health care professionals. Finally, research capacity building to conduct clinical studies for regulatory purposes is an obvious need.

Child-related issues: Knowledge on more effective recruitment strategies and motivation in children is increasing and should be considered by sponsors since initial consent and retention in the study may largely depend on the applied study design. We already referred to the *i*CAN approach early in this chapter (International Children's Advisory Network 2016). An analysis of reasons why children decide not to participate in clinical research in a cohort of 161 children in The Netherlands documented that a lower age, less experience with disease, and less complex research with lower risk were the best predictors for *not* participating. Time constraints and additional burdens were hereby used as reasons not to participate (Hein et al. 2015). These patients were initially considered for 13 different studies (ten randomized controlled trials and three observational studies in the fields of pediatric oncology, gastroenterology, ophthalmology, pulmonology). A similar pattern was observed in a study on adolescents with type 2 diabetes. In this specific setting, monetary incentives and the use of technology turned out to be most effective. Even then, recruitment remained difficult, and the authors suggested that a study design covering concerns of the adolescents (e.g., body image, weight loss, stress management), and that accommodates to their schedules, and is conducted in more convenient locations than medical facilities is much more likely to recruit (Nguyen et al. 2014).

Parent-related issues: In essence, parents consider potential clinical benefit, child safety, practicalities of participation, research for the common good, access to new medicines, and randomization (Woolfall et al. 2013). An analysis on factors that drive decisions of parents to participate or not in clinical research suggested that parents who declined had a higher socioeconomic status, were more anxious about their decision, and found it harder to make their decision compared with consenting parents. Consenting parents expressed more trust and altruism, perceived the potential for enhanced care, reflected better the concept of randomization, and were more certain about the decisions made (Hoberman et al. 2013).

Research capacity: Patient and parents factors cannot be adapted acutely at the time of the individual decision to consent or not, but we should inform public and stakeholders about the need for clinical trials of medicines in children. Consequently, the bottom up approaches to support individual motivation of patients and their parents should be combined with more top down, structural approaches like the European Network of Pediatric Research at the European Medicines Agency (Enpr-EMA) (European Network of Pediatric Research at the European Medicines Agency 2012). Enpr-EMA aims to facilitate studies by establishing a European network of national and specialty networks, investigators, and centers with expertise in performing trials in the pediatric population. Through such a network, one aims to stimulate networking and stakeholder collaboration, to build research capacity, to inform the public and the relevant stakeholders about pediatric clinical trials, and to raise awareness among health care professionals about the need for clinical trials. At a more individual level, parents and health care professionals acknowledge the influence that pediatricians have on the decisions to participate. Pediatricians believe parents balance well the perceived risks and gains when deciding about participation. They thought the child's condition, parents' health beliefs and personal attributes, and the doctors' beliefs and relationship with the investigators influenced parents' attitudes. Perceived risks included inconvenience, inadequate resources, and potential harms to the patient and doctor-patient relationship. Perceived gains included professional benefits for pediatricians, improved patient care, convenience for the families and themselves, and scientific advancement (Caldwell et al. 2002).

Especially at the level of capacity building, there are very recent initiatives in the United States and in Europe that are very promising. In the United States, Institute for Advanced Clinical Trials (iACT) is an independent nonprofit organization and believe that children of all ages deserve innovative medical therapies that are developed with the same level of urgency and commitment afforded adults (International Children's Advisory Network 2016). The institute works with others to assure that studies are designed to generate sufficient data to allow safe and effective use of new medications and devices in pediatric populations. A key factor in making this a reality is to optimize and accelerate biomedical innovation using child-centered clinical trial networks and collaboration with like-minded institutions, trial sponsors, and other stakeholders. Together with parents, patients, investigators, foundations, regulators, other government agencies, biopharmaceutical sponsors, and children's networks, Institute for Advanced Clinical Trials (iACT) for children has the ambition to catalyze improvements in the quality and timely completion of global pediatric studies to address the gap in evidence for best use of therapeutics in children (International Children's Advisory Network 2016). Strategies to achieve this relate (i) to develop a strategy and planning (independent expert advice and guidance to sponsors), (ii) to develop capabilities, tools, and best practices (streamline and improve clinical trial processes), (iii) to develop infrastructure and clinical trial execution capacity (support and manage network of prequalified trial-ready sites and collaborate with regional networks), and (iv) to foster leadership (efforts to assure early and sustained engagement of patients, caregivers, and investigators). A very similar effort is ongoing in Europe within the Innovative Medicines Initiative (IMI2), as part of the 10th call (European Commission 2016). This

call explicitly stated that the paucity of patients available to study the many pediatric indications and the need for multiple capable sites to satisfy enrolment in trials, the clinical trial infrastructure across the EU is not sufficiently organized and lacks adequate funds and scale to consistently and efficiently deliver. This deficiency in clinical research capability also negatively impacts the capacity to complete research in areas of unmet medical need. In addition to improved infrastructure for efficient study execution, collaborative efforts to maximize the coordination of pediatric networks across the EU, utilize innovative study designs, and engage regulators in planning drug development programs are all needed to guarantee that Europe can augment its current capability as a critical region for developing medicines for children.

Pediatric oncology as an illustration of the progress made and the challenges to be handled: Pediatric oncology differs from adult findings both in the patterns of their malignancies as well as in the response to pharmacotherapeutic interventions. Despite these a priori existing hurdles, the success in childhood leukemia illustrates how treatment programs and multidisciplinary, multicenter collaborative efforts were designed using clinical- and biology-based risk factors seen in the patients and resulted in impressive improvement in outcome (Ravindranath 2015). These improvements were first focused on survival, with a subsequent shift to aspects related to long-term outcome and reduced morbidity and are very well described for acute leukemia. However and similar to the setting in adults, advances in cellular and molecular techniques (mechanism of action) have boosted the field of pediatric oncology, both for leukemia, but also for pediatric oncological diseases that are still associated with poor outcome (intracranial, sarcoma). Initiatives like the SIOPE strategic plan (*seven objectives*: innovative medicine, precision medicine, knowledge on the biology, equal access, specific emphasis on adolescents, quality of life in survivors, and causes) should enable further improvement in outcome, either related to mortality (more and sooner access to novel therapies in relapse) or related to morbidity (Vassal et al. 2015).

Pediatric Medicines Research: Keep the Momentum and Aim for Improvement

Clinical trials of medicines in children have resulted in improved product labeling, have increased the identification and quantification of adverse events, and have resulted in the development of new pediatric formulations. Between 1998 and 2012, the FDA issued 401 pediatric study requests. For 189 medicines, studies were completed and exclusivity has been granted. For the majority of medicines (173; 92%), additional and relevant information specific to pediatric pharmacotherapy has been added to the summary of product characteristics, with 108 (57%) receiving a new or expanded pediatric indication (Wharton et al. 2014). The aspects (parents and their children, recruitment challenges, research capacity) earlier discussed are crucial to further boost study conduct to improve knowledge based pediatric pharmacotherapy, and additional progress can also be made by using already existing – but still unpublished – information (Saint-Raymond et al. 2016). Progress may be slower than anticipated or aimed for, but pediatric medicines research has made progress, and we should try to keep this momentum. Initiatives like iCAN, iACT, or the IMI2 are the most recent illustrations of the growing international recognition of the relevance of pediatric trials. A more structured collaboration between the different stakeholders involved (society, regulatory authorities, parents and children, sponsors) at the international level is crucial to use the available, but limited, resources as effective as possible to further improve pharmaceutical care in children.

Despite this positive evolution, still a relevant portion (42%) of completed pediatric trials failed to document either efficacy (n = 38/44) or safety (n = 7/44) (Momper et al. 2015). Interestingly, the dosing evaluated in the study turned out to be a contributing factor for trial failure (all failed efficacy) in ten of these cases. This re-stresses the relevance to explore and validate dosing (phase 2 type of studies) before initiation of larger and pivotal phase 3 trials. Testing for more than one single dose in a specific pediatric population in a

phase 2 type of study design may provide very valuable information for the pediatric medicine development plan. In eight of these development programs, it is very likely that differences in disease characteristics in pediatrics compared to adults have further contributed to the failure to document efficacy. There seems to be also a higher placebo response, resulting in failure to proof efficacy in medicines considered for bipolar disorders in children (Janiaud et al. 2017). Finally, in four cases, the study design (assay sensitivity, control group) was deficient.

An additional aspect to improve this outcome is to consider study design related issues is the use of phase 2 before phase 3 studies. Therapeutic exploratory studies or phase 2 studies may be very useful to raise the likelihood of a valid study design – including dose seeking – before conducting the pivotal phase 3 efficacy trials and are common practice in "adult" product development and should be further explored in children. The starting doses considered are hereby generally based on estimates driven modeling and simulation. It allows the quantitative use of sparse sampling, characterization and prediction of PK and PD, extrapolation from adults to children, interpolation between pediatric age subsets, optimal use of scientific literature, and in vitro/preclinical data (Manolis and Pons 2009). Population modeling, either nonlinear mixed effect modeling (NONMEM) or physiology-based PK (PBPK) approaches, allow and support the analysis and interpretation based on such sparse and unbalanced datasets (European Medicines Agency 2016b). Phase 2 can subsequently be used to validate and optimize these models to further raise the likelihood that the results of subsequent phase 3 studies are indeed valid and useful. To further stress its relevance, an EMA-based survey illustrated that single dosing are the rule (83/97 studies), with commonly (40/97) no prespecified target (40/97) and if so, decisions are guided by PK data only (33/57) (Hampson et al. 2014). This reflects the common assumption of dose proportionality and similar exposure−response relationships in adults and children (Fig. 2). Few development programs prespecify steps to verify these assumptions in children. There is scope for the use of Bayesian methods as a framework for synthesizing existing information to quantify prior uncertainty about assumptions (Hampson et al. 2014).

Importantly, the number of medicines that target major pediatric diseases is still limited. The majority of studies in children are driven by a research program initially developed for adults, with subsequent translation to pediatrics, although some specific products (e.g., lumacaftor or ivacaftor for cystic fibrosis, surfactant for hyaline membrane disease) illustrate the potential of a pediatric pathophysiology focused approach. An emerging concept to facilitate medicine development is repurposing by the use of "old" medicines for new indications, avoiding the need for time- and resource-intensive toxicity studies. The potential relevance to neonates can be illustrated by repurposing projects related to propranolol (hemangioma) or insulin-like growth factor-1 (retinopathy of prematurity) (Léauté-Labrèze et al. 2017; Hellstrom et al. 2016).

Acknowledgments The clinical research activities of Karel Allegaert are facilitated by the agency for innovation by Science and Technology in Flanders (IWT) through the SAFEPEDRUG project (IWT/SBO 130033).

References and Further Reading

Allegaert K, van den Anker J (2015) Neonatal drug therapy: the first frontier of therapeutics for children. Clin Pharmacol Ther 98:288–297

Caldwell PH, Butow PN, Craig JC (2002) Pediatricians' attitudes toward randomized controlled trials involving children. J Pediatr 141:798–803

Cherrill J, Hudson H, Cocking C et al (2010) Clinical trials: the viewpoint of children with a chronic illness compared with healthy children. Arch Dis Child 95:229–232

Cliff-Eribo KO, Sammons H, Choonara I (2016) Systematic review of paediatric studies of adverse drug reactions from pharmacovigilance databases. Exp Opinin Drug Saf 15:1321–1328

Critical Path Institute (2015) https://c-path.org/programs/inc. Accessed 21 Jul 2017

De Cock RF, Piana C, Krekels EH et al (2011) The role of population PK–PD modelling in paediatric clinical research. Eur J Clin Pharmacol 67(Suppl 1):5–16

Doussau A, Geoerger B, Jimenez I et al (2016) Innovations for phase I dose-finding designs in pediatric oncology clinical trials. Contemp Clin Trials 47:217–227

European Commission (2016) http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/imi2-2016-10-04.html. Accessed 21 Jul 2017

European Medicines Agency (2016a) http://ec.europa.eu/health//sites/health/files/files/paediatrics/2016_pc_report_2017/ema_10_year_report_for_consultation.pdf. Accessed 21 Jul 2017

European Medicines Agency (2016b) http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/07/WC500211315.pdf. Accessed 21 Jul 2017

European Network of Paediatric Research at the European Medicines Agency (2012) http://www.ema.europa.eu/ema/index.jsp?curl=pages/partners_and_networks/general/general_content_000303.jsp. Accessed 21 Jul 2017

Gonzalez D, Paul IM, Benjamin DK Jr et al (2014) Advances in pediatric pharmacology, therapeutics, and toxicology. Adv Pediatr Infect Dis 61:7–31

Hampson LV, Herold R, Posch M et al (2014) Bridging the gap: a review of dose investigations in paediatric investigation plans. Br J Clin Pharmacol 78:898–907

Hein IM, Troost PW, de Vries MC et al (2015) Why do children decide not to participate in clinical research: a quantitative and qualitative study. Pediatr Res 78:103–108

Hellstrom A, Ley D, Hansen-Pupp I et al (2016) Insulin-like growth factor 1 has multisystem effects on foetal and preterm infant development. Acta Paediatr 105:576–586

Hoberman A, Shaikh N, Bhatnagar S et al (2013) Factors that influence parental decisions to participate in clinical research: consenters vs nonconsenters. JAMA Pediatr 167:561–566

Hoppu K, Anabwani G, Garcia-Bournissen F et al (2012) The status of paediatric medicines initiatives around the world – what has happened and what has not? Eur J Clin Pharmacol 68:1–10

Horen B, Montastruc JL, Lapeyre-Mestre M (2002) Adverse drug reactions and off-label drug use in paediatric outpatients. Br J Clin Pharmacol 54:665–670

International Children's Advisory Network (2016) https://www.icanresearch.org. Accessed 21 Jul 2017

Janiaud P, Cornu C, Lajoinie A et al (2017) Is the perceived placebo effect comparable between adults and children ? A meta-regression analysis. Pediatr Res 81:11–17

Joseph PD, Craig JC, Caldwell PH (2015) Clinical trials in children. Br J Clin Pharmacol 79:357–369

Kearns GL (2010) Beyond biomarkers: an opportunity to address the 'pharmacodynamic gap' in pediatric drug development. Biomark Med 4:783–786

Léauté-Labrèze C, Harper JI, Hoeger PH (2017) Infantile haemangioma. Lancet 390:85–94

Lepola P, Needham A, Mendum J et al (2016) Informed consent for paediatric clinical trials in Europe. Arch Dis Child 101:1017–1025

Magalhaes J, Rodrigues AT, Rogue F et al (2015) Use of off-label and unlicensed drugs in hospitalized paediatric patients: a systematic review. Eur J Clin Pharmacol 71:1–13

Manolis E, Pons G (2009) Proposals for model-based paediatric medicinal development within the current European Union regulatory framework. Br J Clin Pharmacol 68:493–501

Momper JD, Mulugeta Y, Burckart GJ (2015) Failed pediatric drug development trials. Clin Pharmacol Ther 98:245–251

Mühlbauer B, Janhsen K, Pichler J et al (2009) Off-label of prescription drugs in childhood and adolescence: an analysis of prescription patterns in Germany. Dtsch Arztebl Int 106:25–31

Nguyen TT, Jayadeva V, Cizza G et al (2014) Challenging recruitment of youth with type 2 diabetes into clinical trials. J Adolesc Health 54:247–254

Randolph AG (2009) The unique challenges of enrolling patients into multiple clinical trials. Crit Care Med 37(1 Suppl):S107–S111

Ravindranath Y (2015) Evolution of modern treatment of childhood acute leukemia and cancer: adventures and battles in the 1970s and 1980s. Pediatr Clin N Am 62:1–10

Reed MD (2011) Reversing the myths obstructing the determination of optimal age- and disease-based drug dosing in pediatrics. J Pediatr Pharmacol Ther 16:4–13

Regulatory Affairs Professionals Society (2016) http://www.raps.org/Regulatory-Focus/News/2016/01/29/24212/FDA-NIH-Unveil-Glossary-of-Clinical-Definitions. Accessed 21 Jul 2017

Saint-Raymond A, Pelle B, Zaccaria C et al (2016) Usage of unpublished paediatric data. Arch Dis Child 101:81–84

Sammons HM, Atkinson M, Choonara I et al (2007) What motivates British parents to consent for research? A questionnaire study. BMC Pediatr 7:12

Schachter AD, Ramoni MF (2007) Paediatric drug development. Nat Rev Drug Discov 6:429–430

Stiers JL, Ward R (2014) Newborns, one of the last therapeutic orphans to be adopted. JAMA Pediatr 168:106–108

Tromp K, Zwaan CM, van de Vathorst S (2016) Motivations of children and their parents to participate in drug research: a systematic review. Eur J Pediatr 175:599–612

Turner MA, Portman RJ, Davis JM (2017) Regulatory science in neonates: a framework that supports evidence-based drug therapy. JAMA Pediatr. https://doi.org/10.1001/jamapediatrics.2017.1360

United States National Institutes of Health (2017) https://www.clinicaltrials.gov/ct2/search. Accessed 15 Jul 2017

Van Riet-Nales DA, de Jager KE, Schobben AF et al (2011) The availability and age-appropriateness of medicines

authorized for children in the Netherlands. Br J Clin Pharmacol 72:465–473
Vassal G, Rousseau R, Blanc P et al (2015) Creating a unique, multi-stakeholder Paediatric Oncology Platform to improve drug development for children and adolescents with cancer. Eur J Cancer 51:218–224
Wharton GT, Murphy MD, Avant D et al (2014) Impact of pediatric exclusivity on drug labeling and demonstrations of efficacy. Pediatrics 134:e512–e518
Woolfall K, Shilling V, Hickey H et al (2013) Parents' agendas in paediatric clinical trial recruitment are different form researchers' and often remain unvoiced: a qualitative study. PLoS One 8:e67352
World Health Organization (2012) http://www.who.int/medicines/areas/quality_safety/safety_efficacy/EMP_ConsumerReporting_web_v2.pdf. Accessed 21 Jul 2017
World Health Organization (2017) http://www.who.int/medicines/publications/essentialmedicines/6th_EMLc2017.pdf. Accessed 21 Jul 2017

Clinical Studies in Geriatric Population 20

Petra A. Thürmann

Contents

Abstract

Exclusion of the older population from clinical trials has for long resulted in a knowledge gap about drug efficacy and side effects in older age groups. When considering the demographic change in most industrialized countries and the fact that seniors receive the highest share of all drugs, a need for clinical trials with elderly patients is obvious. Potential pharmacokinetic as well as pharmacodynamic changes in comparison to younger adults have to be described. The safety profile can be different due to several reasons, which may be explained by comorbidities, comedications, or age-related changes. Calendar age serves only as a proxy for age, and geriatric assessment tools should more and more be applied to characterize the heterogenous group of older adults from 65 years to more than 100 years of age. Several ICH guidelines and reflection papers from regulatory bodies do exist and offer advice on how to design and analyze clinical trials including the older population.

P. A. Thürmann (✉)
Philipp Klee-Institute of Clinical Pharmacology, HELIOS University Clinic Wuppertal, Wuppertal, Germany

University Witten/Herdecke, Witten, Germany
e-mail: petra.thuermann@uni-wh.de; petra.thuermann@helios-gesundheit.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_43

Introduction

According to international statistics, the percentage of older adults is steadily increasing in Europe as well as the USA; nearly 20% of the population is above the age of 65 years; particularly in Japan the percentage of this age group comes to 28%. Also in China the older population is increasing, but this percentage is slightly above 10% (China) and only about 6% (India) (GBD Mortality Collaborators 2018a). However, elderly people are prescribed the highest number of drugs (Qato et al. 2016); however, over-the-counter (OTC) drugs have to be considered as well. The reason for this phenomenon is obvious: with increasing age the percentage of patients with one or more chronic conditions increases, and most symptoms and conditions are treated by medicines (GBD Mortality Collaborators 2018b).

Already in 1994 the ICH guideline for "Studies in support of special populations: geriatrics E7" emphasized that the use of drugs in this population requires special consideration due to the frequent occurrence of underlying diseases, concomitant drug therapy, and the subsequent risk of drug interactions. Consequently, "patients entering the clinical trials should be reasonably representative of the population that will be later treated with the drugs" (ICH 1993).

Since then numerous publications report underrepresentation of older patients in clinical trials (Downing et al. 2016). Lewis et al. (2003) reported that only 25% of participants in cancer trials were older than 65 years, whereas 61% of the population affected by cancer are in this age group. Similarly, elderly patients were frequently excluded from cardiovascular trials, where mean age of study patients came to about 60–65 years in most of the trials investigating drugs for treatment of heart failure (Vitale et al. 2017).

Several measures were undertaken by regulatory agencies to increase the inclusion of elderly patients in clinical trials and also to encourage the consideration of age-specific issues such as multiple comorbidities, concomitant medications, changes in pharmacokinetics, and age-specific safety analyses. Given the fact that the older population is the fastest growing age group, enrollment of 100 patients above the age of 65 years is for many indications no longer representative. Moreover, date should also be presented for three senior age groups, 65–74, 75–84, and above 85, according to the 2010 update of ICH E7 (Cerreta et al. 2015; ICH 2010). Sponsors are encouraged to mirror the real-life conditions of multimorbidity while including older patients. In some instances it is more informative to include older patients in the regular phase III trials, because this allows for comparison of the benefit/risk ratio between younger and older patients. However, a particular study with extensive cognitive testing is more likely to be performed in a separate "geriatric" trial (ICH 2010).

The Need for Clinical Trials in Older Adults

Age itself is associated with several biological changes resulting in altered pharmacokinetics and pharmacodynamics of medicines (Reeve et al. 2015; Shah 2004). This may result in different dosing recommendations (usually lower doses) for older patients and an altered risk/benefit ratio for some drugs. Secondly, the majority of older patients has more than one chronic disease (van den Akker et al. 2009) and receives drugs, which may interfere negatively with the investigational medicinal product or the condition, for which the investigational drug should be given (Shah 2004). In summary, old age, multimorbidity, and polypharmacy together increase the likelihood for drug-drug and drug-disease interactions as well as for adverse drug reactions (ADRs) (Shah 2004). However, information about contraindicated drugs and diseases is necessary for the safe use of a new drug.

One issue, which has been neglected so far due to its complexity, is the relevance of a trial endpoint for the older population. Prolongation of life may in certain stages of age and frailty (see below) no longer be a meaningful endpoint, and the ability to perform regular activities in daily life may be more relevant (Bartlam et al. 2012; Mangoni and Pilotto 2016).

The abovementioned reasons explain some of the major challenges faced with clinical trials in the older population. But apart from these considerations, more practical issues such as informed

consent forms for very old persons, their ability to visit the research clinic, and necessary support by caregivers have to be considered by sponsors and trial sites (Bartlam et al. 2012; Mangoni and Pilotto 2016).

As already outlined in the figures about the demographic change and the reference made to the ICH E7 guideline (ICH 1993), the definition of "geriatric population" or "older population" has to be reconsidered.

Aging and Definitions for the Geriatric Population

According to Strehler (1962), aging has been described as a progressive functional decline, as a result of changes made to cells and structural components (e.g., bone) (Table 1).

These changes can be found in all biological systems, e.g., weight loss; a lower metabolic rate; longer reaction times; declines in cognitive and memory functions; decline in audition, olfaction, and vision and all organ functions, as well as changes in the immune system. Interestingly, in old age we can find an increase in inflammation levels, which may contribute to age-related conditions such as arthritis and atherosclerosis (da Costa et al. 2016; Formiga et al. 2016). The process of aging has not yet been fully elucidated; it includes molecular biology (e.g., telomere length, loss of DNA repair mechanisms) as well as environmental factors (da Costa et al. 2016).

These physiological changes have to be considered when performing trials in older patients. One aspect of old age should be highlighted because of its importance for side effects of drugs: impaired homeostasis. This can be illustrated with a few examples as shown in Table 2.

One more feature in the older population are the so-called geriatric syndromes, which are defined as follows:

> Clinical conditions in older persons that do not fit into disease categories but are highly prevalent in old age, multifactorial, associated with multiple co-morbidities and poor outcomes and are only treatable when a multidimensional approach is used. (Cruz-Jentoft et al. 2010; Inouye et al. 2007)

Table 1 Attributes of aging. (Adapted from Strehler 1962)

Feature	Consequence
Cumulative	Effects of aging increase with time
Universal	All members of a species display signs of aging
Progressive	Aging is a series of gradual changes
Intrinsic	Changes would take place even in a "perfect" environment
Deleterious	Changes compromise normal biological functions

Table 2 Examples for relevant side effects of drugs due to impaired homeostasis in old age

Drug effect	Consequence
Antihypertensive	Orthostatic hypotension, risk for falling
Diuretic	Exsiccosis during summer heat or with concomitant diarrhea
Anticholinergic	Delirium, misdiagnosed as dementia
Sedative	Stupor, risk for falling
Antihistaminergic	Oversedation

Examples for commonly accepted geriatric syndromes are delirium, falls, frailty, dizziness, syncope, and urinary incontinence (Inouye et al. 2007). It is obvious that cumulation of these syndromes impairs activities of daily living, may hamper participation in a clinical trial (mobility, urinary incontinence), and increases the risk for ADRs. It is thus necessary to document the presence of these syndromes as "diagnoses" and to document changes in these syndromes throughout any clinical trial in geriatric patients.

Further age-related conditions may compromise the outcome of a study: the tolerance for time-consuming questionnaires is reduced, which may impede the evaluation of a clinical trial's endpoint. In clinical trials with antiinfectives, body temperature is frequently documented and may serve as a predefined endpoint or for decision-making about antibiotic treatment. However, the increase in body temperature is attenuated in old age, which may result in wrong treatment decisions. It is thus necessary to involve geriatric expertise when planning such trials.

Another unique feature for old age is the huge interindividual variability among seniors: in all old age groups, we find the so-called go-go's (extremely fit older adults), slow-go's, and no-go's (usually immobile). For the pediatric population, it is obvious that newborns are a different group than adolescents and it is fairly easy to differentiate according to calendar age. In contrast, the common age definition "above 65 years" (usually in parallel with retirement) does not consider the interinidivdual variability and the necessity to differentiate between "old" and "oldest old." According to the WHO definition from 1980, "elderly adults" were between 61 and 75 years, "old" between 76 and 90 years, and above 90 years the "oldest old." In the ICH E7 Questions and Answers Document (ICH 2010), it is specified that "data should be presented for various age groups (e.g., <65, 65–74, 75–85 and ≥85)" to cover the whole age range. Having in mind the abovementioned interindividual variability with regard to fitness, a more detailed specification for this population is required.

Functional Decline, Multimorbidity, and Frailty

Obviously, above a certain age, calendar age is not the ideal "biomarker" for the vulnerability or fitness of a trial participant. In order to characterize the heterogenous older population, reliable and valid biomarkers could be a tool to differentiate between the go-go, slow-go, and no-go phenotypes. In terms of gerontological science, numerous molecular markers, e.g., telomere length, are associated with longevity but do not reflect physical or cognitive functions. One feature typical for aging is sarcopenia, a progressive loss of skeletal muscle mass and strength resulting in decline of mobility and loss of independence, the latter being one of the most important aspects in old age (Doherty 2003). Measuring muscle strength, e.g., with the handgrip test, is therefore and easy but important way to characterize a patient's physical fitness. Fried et al. developed the concept of frailty, which can be identified by addressing five simple questions (Fried et al. 2001):

- Unintentional weight loss (4,5 kg during the last 12 months)
- Exhaustion (self-reported)
  - "I felt that everything I did was an effort"
    "I could not get going"
- Muscle weakness (handgrip strength, dynamometer)
- Slowness while walking (observed, 4 m >6 s)
- Low levels of activity (self-reported, >4 h sitting/day;
  <1 walk/month)

By adding up predefined points for the results, patients can be classified as frail, pre-frail, or fit. This score has been validated across various geriatric populations (e.g., inpatients, nursing home residents, community-dwelling elderlies) and is significantly associated with falls, morbidity, and mortality.

In general, physical as well as cognitive functions can be measured within a so-called geriatric assessment, as outlined by the Geriatric Expert Committee of the EMA (Cerreta et al. 2016). For example, the Multidimensional Prognostic Index indicates the 1-year risk for dying in an older population after hospitalization and includes a comprehensive geriatric assessment with cognitive tests, activities of daily living (ADL), as well as nutrition, number of chronic conditions, and daily medications (Mangoni and Pilotto 2016). However, a comprehensive assessment requires at least 30 minutes and is not feasible in most clinical trials. A recent meta-analysis in published randomized controlled clinical trials performed in older adults showed that such geriatric testing is not even performed in many trials designed for patients above 70 or 80 years (Van Deudekom et al. 2017). Therefore, the EMA Geriatric Expert Committee suggests a reduced version (Table 3).

These tests are validated and freely available, in contrast to some other test batteries where licenses have to be paid for (EMA 2018). It should be considered that not all of these tests are feasible in every trial, and they do have their shortcomings too. As an example, the SPPB has a good

Table 3 Geriatric assessment as suggested by the Geriatric Expert Committee of the EMA (Cerreta et al. 2016; EMA 2018)

Test	Parameter(s)	References
Short Physical Performance Battery (SPPB)	Balance (10s); ability to rise from a chair gait speed (usually 4 m distance, 1 m/s)	Guralnik et al. (1994)
Montreal Cognitive Assessment (MoCA)	Short-term memory recall; visuospatial abilities (incl. clock-drawing task); executive functions; attention, concentration and working memory; language; orientation to time and place	Nasreddine et al. (2005)
Mini Nutritional Assessment-Short Form (MNA-SF)	Nutrition; anthropometric measurements; global assessment; self-assessment	Guigoz et al. (2002)
Cumulative Illness Rating Scale-Geriatric (CIRS-G)	Number and type of chronic illnesses; severity of chronic diseases; representing individual body systems	Miller et al. (1992)

performance for outcomes as hospitalization, falls, and mortality but cannot be used for patients who are currently at bed rest (e.g., acute illness) and does only partially consider cognition and even interferes with cognition.

Clinical Pharmacology Issues to Consider: Pharmacokinetics, Pharmacodynamics, and Drug Interactions

With advancing age not only the renal but also the hepatic elimination capacity is reduced. In addition, other factors such as higher body fat content and lower plasma protein concentration contribute to pharmacokinetic changes during the process of aging (Reeve et al. 2015; Shah 2004; Tan et al. 2015). Therefore, pharmacokinetic data are required for all age groups and can be sampled during clinical trials to allow for population modeling approaches (Cerreta et al. 2015; ICH 2010). It is of particular importance to describe alterations in pharmacokinetics which cannot be explained by a decrease in renal function alone as pharmacokinetics in patients with impaired renal function are usually covered in additional supporting pharmacokinetic studies. There are only few examples for pharmacokinetic studies in frail elderly patients (Shah 2004); however, Johnston et al. (2014) showed nicely that gentamicin clearance differs among older patients not only depending on renal function but also in correlation with frailty. Likewise, Schwartz and Verotta (2009) investigated the pharmacokinetic differences of atorvastatin in elderly community-dwelling patients and nursing home residents. These studies are challenging what concerns the realization in practice and, however, offer important insight into pharmacokinetic peculiarities of frail elderly patients. Such studies in the geriatric population are requested by the ICH E7 guideline (ICH 2010) particularly when the investigational product will be used particularly in this population, e.g., with Alzheimer's or Parkinson's disease.

It should be noted that for the improvement of dose selection in the pediatric population, physiologically based pharmacokinetic modeling appears to be a useful tool and is frequently used. Unfortunately, only very few publications show the application of these modeling approaches in the older population (Schlender et al. 2016), although a database on physiological parameters has been created (Thompson et al. 2009).

Pharmacokinetic drug-drug interaction studies are required depending on the metabolic properties of the new drug (Shah 2004). Since multiple comedications are to be expected when planning phase III trials, information about the potential for pharmacokinetic interactions is mandatory. Elderly patients are likely to receive drugs with a narrow therapeutic range such as digoxin and oral

anticoagulants (particularly vitamin K antagonists but also the direct oral anticoagulants); therefore, any influence on these drugs should be known. But also for the most relevant cytochrome P450 enzyme interaction, studies are needed, depending on the drug's metabolism. It should be noted that many older patients receive not only one drug metabolized, e.g., via CYP2D6, but sometimes two or even more and even minor interactions may become clinically relevant (ICH 1993, 2010; Shah 2004).

Older adults generally have a higher risk for ADRs due to different and not only pharmacokinetic reasons (Shah 2004). As already described in Table 2, certain side effects call for special attention. As outlined in the ICH E7 document from 2010 (ICH 2010), drugs with antihistaminergic, anticholinergic, and/or sedative effects need to be tested in the geriatric population, and the extent of these effects has to be shown and compared with younger patients. Cognitive effects can be studied with standardized test batteries, whereas effects, e.g., on bladder function or constipation, have to be explored verbally or with specific questionnaires. Drug effects on sarcopenia (e.g., tested as muscle strength) or falls in general must be considered. All these effects have to be recognized in the presence of multimorbidity and polypharmacy, i.e., all comedications have to be considered and evaluated, too. For marketed drugs different tools and lists have been developed to calculate their anticholinergic burden (Hilmer et al. 2007) or to describe their potential risk for side effects in older patients (The 2019 American Geriatrics Society Beers Criteria® Update Expert Panel 2019; Renom-Guiteras et al. 2015).

Drug Administration and Formulation

For the pediatric population, it was obvious that special drug formulations are needed and that taste of a syrup matters for adherence reasons. It may be less common knowledge that dysphagia is a frequently encountered problem with geriatric patients and is associated with negative health outcomes (Baijens et al. 2016). Swallowing a tablet or capsule of a certain size can become virtually impossible and sometimes dangerous. A recent study found that approximately 50% of geriatric inpatients suffer from dysphagia as proven by the volume-viscosity swallow test and the Minimal Eating Observation Form-II (Melgaard et al. 2018). Older patients also face problems when opening drug containers and blisters or trying to divide their medications (Notenboom et al. 2017). To overcome the most prevalent of these problems, the EMA CHMP reflection paper (van Riet-Nales et al. 2016) on the quality aspects (pharmaceutical development) of medicines for older people addresses three major issues: pharmaceutical formulations, package opening, and storage. Although these problems seem to be of less importance, not being able to use the medications appropriately presents a relevant safety and efficacy factor in clinical practice and may also impact on results of clinical trials in the geriatric population.

Ethical Considerations, Recruitment, and Support

Treatment of older patients always involves a specific risk/benefit estimation between expected treatment effects, treatment-related adverse events, and the expected general survival probability of a given patient (Mangoni and Pilotto 2016; Soto-Perez-de-Celis et al. 2018). For example, due to the increase in older cancer patients with multimorbidity and different states of (non)-frailty, it is of paramount importance to gain information about the risks and benefits of a new cancer treatment in the context of old age, high comorbidity, high risk for treatment-associated ADRs, and a limited overall life expectancy. This calls again for the need to characterize physical and cognitive performance of study participants (Mangoni and Pilotto 2016). Moreover, specific trials have to be conducted in the oldest old with certain cancer types, which are particularly prevalent in this age class, e.g., chronic myeloid leukemia (O'Reilly et al. 2018).

To gain insight into patients', caregivers', and health professionals' thoughts about clinical trials in older patients, the PREDICT research project

was carried out in nine European countries (Bartlam et al. 2012). Focus interview groups were conducted across countries and diseases and revealed several astonishing results: Ageism was noticed as a general view about older adults and seniors as well as caregivers expressed that they felt they were excluded from the potential benefits of research. The risks of participating in clinical trials were perceived but accepted. However, the measured endpoints such as mortality/prolongation of life were criticized as being no longer relevant for some of the older participants. They wanted to be involved in the planning of the trial to assure that their preferences and values are considered. The patient-physician relationship proved to be very important for the decision to participate or not. Patient information brochures were found to be too long and hard to understand. Many practical issues were discussed, such as frequent visits at the trial site and necessity for family (travel) support. In conjunction with the guidelines for clinical trials in the geriatric population, participation of older patients must be strengthened by several means: during the design of the study (e.g., inclusion and exclusion criteria, outcomes), the trial conduct and enrollment (e.g., travel support), and at the trial site with a special focus on the needs and preferences of older adults.

In conclusion, the population of elderly patients must be represented in clinical trials, which is mandated by current guidelines of regulatory bodies. Enrolling senior citizens in such studies is a major challenge, and geriatric experts should be involved to assure the design of pragmatic studies is adequate for the requirements and preferences of geriatric patients.

References and Further Reading

Baijens LW, Clavé P, Cras P et al (2016) European Society for Swallowing Disorders – European Union Geriatric Medicine Society white paper: oropharyngeal dysphagia as a geriatric syndrome. Clin Interv Aging 11:1403–1428

Bartlam B, Crome P, Lally F et al (2012) The views of older people and carers on participation in clinical trials. Clin Invest 2:327–336

By the 2019 American Geriatrics Society Beers Criteria® Update Expert Panel (2019) American Geriatrics Society 2019 Updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. J Am Geriatr Soc 67:674–694

Cerreta F, Temple R, Asahin C et al (2015) Regulatory activities to address the needs of older patients. J Nutr Health Aging 19:232–233

Cerreta F, Ankri J, Bowen D et al (2016) Baseline frailty evaluation in drug development. J Frailty Aging 5:139–140

Cruz-Jentoft AJ, Baeyens JP, Bauer JM et al (2010) Sarcopenia: European consensus on definition and diagnosis. Age Ageing 39:412–423

da Costa JP, Vitorino R, Silva GM et al (2016) A synopsis on aging – theories, mechanisms and future prospects. Ageing Res Rev 29:90–112

Doherty TJ (2003) invited review: aging and sarcopenia. J Appl Physiol 95:1717–1727

Downing NS, Shah ND, Neiman JH et al (2016) Participation of the elderly, women, and minorities in pivotal trials supporting 2011– 2013 U.S. Food and Drug Administration approvals. Trials 17:199. https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-016-1322-4. Accessed 26 May 2019

European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) (2018) Reflection paper on physical frailty: instruments for baseline characterisation of older populations in clinical trials. EMA/CHMP/778709/2015. https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-physical-frailty-instruments-baseline-characterisation-older-populations-clinical_en.pdf. Accessed 26 May 2019

Formiga F, Ferrer A, Padros G et al (2016) Evidence of functional declining and global comorbidity measured at baseline proved to be the strongest predictors for long-term death in elderly community residents aged 85 years: a 5-year follow-up evaluation, the OCTABAIX study. Clin Interv Aging 11:437–444

Fried LP, Tangen CM, Walston J et al (2001) Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 56:M146–M156

GBD Mortality Collaborators (2018a) Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392:1684–1735

GBD Mortality Collaborators (2018b) Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392:1789–1858

Guigoz Y, Lauque S, Vellas BJ (2002) Identifying the elderly at risk for malnutrition. The Mini Nutritional Assessment. Clin Geriatr Med 18:737–757

Guralnik JM, Simonsick EM, Ferrucci L et al (1994) A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 49:M85–M94

Hilmer SN, Mager DE, Simonsick EM et al (2007) A drug burden index to define the functional burden of medications in older people. Arch Intern Med 167:781–787

Inouye SK, Studenski S, Tinetti ME, Kuchel GA (2007) Geriatric syndromes: clinical, research, and policy implications of a core geriatric concept. J Am Geriatr Soc 55:780–791

International Conference on Harmonisation of Technical requirements for Registration of Pharmaceuticals for Human Use (1993) ICH Harmonised Tripartite Guideline. Studies in Support of Special Populations: Geriatrics E7.Current Step 4 Version. CPMP/ICH/379/95. https://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. Accessed 26 May 2019

International Conference on Harmonisation of Technical requirements for Registration of Pharmaceuticals for Human Use (2010) E 7 Questions and Answers. EMA/CHMP/ICH/604661/2009. https://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. Accessed 26 May 2019

Johnston C, Hilmer SN, McLachlan AJ et al (2014) The impact of frailty on pharmacokinetics in older people: using gentamicin population pharmacokinetic modeling to investigate changes in renal drug clearance by glomerular filtration. Eur J Clin Pharmacol 70:549–555

Lewis JH, Kilgore ML, Goldman DP et al (2003) Participation of patients 65 years of age or older in cancer clinical trials. J Clin Oncol 21:1383–1389

Mangoni AA, Pilotto A (2016) New drugs and patient-centred end-points in old age: setting the wheels in motion. Expert Rev Clin Pharmacol 9:81–89

Melgaard D, Rodrigo-Domingo M, Mørch MM (2018) The prevalence of oropharyngeal dysphagia in acute geriatric patients. Geriatrics 26(15). https://doi.org/10.3390/geriatrics3020015.

Miller MD, Paradis CF, Houck PR et al (1992) Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. Psychiatry Res 41:237–248

Nasreddine ZS, Phillips NA, Bédirian V et al (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53:695–699

Notenboom K, Leufkens HG, Vromans H, Bouvy ML (2017) Learning from patients: Identifying design features of medicines that cause medication use problems. Int J Pharm 517:128–134

O'Reilly A, Murphy J, Rawe S, Garvey M (2018) Chronic lymphocytic leukemia: a review of front-line treatment options, with a focus on elderly CLL patients. Clin Lymphoma Myeloma Leuk 18:249–256

Qato DM, Wilder J, Schumm PL et al (2016) Changes in prescription and over-the-counter medication and dietary supplement use among older adults in the United States, 2005 vs 2011. JAMA Intern Med 176:473–482

Reeve E, Wiese MD, Mangoni AA (2015) Alterations in drug disposition in older adults. Expert Opin Drug Metab Toxicol 11:491–508

Renom-Guiteras A, Meyer G, Thürmann PA (2015) The EU(7)-PIM list: a list of potentially inappropriate medications for older people consented by experts from seven European countries. Eur J Clin Pharmacol 71:861–875

Schlender JF, Meyer M, Thelen K et al (2016) Development of a whole-body physiologically based pharmacokinetic approach to assess the pharmacokinetics of drugs in elderly individuals. Clin Pharmacokinet 55:1573–1589

Schwartz JB, Verotta D (2009) Population analyses of atorvastatin clearance in patients living in the community and in nursing homes. Clin Pharmacol Ther 86:497–502

Shah RR (2004) Drug development and use in the elderly: search for the right dose and dosing regimen. Br J Clin Pharmacol 58:452–469

Soto-Perez-de-Celis E, Li D, Yuan Y (2018) Functional versus chronological age: geriatric assessments to guide decision making in older patients with cancer. Lancet Oncol 19:e305–e316

Strehler BL (1962) Time, cells, and aging. Academic, New York

Tan JL, Eastment JG, Poudel A, Hubbard RE (2015) Age-related changes in hepatic function: an update on implications for drug therapy. Drugs Aging 32:999–1008

Thompson CM, Johns DO, Sonawane B et al (2009) Database for physiologically based pharmacokinetic (PBPK) modeling: physiological data for healthy and health-impaired elderly. J Toxicol Environ Health B Crit Rev 12:1–24

van den Akker M, Buntinx F, Knottnerus JA (2009) Comorbidity or multimorbidity. Eur J Gen Pract 2:65–70

Van Deudekom FJ, Postmus I, Van der Ham DJ et al (2017) External validity of randomized controlled trials in older adults, a systematic review. PLoS ONE 12(3):e0174053

van Riet-Nales DA, Hussain N, Sundberg KA et al (2016) Regulatory incentives to ensure better medicines for older people: from ICH E7 to the EMA reflection paper on quality aspects. Int J Pharm 512:343–351

Vitale C, Fini M, Spoletini I et al (2017) Under-representation of elderly and women in clinical trials. Int J Cardiol 232:216–221

Personalized Medicine

21

Federico Goodsaid, Felix Frueh, and Michael E. Burczynski

Contents

Abstract

Personalized medicine (PM) has grown in both clinical importance and in cultural recognition and popularity over the last two decades. In this chapter we will first address the different public perceptions of the term "personalized medicine," with the ultimate goal of identifying the definition(s) of PM which best reflect how it is currently prosecuted within modern drug development. A brief historical context will then be provided, followed by a discussion of the general tenets of personalized medicine as employed within the pharmaceutical industry and in clinical practice. The chapter concludes with some tables and a small number of case studies that highlight the PM concept in both oncology and rare diseases and a consideration of what the future holds for the field.

F. Goodsaid · F. Frueh
Opus Three LLC Life Sciences, Pescadero, CA, USA
e-mail: federico@opusthree.com; felix@opusthree.com

M. E. Burczynski (✉)
Department of Translational Medicine, Global R&D, Teva Pharmaceuticals, Frazer, PA, USA
e-mail: tedburczynski@gmail.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_47

Personalized Medicine: One Concept, Many Approaches

A great deal of excitement and hope has been generated by simply articulating the concept of "personalized medicine" within modern society, but this same excitement has also left open the possibility for oversimplification, misinterpretation, and – to some extent – even confusion regarding what PM really means today. For this reason it is useful to first consider the gamut of possible definitions for personalized medicine and then narrow down the definition to the most reasonable and practical one that pharmaceutical companies employ today. The pursuit of personalized medicines has revolutionized our industry's shared quest to create novel medicines that are more precise for treating ailments of patients around the globe. But what is PM, at its core?

Personalized medicine can take many forms, but for the purposes of this chapter, it is sufficient to consider three main categories of PM: (1) truly "individualized" therapies – literally, where a single medicine is created for a single patient; (2) "precision" medicines, which are medicines meant for all patients but which are accompanied by some sort of test to determine the overall likelihood of a beneficial effect; and finally (3) "targeted" medicines which are medicines specific to a single molecular target present within individuals and which can be robustly tested to identify those patients who should or should not receive the medicine.

We will briefly call the reader's attention to the extremely innovative and recently emerging "individualized PM" therapies first. We'll then focus the remainder of the chapter on the latter PM approaches of "precision medicine" and/or "targeted medicine" which are much more commonly pursued in drug development today. The reader should keep in mind that there is no consensus on exactly which of these definitions best reflect the idea of PM today, and we draw these distinctions only for the purpose of presenting the overall topic of PM in clearest fashion.

PM as Individualized Medicine

For the uninitiated, the entire notion of PM may immediately invoke a futuristic state in which a single patient afflicted with a disease might be diagnosed with high accuracy and a customized pill subsequently created which treats that single, particular patient with perfect precision. This is the most literal definition of personalized medicine, in which a truly "personalized" therapy including the necessary chemistry or biology would then be manufactured or synthesized in real time to address that individual patient's specific needs. While this territory has been well-tread in the plots of modern-day biomedical thriller novelists like Michael Crichton, Douglas Preston, and Michael Ransom . . . this is not (yet?) the norm we are living in today. It is likely to remain unrealistic for quite some time, despite the advancements in clinical sciences and in technology (e.g., 3D printing, CRISPR, and other technologies) which are just starting to illustrate how this future state may indeed 1 day evolve.

Despite the limitations, there are a small but growing number of truly personalized (individualized) medicines that have recently been approved by the FDA and/or are being developed today in the form of experimental therapies quite literally tailored to individual patients. While these emerged from academic translational laboratories in research hospitals initially, they have found their way into both the clinic and into pharmaceutical development. The best example of a truly individualized medicine comes from the groundbreaking field of immuno-oncology.

Immuno-oncology is beyond the scope of any single textbook or chapter or section . . . but in an oversimplified nutshell, it is the synthesis of the fields of immunology and oncology for the express purpose of reinvigorating the patient's own immune system to kill the tumor cells present in the host. While several classical monoclonal antibody-based IO therapies have been approved to date (Yervoy, Opdivo, and others). . . the FDA's approval of Yescarta® (axicabtagene ciloleucel, Kite Pharmaceuticals), a chimeric antigen receptor therapy (known as CAR-T) for the treatment

of adult patients with relapsed or refractory large B-cell lymphoma, officially ushered in the era of individualized therapy.

The CAR-T approach uses genetic engineering and adoptive transfer of a patient's own T-cells back into the host to attack the very tumors which initially evaded that patient's immune system and flourished in its presence. This complex and laborious research approach was the first of its kind to be evaluated as therapeutic at large pharmaceutical companies. The development process was used to determine whether it could be adapted out of the research hospital setting and used in standard clinical medicine environments and made available to patients with responsive forms of disease.

While CAR-T is technically a procedure, rather than a standardly administered medication, it is nonetheless the best example to date of how PM can actually refer to a truly individualized medicine concept. By its very nature, CAR-T can only work if it is used to create and express chimeric antigen receptors in the patient's own T-cells. Thus any future treatments based on an analogous approach – namely, of removing a patient's cells, genetically engineering them, and transferring them back into the patient, will constitute additional examples of individualized therapies in the years to come.

As will be discussed in the next section, however, in the vast majority of cases, PM does not refer to a single therapy being developed for a single individual, but instead reflects an approach in drug development to create medicines that will work well in specific subsets of patients.

PM as Precision Medicine

While we may still be a long way from routinely experiencing the futuristic scenarios of individualized medicines described above, recent progress has provided exciting new interventions, insights, and knowledge toward a state where we will no longer be satisfied with a one-size-fits-all approach to medicine: a state the pharmaceutical industry has started to accept and is embracing as illustrated by the rapidly increasing number of personalized medicines reaching the market.

As stated previously, the other definitions or views on PM include "precision medicine," which is a more technical view on delivering highly precise interventions to patients that are calculated based on a set of predefined variables (e.g., biomarkers, likelihood ratios of outcomes, etc.), or "targeted medicine" which refers to the use of therapeutic intervention for which the specific target is known, is present in a patient, and can be assayed using a test to determine the suitability of the therapy for that patient (there are many examples for this type of PM). These approaches are highly similar conceptually, and the main difference between them is that precision medicine ultimately relies on a test that gives a range of likely responses to a therapy across a population of patients, whereas a targeted medicine relies on a test that determines whether a patient should or should not be a candidate for therapy. Ultimately the key aspect common to both of these PM approaches is that they combine a patient's (measurable) characteristics (most often biomarkers) with the choice of a therapy to increase the likelihood of a positive outcome in the patient, both from the perspective of safety and efficacy.

How has PM impacted drug development? The pharmaceutical industry has been facing a challenge with the fact that new technologies capable of stratifying patients into subgroups with specific characteristics that render them more or less susceptible to response or harm are rapidly becoming available. The era of developing a chemical for the average patient population of which as many as 80% may derive no or limited benefit is coming to an end, which means that the market is becoming more segmented, but also bigger: diseases are no longer just described as "breast cancer" but as Her2-positive, ER-negative breast cancer, indicating that there are subtypes of disease with molecular markers that can be used to better characterize the disease and may resemble markers useful for targeted or individualized therapy, i.e., PM. Obviously this led to a significant and still present concern that PM by

creating subpopulations may decrease the size of the market and that it will be more challenging to develop new blockbuster therapies. This is largely accurate, but at the same time new opportunities have been created for the development of therapies that are more likely to succeed as the appropriate patient population can be more easily identified and, therefore, the chance of response to the therapy can be increased. In addition, the introduction of a truly personalized medicine into a highly fragmented market should in theory result in a greater share of the market than introduction of yet another "me-too" therapy. Ultimately several examples have illustrated that PM can be a powerful way to decrease the development time and cost for new therapies and a solid incentive for developing therapies in existing areas of medical need.

In addition to the challenges posed by discoveries that may be useful to identify responders, nonresponders, and/or patients at risk based on the presence or absence of certain markers, the field has provided new insights into targets and our understanding of disease and mechanisms leading to disease. Drugs have been developed targeting phenotypes with specific underlying genetic causes (e.g., Vertex CFTR) that wouldn't have been possible if not for a much deeper understanding of the molecular mechanisms causing the disease. The definitions and views on PM described above are therefore intricate to the goal of better understanding disease and using this knowledge in drug discovery and development. It is reasonable to assume that such understanding will provide the foundation of future drug development and that, therefore, it will become increasingly difficult for pharmaceutical manufacturers to ignore the scientific and clinical advancements that enable personalized medicine and the development of new therapies based on such knowledge.

Similar to drug development – or perhaps ahead of it – clinical practice has been impacted by the advent of new technologies allowing a better characterization of patients and disease. Based on increasingly accurate molecular diagnostics, the assessment of disease has become significantly more precise. On the one hand, this has allowed us to identify therapies that are more likely to benefit patients. On the other hand, it has created a situation where certain patient groups are now left with few therapeutic options: if it can be predicted that a therapy has no or only a very limited chance to improve a patient's ailment, it may leave the patient (and physician) without a reasonable option for treatment. Although this may be a scientifically sound argument, it leaves the compassionate aspect of treatment largely out of consideration. In addition, most therapies including targeted therapies do not portray a black-and-white response profile: therefore, there is always a (small) chance that a patient may benefit in the absence of a positive marker signal. Clinical practice is significantly more complicated than drug development from this perspective: how much better off are patients with the availability of PM? The dissection of disease in ever-finer subtypes is shining a light of how much (or little) we really understand about disease and the underlying pathophysiology. There is a transition phase – we are in the middle of it – where the gap between knowing whether or not a patient has a chance to respond to a certain treatment is ahead of the availability of treatments addressing all subtypes of disease.

Still, the knowledge that a patient has a greater likelihood of responding to a certain therapy is invaluable. Not only does it provide assurance to the treating physician that a treatment decision can be made based on up-to-date accurate information, but it also provides the rationale for payers to cover increasingly expensive therapies. It would not be feasible to reimburse therapies that cost many thousands of dollars if the response to no-response profile could not be improved. Taken together, the pharmaceutical industry has an opportunity to benefit from significant scientific advancements to develop more precise therapeutic options, which at the same time are more likely to be paid for if they address a truly unmet medical need. Patients benefit from a more precise assessment (diagnosis) of their disease and will have more and better

treatment options that are addressing the underlying pathophysiology, and payers have a better foundation for making coverage and reimbursement decisions.

The Path from One Gene–One Drug to Many Genes–One Drug

The past decade has seen the concurrent development – and success – of more targeted therapeutic approaches and molecular diagnostics, leading to better, more precise assessment of disease and therapeutic interventions. Precision medicine has transformed therapeutic alternatives in oncology and other clinical areas, driven by the rapid evolution of molecular diagnostics based on a more detailed understanding of disease and disease mechanism (pathophysiology).

The practice of medicine has always focused on specific phenotypic characteristics of each patient to determine how to treat them. However, the first demonstration of precision medicine requiring molecular diagnostics was the application of molecular diagnostics in the pharmacogenomics of drug metabolism enzymes (https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm079849.pdf). Mutations in the genes for these enzymes lead to incorrect dosing for many therapeutic drugs across multiple clinical areas (Phillips et al. 2001). The Critical Path for Innovation documents issued by the FDA in 2004 (http://wayback.archive-it.org/7993/20180125035500/https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/UCM113411.pdf) both acknowledged and encouraged the development of molecular diagnostic tests for drug metabolism enzyme mutations which would lead to accurate dosing. The list for therapies which the FDA recognizes today (https://www.fda.gov/drugs/scienceresearch/ucm572698.htm) as requiring testing to determine accurate dosing includes over 90 entries. Over the next decade, a second set of molecular diagnostics was developed, reflecting the development of targeted therapies in oncology (https://www.fda.gov/drugs/scienceresearch/ucm572698.htm) (Table 1).

In oncology, the relationship between molecular diagnostics and precision medicine was triggered by the approval of trastuzumab (https://www.accessdata.fda.gov/drugsatfda_docs/appletter/1998/trasgen092598L.pdf) targeting overexpression of the Her2 gene. Whether measuring genomic overexpression or genomic variants for patient selection, these tests identified patients likely to benefit from targeted therapies. Targeted therapies initially led to the concept that precision medicine was all about "one gene–one drug" (Jørgensen 2013). This simplified view of precision medicine kept corresponding molecular diagnostics simple in platform, analysis of results, and validation. Technologies (Brooks 1982; Luthra and Medeiros 2006) for the platforms of these molecular diagnostics predated precision medicine and drove a strictly incremental adoption of targeted therapies. Several studies (https://www.nature.com/articles/537S106a; https://www.nytimes.com/2010/06/15/health/15canc.html) published over the first decade of precision medicine questioned the therapeutic value of targeted therapies with companion diagnostics identifying patients with "one gene–one drug" molecular diagnostics.

Sanger sequencing for DNA (Cheong and Caramins 2014) was available at the onset of precision medicine. While it was the original platform for much of the knowledge available at the time about DNA sequences, it was also slow, expensive, and not suitable for molecular diagnostic applications. Next-generation sequencing (NGS) (Loewe 2013) developed throughout the first decade of precision medicine and transformed the speed, cost, and molecular diagnostic value of DNA sequences. For example, NGS results have allowed pathologists and oncologists in Molecular Tumor Boards optimizing therapeutic opportunities for their patients to look not only at single gene expression and mutations associated with targeted therapies, but to genes in molecular pathways related to target genes in oncology (Haslem et al. 2018). Decision-making by Molecular Tumor Boards is now

Table 1 Targeted therapies and biomarkers in oncology in the FDA table of pharmacogenomic biomarkers in drug labeling

Drug	Biomarker	Drug	Biomarker
Abemaciclib (1)	ESR (hormone receptor)	Letrozole	ESR, PGR (hormone receptor)
Abemaciclib (2)	ERBB2 (HER2)	Mercaptopurine	TPMT
Ado-trastuzumab Emtansine	ERBB2 (HER2)	Midostaurin (1)	FLT3
Afatinib	EGFR	Midostaurin (2)	NPM1
Alectinib	ALK	Midostaurin (3)	KIT
Anastrozole	ESR, PGR (hormone receptor)	Neratinib (1)	ERBB2 (HER2)
Arsenic trioxide	PML-RARA	Neratinib (2)	ESR, PGR (hormone receptor)
Atezolizumab	CD274 (PD-L1)	Nilotinib (1)	BCR-ABL1 (Philadelphia chromosome)
Avelumab	CD274 (PD-L1)	Nilotinib (2)	UGT1A1
Belinostat	UGT1A1	Niraparib	BRCA
Blinatumomab	BCR-ABL1 (Philadelphia chromosome)	Nivolumab (1)	BRAF
Bosutinib	BCR-ABL1 (Philadelphia chromosome)	Nivolumab (2)	CD274(PD-L1)
Brentuximab vedotin	ALK	Nivolumab (3)	Microsatellite instability, mismatch repair
Brigatinib	ALK	Obinutuzumab	MS4A1 (CD20 antigen)
Busulfan	BCR-ABL1 (Philadelphia chromosome)	Olaparib	BRCA
Cabozantinib	RET	Olaratumab	PDGFRA
Capecitabine	DPYD	Omacetaxine	BCR-ABL1 (Philadelphia chromosome)
Ceritinib	ALK	Osimertinib	EGFR
Cetuximab (1)	EGFR	Palbociclib (1)	ESR (hormone receptor)
Cetuximab (2)	RAS	Palbociclib (2)	ERBB2 (HER2)
Cisplatin	TPMT	Panitumumab (1)	EGFR
Cobimetinib	BRAF	Panitumumab (2)	RAS
Crizotinib (1)	ALK	Pazopanib (1)	UGT1A1
Crizotinib (2)	ROS1	Pazopanib (2)	HLA-B
Dabrafenib (1)	BRAF	Pembrolizumab (1)	BRAF
Dabrafenib (2)	G6PD	Pembrolizumab (2)	CD274 (PD-L1)
Dabrafenib (3)	RAS	Pembrolizumab (3)	Microsatellite instability, mismatch repair
Dasatinib	BCR-ABL1 (Philadelphia chromosome)	Pertuzumab (1)	ERBB2 (HER2)
Denileukin diftitox	IL2RA (CD25 antigen)	Pertuzumab (2)	ESR, PGR (hormone receptor)
Dinutuximab	MYCN	Ponatinib	BCR-ABL1 (Philadelphia chromosome)
Durvalumab	CD274 (PD-L1)	Rasburicase (1)	G6PD
Enasidenib	IDH2	Rasburicase (2)	CYB5R
Erlotinib	EGFR	Ribociclib (1)	ESR, PGR (hormone receptor)
Everolimus (1)	ERBB2 (HER2)	Ribociclib (2)	ERBB2 (HER2)
Everolimus (2)	ESR (hormone receptor)	Rituximab	MS4A1 (CD20 antigen)
Exemestane	ESR, PGR (hormone receptor)	Rucaparib (1)	BRCA
Fluorouracil (2)	DPYD	Rucaparib (2)	CYP2D6

(continued)

Table 1 (continued)

Drug	Biomarker	Drug	Biomarker
Fulvestrant (1)	ERBB2 (HER2)	Rucaparib (3)	CYP1A2
Fulvestrant (2)	ESR, PGR (hormone receptor)	Tamoxifen (1)	ESR, PGR (hormone receptor)
Gefitinib	EGFR	Tamoxifen (2)	F5 (factor V Leiden)
Ibrutinib (1)	Chromosome 17p	Tamoxifen (3)	F2 (prothrombin)
Ibrutinib (2)	Chromosome 11q	Thioguanine	TPMT
Imatinib (1)	KIT	Trametinib (1)	BRAF
Imatinib (2)	BCR-ABL1 (Philadelphia chromosome)	Trametinib (2)	G6PD
Imatinib (3)	PDGFRB	Trametinib (3)	RAS
Imatinib (4)	FIP1L1-PDGFRA	Trastuzumab (1)	ERBB2 (HER2)
Inotuzumab ozogamicin	BCR-ABL1 (Philadelphia chromosome)	Trastuzumab (2)	ESR, PGR (hormone receptor)
Irinotecan	UGT1A1	Tretinoin	PML-RARA
Lapatinib (1)	ERBB2 (HER2)	Vemurafenib (1)	BRAF
Lapatinib (2)	ESR, PGR (hormone receptor)	Vemurafenib (2)	RAS
Lapatinib (3)	HLA-DQA1, HLA-DRB1	Venetoclax	Chromosome 17p

based not only on that which is as certain as companion diagnostics but also on that which is less certain, on sequencing information which can contribute to the accuracy of this therapeutic decision-making.

The Path from Solid Tumor Biopsy Specimens to Liquid Biopsy Specimens

Accessibility to specimens that can be used to measure a patient's individual characteristic is the foundation of precision medicine. In oncology, such specimen is most often based on solid tumor biopsies. For the most part, therapeutic decision-making is possible with these specimens. However, there are tissues (such as lung tissue) for which solid tumor biopsies are challenging (Esposito et al. 2017) to obtain for molecular diagnostic therapeutic decision-making in oncology patients. In addition, tumor biopsy specimens are also difficult to use to monitor response and recurrence in oncology.

Therefore, liquid biopsies, i.e., the use of blood as a specimen to identify markers that are representative of a specific pathophysiology (e.g., circulating nucleic acids), have been used for prognostic applications in oncology over two decades. Figure 1 shows the different permutations for indications/uses, signal sources, and platforms for liquid biopsies.

Today, there is one FDA-approved product (https://www.accessdata.fda.gov/cdrh_docs/pdf15/P150047B.pdf) for EGFR-targeted therapeutic decision-making using cell-free DNA and qPCR. However, liquid biopsies are subject to fundamental limitations. For example, the sensitivity for detection of cell-free DNA from tumors as a fraction of total plasma cell-free DNA can be challenging, and the pipeline sof tware for the accurate identification of variants can be difficult to calibrate, leading to very low congruence for same patient-paired samples and differences between platforms (Torga and Pienta 2018).

Prognostic and Predictive Testing in Oncology: Case Study for Oncotype Dx

Precision medicine has been enabled by molecular diagnostics with prognostic and predictive indications. Oncotype DX is a case study for product development in these areas which illustrates the value of these tests for cancer patients

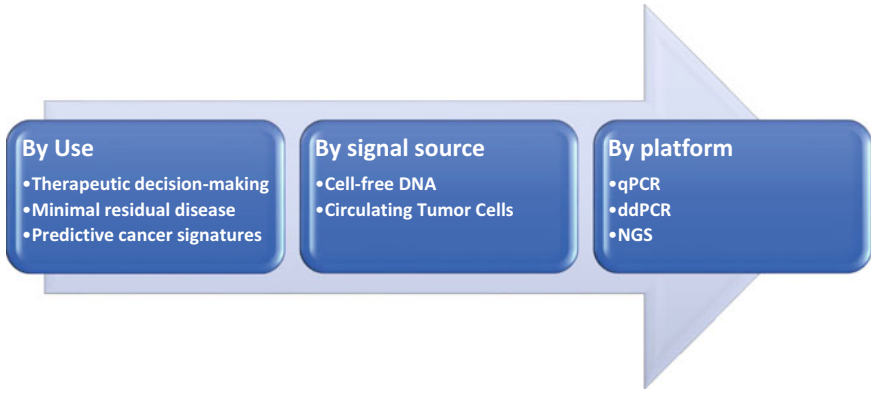


Fig. 1 The new frontier in precision medicine for oncology: liquid biopsies. Different uses, signal sources, and platforms define the products currently developed or reaching the market

and the complexity of product development, regulatory review, and product acceptance strategies to make them viable.

As of 2018, over 400 peer-reviewed papers reference this assay platform. The Oncotype DX® assay (https://preview.ncbi.nlm.nih.gov/pubmed/?term=Oncotype+DX) was developed to optimize the selection of adjuvant systemic therapy for patients with estrogen receptor (ER)-positive, lymph node-negative breast cancer. The Oncotype DX result is presented as the recurrence score which is a continuous score that predicts the risk of distant disease recurrence.

Implementation of the Oncotype DX assay has led to a change in how chemotherapy is utilized in patients with early stage, estrogen receptor (ER)-positive, lymph node-negative breast cancer (Mamounas et al. 2018). This test generates a recurrence score between 0 and 100, which correlates with probability of distant disease recurrence. Patients with low-risk recurrence scores (0–17) are unlikely to derive significant survival benefit with adjuvant chemotherapy and hormonal agents derived from using adjuvant hormonal therapy only. Conversely, adjuvant chemotherapy has been shown to significantly improve survival in patients with high-risk recurrence scores ($\geq$31). Patients at highest risk of recurrence are prescribed systemic treatment. Low-risk patients avoid adverse events from therapies unlikely to influence their survival.

Over the past decade, this test and others on this platform have opened genomic-based personalized cancer care for breast cancer in the USA. It is now widely utilized in various parts of the world (Giuliano et al. 2017; Jaafar et al. 2014; Rouzier et al. 2013). Together with several other genomic assays, Oncotype DX has been incorporated into clinical practice guidelines on biomarker use to guide treatment decisions. The assay has been validated for use in the prognostication and prediction of degree of adjuvant chemotherapy benefit in both lymph node-positive and lymph node-negative early breast cancers (Siow et al. 2018). Clinical studies have consistently shown that the Oncotype DX has a significant impact on decision-making in adjuvant therapy recommendations and appears to be cost-effective in diverse health-care settings.

Oncotype DX has succeeded, not only as a transformational tool for oncology patients but also as the most successful commercial platform, by far, in this class. This success, however, has been achieved independently of regulatory approval in the USA (Ross et al. 2008). Genomic Health considered the development of the algorithm used to determine the test score proprietary. A similar assay (MammaPrint) developed by Agendia received approval by the FDA (Brandão et al. 2018). Agendia agreed to submit its platform for approval by the FDA and to share the development data for its algorithm. After this approval, the FDA proposed a draft IVDMIA

guidance (later withdrawn) to address these types of assays (https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm079148.htm; https://www.alston.com/-/media/files/insights/publications/2010/08/recent-fda-actions-on-ldts-and-genetic-testing-whe/files/kazon_bna_art_8_25_10/fileattachment/kazon_bna_art_8_25_10.pdf).

Oncotype DX and MammaPrint are important case studies on the power of molecular diagnostics to transform cancer treatments and perhaps even transform regulatory policy: they redefined the relationship between successful analytical and clinical development and validation and its adoption by oncologists throughout the world for the benefit of their patients.

The Power of Precision Medicine in Rare Disease Therapeutic Development

Precision medicine has found one of its most successful applications in rare disease therapeutic product development. A rare disease often is originally defined through its clinical symptoms. In the absence of other information about the molecular mechanism of a rare disease, DNA sequencing data can empirically show germline mutations which correlate with the clinical symptoms associated with a rare disease (Bacchelli and Williams 2016). This empirical approach can be followed, as it would in oncology, to identify enrichment biomarkers or companion diagnostics for patient selection.

Cystic Fibrosis

Rare diseases for which molecular mechanisms have been exhaustively studied open novel applications for precision medicine. For example, cystic fibrosis is a rare disease (30,000 patients in the USA) (https://www.cff.org/What-is-CF/About-Cystic-Fibrosis/) for which the molecular mechanism is well established. Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene (https://www.cff.org/What-is-CF/Genetics/Types-of-CFTR-Mutations/). Mutations in CFTR affect gating and other structural features and the number of active copies in the membrane (https://cftr2.org/).

There are two types of biomarkers associated with these mutations. These mutations can be used as patient selection markers in clinical studies for targeted therapies (Merk and Schubert-Zsilavecz 2013). From a regulatory perspective, these mutations could develop into CDx tests, but over 95% of CF patients in the USA have been genotyped for CFTR mutations by the Cystic Fibrosis Foundation (Wiencek and Lo 2018), and the FDA has also cleared (https://www.accessdata.fda.gov/cdrh_docs/reviews/K132750.pdf) the Illumina CFTR Clinical Genotyping NGS Test on the MiSeq Dx to identify any previously unknown variants in these patients.

Cystic fibrosis may represent an extreme opportunity for precision medicine as there may be as little as a single individual associated with a specific CFTR mutation (sporadic or familial mutations). Clinical study designs focused on individual CFTR mutations are not possible for most CFTR mutations (as they wouldn't be for any other such disease in which sporadic mutations occur). Therefore, a clinical study design strategy consistent with this limitation is one where patients with multiple CFTR mutations are combined into patient populations with a shared mutational class. While this strategy has succeeded in designing clinical studies for therapeutic products such as Kalydeco (Moran 2017; Linsdell 2017), it has still not been broadly accepted by the FDA for label expansion.

The second type of biomarker associated with these mutations is the in vitro electrophysiological measurement of chloride transport in cells isolated from CF patients (Cholon and Gentzsch 2018). This test is used to quickly identify patients who are likely to benefit from Kalydeco and other CF therapies (https://www.fda.gov/Drugs/NewsEvents/ucm559051.htm) and may be viewed as a surrogate for a specific phenotype of the disease. Importantly, this test has also led to Kalydeco label expansion claims for 23 mutations represented in a population of at least 900 patients

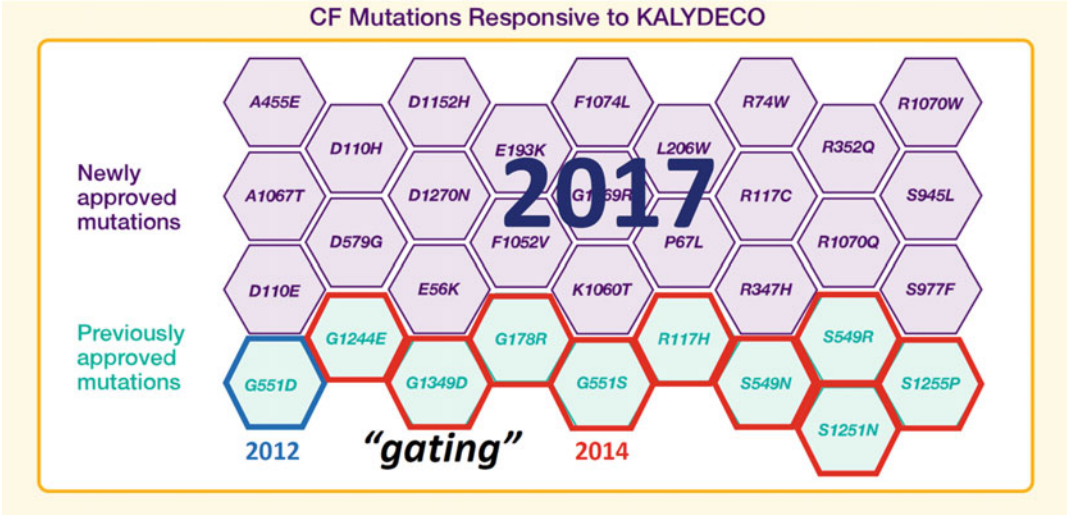


Fig. 2 Cystic fibrosis mutations responsive to Kalydeco

(https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm559212.htm). This unique precedent at the FDA could potentially lead in the future to the application of such surrogate tests for the initial approval of new therapies for rare diseases (Fig. 2).

Duchenne Muscular Dystrophy

The molecular defect in Duchenne muscular dystrophy is a defective dystrophin molecule (Hoffman 1989). This is a very large protein, transcribed from multiple exon sections (Gao and McNally 2015). Exon-skipping therapy is a structural strategy to partially correct the defect in dystrophin by splicing out exon 51 from transcription of the complete dystrophin molecule (Yokota et al. 2007). The resulting (shorter) dystrophin molecule does not have the full activity of normal dystrophin, but shows improved performance compared to the version with a detective exon 51 (Lim et al. 2017). Patients with this exon 51-skipped version have improved outcomes in their physiology and life expectancy. There are other defective exons transcribed into dystrophin for which this same strategy can be applied (Lee et al. 2017).

The exon 51-skipping therapy has been available throughout the past decade. However, its regulatory approval pathway throughout this period highlights the major challenges faced in the development and approval of novel precision therapies for rare diseases. There are two broad areas in which this challenge has had a major impact on how – and how long – this regulatory pathway has been drawn:

1. Acceptance by the FDA of a dystrophin activity as a surrogate biomarker for exon-skipping therapies
2. Clinical study designs which reflect both the small patient populations available for rare disease clinical development

To this date, a standard endpoint in clinical trials for the approval of rare disease therapies associated with at the FDA is the 6-minute walk (Hamuro et al. 2017). This surrogate reflects the regulatory need to show an improvement in how a patient feels, functions, or survives. A decade ago, this surrogate biomarker was required by the FDA to show therapeutic efficacy in Duchenne muscular dystrophy. This surrogate, however, is unable to capture accurately the immediate mechanistic effects of exon-skipping therapies in

Duchenne muscular dystrophy. The FDA now allows the use of dystrophin activity as an alternative surrogate to show efficacy in this disease.

As with other rare diseases, clinical study designs in Duchenne muscular dystrophy are challenged by the limited number of patients available, the broad heterogeneity in disease symptoms, and the difficulty in the delivery of therapies such as exon skipping to their specific tissue targets. Studies to test the efficacy of exon-skipping therapies have been crippled by their heterogenous results (Randeree and Eslick 2018). Notwithstanding the broad range of therapeutic responses in these studies, the FDA approved the first indication for exon 51-skipping therapy in 2016 (Syed 2016).

Precision Medicine Biomarkers

Precision medicine biomarkers start from two main sources:

- Biological hypotheses associated with the therapeutic target
- Empirical data obtained from early clinical studies with the proposed therapy

A biological hypothesis for a targeted therapy leads to biological pathways where specific patient selection biomarkers can be initially proposed (Jones and Libermann 2007). Early clinical studies with a limited number of patients, such as Phase 2a studies, can be used to prioritize possible patient selection biomarkers (Le Tourneau et al. 2008). In precision medicine applications, these biomarkers are detected using primarily genomic (variants or expression-level measurements) (Lin et al. 2017) or immunohistochemistry (https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm094002.htm) platforms. As shown above, genomic platforms for patient selection biomarkers have evolved from those assays which genotype single genes to those which genotype multiple genes through NGS. Immunohistochemistry platforms continue to be important for semiquantitative assessments of cell surface proteins.

Cell sorting is required for liquid biopsy assays for circulating tumor cells (CTCs) (Alix-Panabières and Pantel 2013). This platform has been in use for several decades (Krebs et al. 2011). Basic cell-sorting technology is mature, but derivative technologies required for accurate selection and detection of specific CTC types have been developed over this period (Andree et al. 2018). The original parameter measure in CTC liquid biopsies was cell enumeration for nonspecifically or specifically labeled cells. More recently, NGS sequencing for DNA isolated from CTCs has also been considered as a possible liquid biopsy platform (Dawson 2018).

Digital devices are used in precision medicine to assess the physiological status of patients with neurodegenerative diseases like multiple sclerosis (Marziniak et al. 2018). Devices as simple as smartphones can be used either as substitutes for paper-based PROs and CROs or as real-time physiological biomarkers (Shah et al. 2016). These digital devices can measure real-time physiological biomarkers such as gait or other measures of activity. These parameters can be measured at baseline and then followed throughout treatment, providing precision measurements of therapeutic efficacy.

The Path from Enrichment Biomarkers to Companion Diagnostics

Clinical study designs in precision medicine for therapeutic product approval at the FDA make use of proposals for enrichment biomarkers (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf) for patient selection to maximize therapeutic effects in the selected patient population. Patient selection biomarker proposals often start with a biological hypothesis linked to the targeted therapy. The major hurdle with these initial proposals is the limited number of patients available for validation of these biomarkers in preliminary (e.g., Phase 2a) studies with a novel

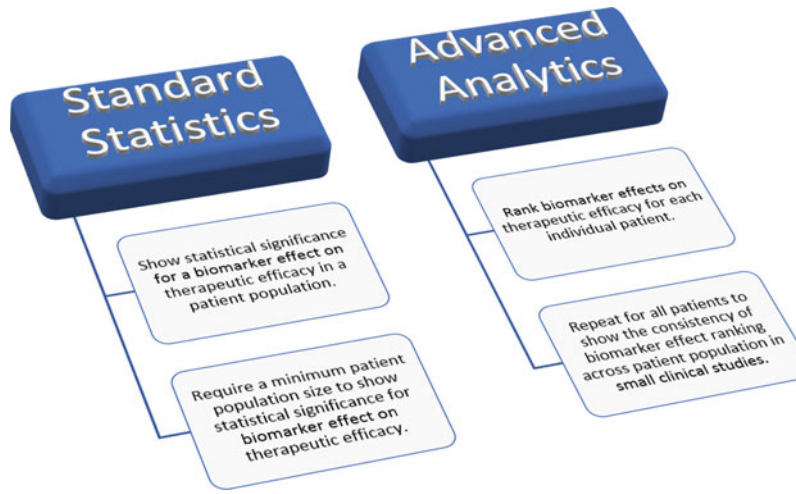
therapy. Standard statistical methods face challenges in ranking biomarkers tested in these preliminary studies because of the limited number of patients in these preliminary studies. This is one of the most difficult steps in the development of therapeutic products for precision medicine. In particular, a decision *how* to include biomarker candidates into clinical trial designs is critical: for the most part, it is advisable that specimens are collected to ensure the analysis of markers at least retrospectively, but if the evidence is strong enough, the marker(s) can be used for stratification, in which case the first steps toward the development of a companion diagnostic has been taken.

In order to address the issue of limited number of samples, advanced analytics platforms such as KEM® (Ariana Pharma) (http://www.arianapharma.com/2017/10/anavex-life-sciences-reports-pk-and-pd-data-from-phase-2a-trial-of-anavex2-73-in-mild-to-moderate-alzheimers-disease-patients/) focus on ranking the biomarkers tested in preliminary studies. The platform does not focus on a standard statistical significance for specific biomarker results but rather measures the link between observed therapeutic effects and different biomarkers included in a preliminary study. The tool then ranks all biomarkers across all patients in a study and allows a quick assessment about the homogeneity of biomarker effects across small patient populations:

This approach can allow one to select a potential companion diagnostic biomarker accurately from preliminary (including pilot) studies such as Phase 1 in patients or Phase 2a studies. Subsequently, a new study, e.g., a Phase 2b study, can be designed and tailored toward demonstrating the therapeutic efficacy in the newly selected patient population. Simultaneously, this study can then also serve as validation of the test as a companion diagnostic.

Conclusion

Personalized medicine has already had a significant impact in today's pharmaceutical development and patient care. By better characterizing patients and disease through sophisticated diagnostics and the use of advanced statistical tools and clinical trial designs, we are now routinely optimizing therapies. Personalized medicine can lead to the development of truly individualized medicines, as is the case with CAR-T therapies, or it can lead to the development of precision and/or targeted medicines, which increase the likelihood that therapies will work in specific patients. Personalized medicine is also no longer an oncology-delimited consideration, and there are now many examples of these approaches having success in many other therapeutic areas. Ever-improving methods for detecting genetic alterations, biomarkers of



many different types, and other measurable characteristics are enabling cutting-edge translational research in these disease areas that will ultimately lead to personalized therapies for years to come. The question of whether to pursue a personalized medicine approach in drug development has evolved in the past few years, growing from an esoteric possibility into an absolutely critical-path consideration for every new target and associated therapeutic candidate entering pharmaceutical development today. Our ability to embrace and implement personalized medicine approaches in pharmaceutical development will be a key determinant of our collective ability as drug developers to address truly unmet medical needs of patients in the future.

References and Further Reading

Alix-Panabières C, Pantel K (2013) Circulating tumor cells: liquid biopsy of cancer. Clin Chem 59(1):110–118

Andree KC, Mentink A, Zeune LL, Terstappen LWMM, Stoecklein NH, Neves RP, Driemel C, Lampignano R, Yang L, Neubauer H, Fehm T, Fischer JC, Rossi E, Manicone M, Basso U, Marson P, Zamarchi R, Loriot Y, Lapierre V, Faugeroux V, Oulhen M, Farace F, Fowler G, Sousa Fontes M, Ebbs B, Lambros M, Crespo M, Flohr P, de Bono JS (2018) Towards a real liquid biopsy in metastatic breast and prostate cancer: diagnostic Leukapheresis increases CTC yields in a European prospective multi-center study (CTCTrap). Int J Cancer. https://doi.org/10.1002/ijc.31752

Bacchelli C, Williams HJ (2016) Opportunities and technical challenges in next-generation sequencing for diagnosis of rare pediatric diseases. Expert Rev Mol Diagn 16(10):1073–1082

Brandão M, Pondé N, Piccart-Gebhart M (2018) Mammaprint™: a comprehensive review. Future Oncol. https://doi.org/10.2217/fon-2018-0221

Brooks JJ (1982) Immunohistochemistry of soft tissue tumors: progress and prospects. Hum Pathol 13(11):969–974

Cheong PL, Caramins M (2014) Approaches for classifying DNA variants found by Sanger sequencing in a medical genetics laboratory. Methods Mol Biol 1168:227–250

Cholon DM, Gentzsch M (2018) Recent progress in translational cystic fibrosis research using precision medicine strategies. J Cyst Fibros 17(2S):S52–S60

Dawson SJ (2018) Characterizing the cancer genome in blood. Cold Spring Harb Perspect Med. https://doi.org/10.1101/cshperspect.a026880. pii: a026880

Esposito A, Criscitiello C, Trapani D, Curigliano G (2017) The emerging role of "Liquid Biopsies," Circulating tumor cells, and circulating cell-free tumor DNA in lung cancer diagnosis and identification of resistance mutations. Curr Oncol Rep 19(1):1

Gao QQ, McNally EM (2015) The dystrophin complex: structure, function, and implications for therapy. Compr Physiol 5(3):1223–1239

Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, Weaver DL, Winchester DJ, Hortobagyi GN (2017) Breast cancer-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. CA Cancer J Clin 67(4):290–303

Hamuro L, Chan P, Tirucherai G, AbuTarif M (2017) Developing a natural history progression model for Duchenne muscular dystrophy using the Six-Minute Walk Test. CPT Pharmacometrics Syst Pharmacol 6(9):596–603

Haslem DS, Chakravarty I, Fulde G, Gilbert H, Tudor BP, Lin K, Ford JM (2018) Nadauld LD precision oncology in advanced cancer patients improves overall survival with lower weekly healthcare costs. Oncotarget 9(15):12316–12322

Hoffman EP (1989) Human molecular genetics and the elucidation of the primary biochemical defect in Duchenne muscular dystrophy. Cell Motil Cytoskeleton 14(1):163–168

Jaafar H, Bashir MA, Taher A, Qawasmeh K, Jaloudi M (2014) Impact of oncotype DX testing on adjuvant treatment decisions in patients with early breast cancer: a single-center study in the United Arab Emirates. Asia Pac J Clin Oncol 10(4):354–360

Jones J, Libermann TA (2007) Genomics of renal cell cancer: the biology behind and the therapy ahead. Clin Cancer Res 13(2 Pt 2):685s–692s

Jørgensen JT (2013) Companion diagnostics in oncology – current status and future aspects. Oncology 85(1):59–68

Krebs MG, Sloane R, Priest L, Lancashire L, Hou JM, Greystoke A, Ward TH, Ferraldeschi R, Hughes A, Clack G, Ranson M, Dive C, Blackhall FH (2011) Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. J Clin Oncol 29(12):1556–1563

Le Tourneau C, Vidal L, Siu LL (2008) Progress and challenges in the identification of biomarkers for EGFR and VEGFR targeting anticancer agents. Drug Resist Updat 11(3):99–109

Lee T, Awano H, Yagi M, Matsumoto M, Watanabe N, Goda R, Koizumi M, Takeshima Y, Matsuo M (2017) 2′-O-methyl RNA/ethylene-bridged nucleic acid chimera antisense oligonucleotides to induce dystrophin exon 45 skipping. Genes (Basel) 8(2):67

Lim KR, Maruyama R, Yokota T (2017) Eteplirsen in the treatment of Duchenne muscular dystrophy. Drug Des Devel Ther 11:533–545

Lin CH, Lin E, Lane HY (2017) Genetic biomarkers on age-related cognitive decline. Front Psych 8:247

Linsdell P (2017) Structural changes fundamental to gating of the cystic fibrosis transmembrane conductance regulator anion channel pore. Adv Exp Med Biol 925:13–32

Loewe RP (2013) Combinational usage of next generation sequencing and qPCR for the analysis of tumor samples. Methods 59(1):126–131

Luthra R, Medeiros LJ (2006) TaqMan reverse transcriptase-polymerase chain reaction coupled with capillary electrophoresis for quantification and identification of BCR-ABL transcript type. Methods Mol Biol 335:135–145

Mamounas EP, Russell CA, Lau A, Turner MP, Albain KS (2018) Clinical relevance of the 21-gene recurrence score® assay in treatment decisions for patients with node-positive breast cancer in the genomic era. NPJ Breast Cancer 4:27

Marziniak M, Brichetto G, Feys P, Meyding-Lamadé U, Vernon K, Meuth SG (2018) The use of digital and remote communication technologies as a tool for multiple sclerosis management: narrative review. JMIR Rehabil Assist Technol 5(1):e5. https://doi.org/10.2196/rehab.7805

Merk D, Schubert-Zsilavecz M (2013) Repairing mutated proteins – development of small molecules targeting defects in the cystic fibrosis transmembrane conductance regulator. Expert Opin Drug Discovery 8(6):691–708

Moran O (2017) The gating of the CFTR channel. Cell Mol Life Sci 74(1):85–92

Phillips KA, Veenstra DL, Oren E, Lee JK, Sadee W (2001) Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. JAMA 286(18):2270–2279

Randeree L, Eslick GD (2018) Eteplirsen for paediatric patients with Duchenne muscular dystrophy: a pooled-analysis. J Clin Neurosci 49:1–6

Ross JS, Hatzis C, Symmans WF, Pusztai L, Hortobágyi GN (2008) Commercialized multigene predictors of clinical outcome for breast cancer. Oncologist 13(5):477–493

Rouzier R, Pronzato P, Chéreau E, Carlson J, Hunt B, Valentine WJ (2013) Multigene assays and molecular markers in breast cancer: systematic review of health economic analyses. Breast Cancer Res Treat 139(3):621–637

Shah KN, Hofmann MR, Schwarzkopf R, Pourmand D, Bhatia NN, Rafijah G, Bederman SS (2016) Patient-reported outcome measures: how do digital tablets stack up to paper forms? A randomized, controlled study. Am J Orthop (Belle Mead NJ) 45(7):E451–E457

Siow ZR, De Boer RH, Lindeman GJ, Mann GB (2018) Spotlight on the utility of the Oncotype DX® breast cancer assay. Int J Womens Health 10:89–100

Syed YY (2016) Eteplirsen: first global approval. Drugs 76(17):1699–1704

Torga G, Pienta KJ (2018) Patient-paired sample congruence between 2 commercial liquid biopsy tests. JAMA Oncol 4(6):868–870

Wiencek JR, Lo SF (2018) Advances in the diagnosis and management of cystic fibrosis in the genomic era. Clin Chem 64(6):898–908

Yokota T, Pistilli E, Duddy W, Nagaraju K (2007) Potential of oligonucleotide-mediated exon-skipping therapy for Duchenne muscular dystrophy. Expert Opin Biol Ther 7(6):831–842

Noninvasive Methodology (NMR)

22

Mitul A. Mehta

Contents

Abstract

Neuroimaging with MRI provides a noninvasive means to assess drug effects in vivo. In addition to the discovery of potential markers of psychopathology, MRI methods can be used to test existing and novel compounds. The assessments can be of metabolite levels, task-based brain activation, brain connectivity, drug-related activation, and quantitative perfusion. These methods are pharmacodynamic in nature and can also be used to describe pharmacokinetic–pharmacodynamic relationships. They complement emission tomography assessments of brain penetration and dose-occupancy relationships and can extend or even be a substitute for these methods when ligands are not available, with particular value when the desired outcome is intermediate markers of function. Limitations and challenges of these methods are intrinsic to the

M. A. Mehta (✉)
Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
e-mail: mitul.mehta@kcl.ac.uk

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_46

measurement, such as vascular artifacts, but they can be overcome with additional assessments.

Purpose and Rationale

Neuroimaging with nuclear magnetic resonance (NMR) technology, commonly referred to as magnetic resonance imaging (MRI), is a noninvasive method to record from the brain in vivo. In this respect, it differs from intracortical recording techniques – which require surgery and necessarily measure highly localized activity – and emission tomographic techniques which require the injection of radiotracers. MRI has better spatial resolution than surface-based brain functional measurement methods such as electroencephalography and near-infrared spectroscopy, and MRI is able to measure from both deep and superficial structures. The potential role in drug discovery and evaluation of NMR technology is in three main areas:

1. Discovery and validation of neuropharmacological mechanisms associated with pathophysiology
2. Developing imaging markers associated with dysfunction as targets for drug evaluation
3. Testing compounds using imaging markers to determine their pharmacodynamic effects

These methods can be used to answer key questions, principally including, dose-response evaluation of the *functional* effects of compounds and mechanistic validation of compounds. The utility of MRI can be at any phase of drug discovery and development. Preclinical imaging can be used in both discovery and evaluation, particularly when used in combination with animal models of dysfunction. There are strengths in preclinical studies in the ability to test multiple doses of drugs, with pharmacokinetic ranges that may not be possible in humans. Moreover, preclinical MRI can be combined with other techniques to provide comprehensive assessments and control for potential confounds such as measurement specific artifacts (Coimbra et al. 2013; Jonckers et al. 2015). While remaining valuable as a methodology, good examples of translational viability from animal models into successful human trials are limited. When using markers of function, such as functional neuroimaging with MRI, drugs can potentially be tested very early in the developmental pathway in humans, and the focus of this chapter will be on human studies. Such studies will typically be in Phase 1b or early in Phase II. Examples of potential utility include (i) the testing that a compound aimed at improving cognitive dysfunction mediated by a specific brain circuit, actually modulates these circuits in healthy volunteers, giving an early indicator of the potential for efficacy before the drug is administered to patients, (ii) testing that a compound reverses the effects of a drug known to model a component of a disorder, providing evidence of a specific pharmacological mechanism in vivo, and (iii) testing the effects on specific brain circuits in patients with known impairment in such systems, giving an early indicator of the potential for efficacy in the target population.

Procedures

There are five principal methods of imaging with MRI to evaluate drug effects, which are depicted in Fig. 1. These methods offer the ability to perform multimodal studies, in that the different methods are sensitive to different aspects of a drug effect, and can be deployed in the same study sessions. Importantly, these methods can also be tailored to each study to ensure that the pharmacodynamic measures capture the desired window within the pharmacokinetic profile (Deakin et al. 2008; Paloyelis et al. 2016; De Simoni et al. 2013). They each have limitations and advantages; for example, all the methods in Fig. 1 (except MRS) rely on the concept of neurovascular coupling (Mathias et al. 2018); when a brain area is more physiologically and electrically active, then there is a cascade of changes which result in more blood flow being delivered to the active area. As another example, task-based fMRI, typically collected using blood-oxygen-level dependent (BOLD) imaging, is

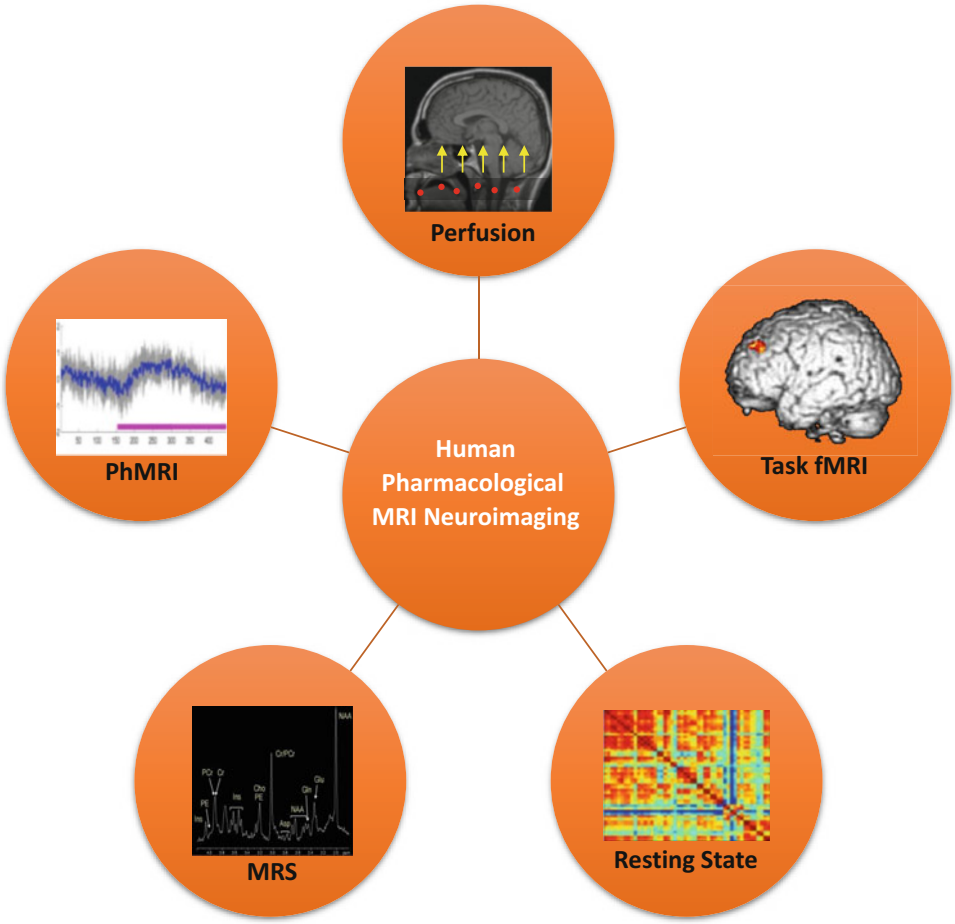


Fig. 1 Five principal methods for measuring drug effects with MRI. Perfusion imaging or cerebral blood flow mapping can be used to quantify drug effects, as is particularly useful in comparing across sessions and correcting within session for baseline differences on the vasculature; task-based fMRI is useful when the effect on the brain activity or connectivity during specific contexts is required; resting state imaging is useful to investigate network level effects on connectivity between regions in a context independent state; MRS is useful in measuring specific metabolite levels (e.g., glutamate); phMRI is a method to track the main effects of typically rapidly acting drugs within a session

nonquantitative but can be applied with reasonable temporal (2s) and spatial (~2–3mm) resolution.

Perfusion Imaging

Perfusion imaging with MRI can be achieved through the use of a contrast agent such as gadolinium, but for research purposes, it is the advent of the endogenous contrast agent, allowing non-invasive data acquisition that has seen this method grow in use for pharmacological imaging studies (Zelaya et al. 2015).

Arterial spin labelling (ASL) is the technique of choice for perfusion imaging, in which two types of image are acquired, label and control. The label image is acquired after encoding of the inflowing blood, and the control image will have the same acquisition parameters, but different encoding. The difference between these two images is regionally weighted by the perfusion in the brain tissue, often referred to as the regional cerebral blood flow (rCBF). A separate image

with the same acquisition parameters but sensitive to the proton density (an Mzero image) is required for quantification (Wong et al. 1997). There are two main forms of ASL:

(i) Pulsed ASL (or pASL) involves the labelling of arterial blood by inverting signal over a relatively wide volume (~100 mm) below the area of data acquisition, typically in the vicinity of the carotid arteries. This is achieved with tailored pulses where the B1-field profile is homogeneous over a wide region (Wong et al. 1997). After a short delay, the imaging data is collected. If this is implemented with a single-shot acquisition, such as echo-planar imaging, then the time for volume acquisition can be similar to functional magnetic resonance imaging (see below). The control image can be collected using double inversion of the arterial blood and collected interleaved with the label images. These "control-label pairs" can be treated as a timeseries, although a single pair is insufficient to provide a robust and reliable perfusion map. The timeseries of pairs can also be separated into blocks to produce a series of perfusion-weighted volumes with significantly enhanced contrast-to-noise ratio (CNR) over the individual pairs.

(ii) Continuously labelled ASL (CASL) or "pseudo-continuously" or "pulsed-continuously" labelled ASL (pCASL) has superior labelling efficiency to pASL. This is achieved by arterial blood being labelled by flow-driven inversion of arterial blood in the presence of a magnetic field gradient, applied in the direction of the moving blood into the brain. A labelling plane is used which leads to a rotation of the longitudinal magnetization vector about the effective field, which in turn changes the orientation of the spins by 180° as they cross the labelling plane (Dai et al. 2008). This method is extremely effective at labelling a substantial amount of arterial blood, which allows for the collection of image data using multishot acquisition techniques, which can give whole brain coverage.

ASL carries a number of advantages for pharmacological imaging (Zelaya et al. 2015).

1. Determine the average effects of a drug on rCBF across the brain, referred to as "global effects."
2. Quantification of the changes induced by a drug on rCBF. The most common technique for functional brain imaging, BOLD, is non-quantitative and dependent on the cerebral metabolic rate of oxygen utilization, blood flow, and volume, whereas perfusion imaging reports on rCBF. Indeed, ASL could be used to "calibrate" BOLD imaging.
3. Assessment of drug effects over multiple sessions. These sessions may be separated by seconds, minutes, hours, days, or longer. It is the quantitative and noninvasive nature of the method that allows multiple assessments over time, which may be not be possible when using alternative methods. Emission tomography may be limited by the half-life of some tracers and radiation dosimetry restrictions. Electroencephalography (EEG) may be restricted by the number of times the electrode cap can be removed and refitted without significant artifacts. BOLD imaging is limited due to low frequency fluctuations within sessions and the nonquantitative nature confining comparisons across sessions to interaction effects.
4. The determination of context-independent drug effects acquired over a number of minutes. These so-called tonic effects offer the advantage when assessing interactions between drugs and the state of the system, notably including pain states and affective states.
5. The determination of context-dependent drug effects on functions/processes which change at very low frequencies. For example, learning paradigms, where the learning occurs over a period of minutes when BOLD contrast data are dominated by low frequency fluctuations. The use of control-label pairs markedly reduces sensitivity to the sources of low frequency noise (e.g., scanner temperature fluctuations).

ASL has very good test-retest reliability (intraclass correlation coefficients >0.80) (Li et al. 2018) making a stable measurement, useful for within-subjects studies, where a drug can be compared to placebo or other active controls. Indeed, ASL methodology is sensitive to a variety of drugs. With fast-acting compounds such as fentanyl, pulsed techniques have been used to chart the rapid changes in the brain (Zelaya et al. 2015). With oral drugs, cASL methods have been used to both describe and discriminate different compounds. For example, different antipsychotic drugs are poorly discriminated by their efficacy in patients with few exceptions, but cASL has demonstrated a difference between the older dopamine D2 receptor antagonist haloperidol and the partial agonist aripiprazole in healthy volunteers (Handley et al. 2013). Aripiprazole had more effects in the cortex, while haloperidol produced a greater increase in rCBF in the ventral striatum. These localized effects are important for understanding potential differences in regional functionality between drugs for the same indication. Another example is from a study of methylphenidate and atomoxetine, both indicated as a treatment for attention deficit hyperactivity disorder, but with known differences in their pharmacology. ASL has been used to demonstrate the cortically and subcortically distributed changes in rCBF that allow the drugs to be discriminated from each other and placebo (Marquand et al. 2012), with methylphenidate have more prominent effects in a set of regions including subcortical structures and atomoxetine having more prominent effects in a set of cortical regions.

It is only relatively recently that there have been attempts to harmonize ASL methods across scanner manufacturers with recommendations for multisite studies (Alsop et al. 2015). Indeed, the number of studies using ASL to describe drug effects is low compared to BOLD fMRI. The tension between optimization and harmonization is present in individual studies, although sharing of data across centers is a drive to standardize the acquisitions as much as possible to address questions that are beyond individual studies. This has been successfully applied to analgesic compounds to demonstrate shared and distinct regions in the brain for their functional effects (Duff et al. 2015).

Task-Based fMRI

Task-based neuroimaging with MRI is almost exclusively implemented with BOLD contrast (Huettel et al. 2014). This method relies on functional hyperemia, that is the arterial delivery of oxygenated blood at a level that exceeds the need in the target region. The venous outflow from an active region will necessarily have a higher level of deoxygenated hemoglobin, but the hyperemia leads to a paradoxical decrease in the relative concentration of deoxyhemoglobin. Because deoxyhemoglobin is paramagnetic, the local distortions in the magnetic field are reduced in more active regions increasing the MRI signal that is sensitive to spin-spin relaxation times (T2* and T2*-weighted signal). When combined with a fast readout, whole brain images of the deoxygenated hemoglobin sensitive contrast can be collected with a temporal resolution of about 2 s and a reasonable spatial resolution, with voxels of approximately 3 mm isotropic. The advantages of the BOLD method are the rapid acquisition of whole brain images, sensitivity to deep brain structures, and standardized processing pipelines that are publicly available. The typical signal change with task activation can be as high as 3%, usually in primary sensory cortices, but is typically fractions of a percent. A disadvantage of the BOLD contrast is that following neuronal activity the signal change peaks a number of seconds later before slowly returning to baseline. This is known as the hemodynamic response function and needs to be factored in to any data modelling for task-based fMRI. Another challenge to BOLD imaging is the influence of movement which must be accounted for in the analysis pipeline. Susceptibility artifacts result from areas of the imaging field of view that alter the local gradients in the magnetic field leading to loss of contrast. Air-tissue boundaries produce loss of contrast in orbitofrontal cortex and the temporal pole, often extending into the basal forebrain. This can result in additional loss of signal in the ventral portions of the striatum and the amygdala. This

potential for artifacts and data loss emphasizes the need for excellent quality control processes in data processing in order to account for scans with loss of contrast, or problematic movement. In some cases, scan volumes or entire scan runs will need to be removed from the analysis. Building such controls into the analysis pipeline is crucial to ensure decisions about data quality and inclusion are made independent of the actual analysis (Schwarz et al. 2011).

Timeseries Correlations and Resting State BOLD

Resting state BOLD imaging involves the acquisition of BOLD contrast timeseries data without the use of a task. Participants typically are asked to lie still in the scanner with their eyes open or closed and not to engage in any particular cognitive activity. The most common use of resting state BOLD is to derive connectivity metrics from the data. Connectivity is a term used to describe the correlation, or coherence, in signals from different brain regions. If regions correlate, they are considered functionally connected. The correlation between any two defined regions can be examined at the level of a region of interest (usually derived from a predefined atlas) or at the voxel level. The connectivity described can be a summary of the entire timeseries, or index dynamic variations in connectivity over time. An advantage of resting state imaging is the simplicity with which it can be implemented, requiring little or no training of the subject or experimenters. A challenge is the potential for artifacts in the correlation of timeseries, such as noise in the BOLD contrast data from the scanner (e.g., low frequency fluctuations) or from the subject (e.g., physiological noise from the pulsatile nature of blood flow, or from respiratory noise). A method to account for these artifacts must be included in the data analysis. A common method is to band pass filter the data.

Sophisticated data processing and analysis methods can be applied to resting state BOLD data to extract network or topological parameters. Independent component analysis can be used to extract components in the data which are separated at spatial or temporal scales. These have two important uses. The first is to derive separable networks that have been linked to distinct functional domains (Smith et al. 2013) and the second is a means to identify and remove components of the data that are related to noise, such as movement or respiration (Pruim et al. 2015). Graph theory is applied to resting state BOLD timeseries data and can be used to calculate the correlation of each brain area to all others, the pathways between any two areas, the efficiency of the system in terms of connection pathways, and other metrics such as small word topology (Bullmore and Sporns 2009).

Connectivity analysis can also be applied during the performance of cognitive tasks. The correlation of the timeseries BOLD data can be hypothesized to alter during changes in cognitive demands (Friston et al. 1997).

Magnetic Resonance Spectroscopy

MRS is used to study metabolite levels in brain tissue (for review, see Buonocore and Maddock 2015). While protons can align along the applied magnetic field, 31P and 13C MRS can also be utilized, and 23Na and 19F are options for MRS investigation, but additional equipment may be needed. At the core of MRS is the fact that the magnetic field a particular atom experiences is affected by its local chemical environment, specifically the magnetic field from nearby motion of electrons. The consequence of experiencing slightly different applied fields due to the varied chemical environment is that the atoms resonate at slightly different frequencies. This "chemical shift" effect can be used to detect different environments of relevant nuclei. It is the chemical shift that gives rise to a MR frequency spectrum consisting of nuclei which resonate at different frequencies, depending on their local environment. The frequencies are not related to the exact concentration of nuclei and depend on the exact magnetic field strength. Therefore, MRS peaks are usually expressed in dimensionless units (parts per million, ppm), with reference to a

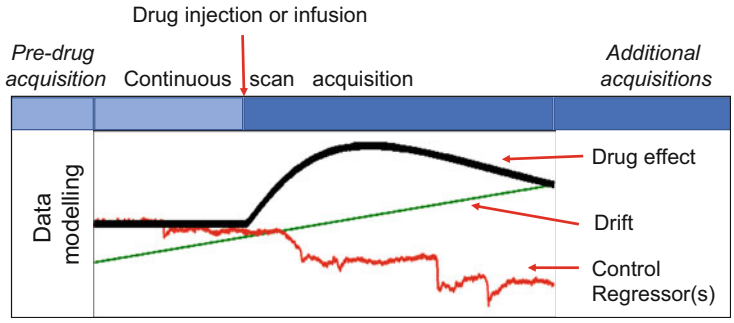


Fig. 2 An example of a phMRI experiment implementation. The blue bar represents the scan acquisition which may involve scans both before and after the phMRI scan. An example modelling approach is shown below for the phMRI scan only. Here regressors for the main drug effect (thick black line) and common confounds of drift (green line) and movement (represented by a single red line, can include multiple regressors) are shown (based on method used in De Simoni et al. 2013)

specific molecule. For proton MRS (1H-MRS), water is commonly used as a reference at 4.7 ppm. Proton MRS is of particular interest in neuroscience, because it can be used to measure the important central nervous system amino acids glutamate and glycine. The main inhibitory neurotransmitter GABA can also be assessed with more elaborate acquisition and analysis.

PhMRI

Pharmacological MRI or phMRI is a method used to describe the changes in the MRI timeseries following compound administration (Breiter et al. 1997). Any timeseries data acquisition method can be combined with compound administration in phMRI. In rodents, the use of a blood-pool contrast agents such as super-paramagnetic iron oxide particles can be used to collect cerebral blood volume timeseries data (Mandeville 2012). These use the same T2* imaging sequences and fast acquisition as BOLD imaging. PhMRI methods in humans have almost exclusively used BOLD contrast. Participants are normally at rest in the scanner when acquisition begins and after sufficient volumes have been acquired to define a stable baseline, the drug is administered without a break in the data acquisition. The expectation of phMRI is that following drug administration, a rapid change in MRI signal can be detected. Due to practical considerations with pharmacokinetics and difficulties with multiple administrations of a drug, this technique is limited to only one timeseries per participant per session.

The main challenge with this method is how to describe the effect of the drug on the MRI signal. One option is to compare the MRI signal before and after the drug administration (Breiter et al. 1997). However, this does not capitalize on temporal variation. Instead, an input function can be used, derived from the profile of the drug plasma levels. This will identify brain areas where the neuroimaging signal matches that of the plasma drug level over many minutes. The implicit assumption here is that the plasma and brain pharmacokinetics are matched and that the functional effects of the compound are temporally matched to the pharmacokinetics. Alternative input functions can be used to enhance the interpretation of the findings, such as the profile of subjective effects (Breiter et al. 1997). Mathematical models of the profile of predicted drug effect derived from pilot, or independent are useful when other input functions are not available or considered inappropriate (Fig. 2).

Critical Evaluation of MRI Approaches in Drug Development

There are a number of questions that NMR methods can be used to address when used in drug development.

Defining and Confirming the Marker for Assessment of a Drug Effect

These markers may be the level of particular metabolites in a specific brain region, a task-related activation, or connectivity change, or a localized difference in perfusion. For example, glutamate is elevated in patients with schizophrenia (Merritt et al. 2016; Schwerk et al. 2014) and MRS would be a viable methodology to test compounds for their ability to reduce glutamate levels in vivo. Hyperperfusion has been noted and replicated in patients at risk of psychosis (Allen et al. 2017) and is therefore a candidate target for interventions to reduce such risks. Amygdala reactivity to negative affective stimuli is elevated in patients with depression and a reduction this response has been used as a target to study antidepressants (Godlewska et al. 2012).

There are significant challenges here that speak to the entire effort of those aiming to describe the neural correlates of brain disorders. The first challenge is the definition of the patient groups (Insel and Cuthbert 2015). The diagnoses are often based on clinical presentation (and in some cases additional tests), which do not define the patients by their abnormal glutamate, perfusion or brain reactivity, etc. That is, using the first example above, there will be individuals with elevated glutamate without schizophrenia and patients with schizophrenia with apparently normal MRS-measured glutamate levels. Moreover, the existence of differences which are sensitive to specific subgroups may lead to dilution of effects unless these subgroups are taken into account. These subgroups may be relatively easy to define, such as first-episode vs. chronic patients, or treatment-responsive vs. treatment-resistant, or prove difficult to define, such as patients with similar symptoms, but different etiologies, or patient groups that separate based on the biological data collected, so-called biotypes (e.g., Drysdale et al. 2017). Such stratifications may or may not be relevant for later assessment of potential treatments and represent an area of significant research need.

Despite these challenges, there are examples of robust, evidence-based neuroimaging markers in use for testing drug effects. Many patients with schizophrenia show elevated presynaptic dopamine markers (Mccutcheon et al. 2017), altered brain connectivity during working memory tasks, and differential activity during reward processing; many patients diagnosed with major depression show hyperreactivity in the amygdala (Leppanen 2006) and altered activity in the anterior cingulate and frontal cortex that may be predictive of treatment response (Fonseka et al. 2018).

Important contributions from the literature with existing drugs will be to provide clear demonstrations that changes to these imaging markers leads to improvements in patient outcomes.

Provide Supportive Evidence for the *Functional* Mechanism of Action of a Drug on a Pharmacological or Disease-Relevant Mechanism

When the effects of a drug are tested on the fMRI signal, it can provide evidence for mechanisms of interest. For example, a drug with potential to enhance reward-related activity in the brain can be tested using a paradigm sensitive to reward processing. An example of this is the emerging evidence for the dopamine antagonist, lurasidone, to act as an antidepressant. A group of individuals were scanned during the performance of the monetary incentive delay (MID) which can be used to ascertain the brain regions involved in response to a reward predicting cue and later in response to actually receiving the reward. Such anticipatory and outcome events have been associated with overlapping brain regions. Increased activity of the key ventral striatum region responsive to reward anticipation was improved by lurasidone, but only in those with higher levels of anhedonia (Wolke 2018). Here the mechanism is captured by the brain regions responsive and the task phase.

The change in fMRI signal indicates support for the hypothesized reward mechanism being modulated by the drug. While this is a valuable goal in proof of concept experimental studies, understanding if the effect size of the fMRI signal change is clinically meaningful would provide a stronger platform for translating findings for patient benefit and potentially aid outcome

measure selection and effective dose selection. In other words, if a change in brain activity is observed, is that predicted to lead to clinical benefit? There are many examples of treatment studies in which fMRI changes are accompanied by symptomatic improvement. For example, in patients with psychosis and other psychiatric disorders such as depression studies show that activation, connectivity and multivariate patterns can normalize with treatment effects (Abbott et al. 2013; Hart et al. 2012). When studies have reported brain activation effect sizes, existing treatments do not fully reverse but instead reduce impairments (Snitz et al. 2005; Williams et al. 2015). For example, responders to antidepressants show a two-third reduction in amygdala activation abnormality compared to no change in nonresponders (Williams et al. 2015) and frontal cortex impairments reduce by a similar amount after successful antipsychotic treatment in patients with schizophrenia (Snitz et al. 2005). While these studies require replication, they are important in defining the goals of these proofs of concept studies such that they can be powered appropriately.

Proof of concept fMRI studies can include either healthy volunteers or patient groups. The decision about which group to focus on depends on a number of factors. Healthy volunteer studies can be conducted faster than patient studies, and the group is theoretically more homogeneous, without potential underlying variation in disorder related processes. In addition, healthy volunteers can be recruited for pharmacological modelling of a potential disease mechanism, against which novel treatments can be assessed.

Ketamine is an example of a model for the glutamatergic system contributions to schizophrenia (Krystal et al. 1994). When given as an infusion to healthy volunteers, it produces a rapid increase in rating scale scores of both positive and negative symptoms of schizophrenia and cognitive impairment (Javitt et al. 2018). Using the phMRI methodology (see above), ketamine produces a robust and reliable neuroimaging response (Deakin et al. 2008; De Simoni et al. 2013). This neuroimaging response can be considered as an assay for the acute disturbance of glutamate transmission and is thus suitable for testing the effect of compounds which may act to attenuate this disturbance, such as lamotrigine and risperidone (Doyle et al. 2013). This assay has been used to test the effectiveness of two group 2 metabotropic glutamatergic receptor agonists (mGlu2 agonist prodrug LY2979165 and the mGlu2/3 agonist prodrug LY2140023). Both compounds were able to attenuate the phMRI ketamine response (Mehta et al. 2018), providing evidence in vivo in humans that the drugs are able to reduce an acute glutamatergic disturbance (as modelled by ketamine). The use of multiple doses of each compound was able to inform on dose response relationships, with only the highest doses of each compound used showing statistically significant attenuation of the ketamine response.

Direct investigation in patients does not require pharmacological modelling of the disease or disorder and thus task-related fMRI, resting state fMRI or perfusion can be used to assess the effects of a drug relative to a placebo or other comparison condition. While the imaging methods have been addressed above, additional considerations of existing treatment, treatment variation and drug-drug interactions must also be addressed. The advantage of these investigations is that they provide direct evidence in the target group of modification of brain mechanisms relevant to the pathology. A series of studies in ADHD suggest that psychomotor stimulant medication normalizes the task-related fMRI impairment (Hart et al. 2012). This represents a candidate assay in ADHD to assess putative novel treatments. Positive results in these studies can provide key input into decisions about larger trials in patient populations and may even inform on assessments for inclusion and target dose ranges.

Drug Comparisons

Different compounds may have similar clinical efficacy, despite having variations in precise mechanisms of action. Antipsychotic drugs are a case in point as they cannot be easily separated in terms of clinical efficacy to reduce positive symptoms of psychosis. Neuroimaging can be used to demonstrate that the circuits modulated by these

compounds do not fully overlap (Handley et al. 2013). Another case is that of medication for ADHD, which includes the archetypal stimulant drug methylphenidate and the non-stimulant atomoxetine. Methylphenidate is an inhibitor of the dopamine and noradrenaline transporter, whereas atomoxetine is selective for the noradrenaline transporter. Arterial spin labelling and task-related fMRI have both been used to show clear discrimination of these medications in healthy volunteers where each participant receives each drug (Marquand et al. 2012). Perfusion mapping showed a preferential response in widespread areas including the caudate and putamen to methylphenidate and more cortical regions for atomoxetine. Task-related fMRI has also been used to derive candidate markers for superior response to each of these compounds in youth with ADHD, with caudate response suggested as a marker for superior response to methylphenidate and motor cortex as a marker for superior response to atomoxetine (Schulz et al. 2017).

The value of testing different compounds, that have proven clinical utility, is the potential to use their responses to define core signatures of efficacy and identify candidate markers for superiority of response to individual compounds. The inclusion of novel compounds in such designs during phase 1b/II can allow both early assessment against a core response signature and also assessment of discrimination from existing treatments *outside* of this core signature (e.g., in other brain networks). These functional "off-target" effects can then be tested for their potential predictive value in clinical efficacy or side effects. In summary, neuroimaging markers can be used to compare different compounds to demonstrate that they operate through difference functional mechanisms or show that despite difference pharmacological profiles there is overlap in the functional effects.

Confounds and Limitations

Functional MRI methods are indirect methods used to infer changes in neuronal function. They are reliant on a process known as neurovascular coupling. This involves neurons, the vasculature and glial cells, in particular astrocytes. Projections known as astrocytic feet modulate vasoconstriction and vasodilation, and these processes are mediated and moderated by a range of neurochemical cascades (Petzold and Murthy 2011). Receptors on the smooth muscle or the astrocytes can also modulate vascular responses. Assessment of the potential alterations in neurovascular coupling and the vascular response following drug administration would be included in optimized protocols. The confounding effects can manifest as (i) changes in baseline perfusion following the drug, which may alter the amplitude of activations during tasks (Cohen et al. 2002), and this can be assessed by including an additional scan with arterial spin labelling, (ii) changes in the vascular reactivity, which may also alter the amplitude of activations, and this can be assessed via non-neuronal methods for vasodilation, such as breathhold-induced hypercapnia (Kannurpatti et al. 2014), and (iii) changes in the variance of the fMRI contrast similarly representing changes in vascular reactivity, which can be correlated with activations or included as a confounding factor in the analyses (Kannurpatti et al. 2014). The increased use of these methodologies will lead to a growth in the number of examples where these additional control steps refine the results and can help define drug classes and neurotransmitter systems where assessing such confounds is imperative. The global signal across the brain can also be affected by physiological noise from normal breathing and cardiac motion, which can be assessed with bellows and pulse monitors. Fluctuations in breathing and heart rate may have direct effects but also alias to lower frequencies. While this can be dealt with in part by filtering the data to include frequencies between ~0.009 Hz and ~0.1 Hz, these effects can also be included in the modelling of the data by calculating a basis set from the measured respiration and heart rate monitors (Murphy et al. 2013).

Movement is a major confound in functional neuroimaging with MRI and must be accounted for in the data modelling. During data collection, movement is limited by ensuring participants are comfortable and securing the head with straps and foam. Care must be taken that this is balanced

against participant comfort otherwise there is a risk that the participants may not tolerate the entire scanning protocol resulting in missing or incomplete data. Prior to data collection, participants can be acclimatized to the neuroimaging environment by the use of a mock-up scanner. Although movement during scans cannot be eliminated, predefining standards for acceptable movement can ensure that artifactual results or reduced sensitivity due to movement is mitigated. Exploratory analyses can be added to the protocol to ensure that any drug effects are not explained by movement.

Multimodal Neuroimaging

FMRI can be combined successfully with other methods, most notably electroencephalography and positron emission tomography. While there are no clear examples of such combinations enhancing the use of MRI methodology in drug development, advances in these methods and the availability of new hardware create an opportunity. For example, EEG and fMRI have been combined in many studies, but for pharmacological studies, there are two main uses that that could provide benefit. The first is to contribute to understanding potential neurovascular artifacts in the fMRI data from the addition of a drug, by providing information on the neuronal activity (e.g., Arthurs et al. 2004), although considerable work is required on the underlying models that link the EEG and fMRI outcomes. The second is to use one method to enhance the analysis of the other. For example, using fMRI to narrow the solutions to the inverse problem in EEG and define potential sources of signal and EEG-informed fMRI analyses, where the EEG response to single conditions can be included in the fMRI model. An example of a single trial model is where the EEG derived P3 response to trials in an oddball paradigm were included in the fMRI analysis as a marker of attention to the task. By informing the fMRI model – in a trial-specific manner – from the EEG response the sensitivity to nicotine modulation of attentional responses in the brain was dramatically enhanced (Warbrick et al. 2012).

Other methods such as emission tomography or postmortem receptor or gene expression data can be combined with MRI-based outcomes to link the fundamentally hemodynamic responses to molecular targets (Dukart et al. 2018). Such methods are important in the development of our understanding of the molecular mechanisms behind the MRI changes with drugs, but as yet do not have a role in supporting MRI for drug development.

Conclusions and Future Directions

The multimodal nature of functional MRI allows for assessment of different aspects of pharmacological modulation in the brain, including changes in task-related activation, functional connectivity between brain regions and network topology, metabolite levels, drug-related activation as assessed using phMRI and quantitative perfusion. The key advantages and challenges of these methods are recapitulated below along with comments on future directions:

(i) MRS can be used to provide evidence of the neurochemical changes induced by drugs but are limited to the metabolite levels that MRS is sensitive to with poor temporal (minutes) and spatial (cm) resolution. Dynamic changes in metabolite levels using, so-called functional MRS (Jelen et al. 2018) will be a significant advantage allowing testing of the responsivity of the system and the role of treatment in potentially normalizing abnormalities. Newer methods that allow whole brain assessment rapidly with good signal to noise ratio will be important developments.

(ii) Task-related fMRI can be used to provide evidence of context-dependent modulation of brain function and thus identify areas of greater and smaller sensitivity and indicate the neural mechanisms underpinning the drug effects. The implementation of such methods requires expertise in the specific task for delivery during the scan and appropriate analysis of the data. Participants may

have to be dropped because of poor performance and those who are cognitively challenged, lack motivation or are otherwise unwell may not be able to engage with the tasks. Typically, a very small number of cognitive domains can be assessed because of the time required to collect sufficient data while maintaining participant engagement and comfort. Methods to reduce data acquisition time will be important to enhance the participant experience and allow the protocols to be more efficient or contain more measurements.

(iii) Resting state imaging has proved extremely versatile as a technique for assessing pharmacological effects of compounds, demonstrating relatively local changes in specific connections (i.e., specific pathway) and topological reorganization (i.e., widespread network effects). It has been used to validate assays for drug development (e.g., Joules et al. 2015 for ketamine). Its great advantage is that it provides task-independent descriptions of drug effects, but it lacks functional specificity and it is not clear if a drug effect observed during a resting scan translates into a functional effect in the same brain areas.

(iv) Scanning at rest is also the basis of phMRI which is used to measure the effect of a fast-acting drug on the fMRI signal. It is particularly well suited to study drugs which are injected and is a translatable technique as the same methods can be used in small animal imaging. In drug development, in humans, this method has been used to develop an assay of glutamatergic dysfunction against which novel drugs can be tested (Mehta et al. 2018). The translational value is an advantage of this method (Coimbra et al. 2013), which accompanies its growing use in human studies (Javitt et al. 2018). Being constrained to only fast-acting drugs does however limit the number of drugs that can safely be used in this assay.

(v) Perfusion imaging can be used to track the effects of injected compounds with the phMRI methodology and also to quantify regional cerebral blood flow in order to compare drug effects across sessions. The signal to noise ratio is lower than that of BOLD fMRI, but the sensitivity and reliability is very good, with effects of many different drug classes demonstrated to date. There are two main uses of perfusion imaging – to assess the baseline state of the system, with particular value in understanding the vascular contributions to task-based and resting state fMRI changes, and to quantify the effects of compounds directly.

Overall, these methods can be used as described above, by themselves, or in combinations within the same imaging sessions to answer questions on neuropharmacological mechanisms of psychopathology, development of assays and describing pharmacodynamic effects, and pharmacokinetic–pharmacodynamic relationships of compounds. When testing novel compounds, these methods are fundamentally pharmacodynamic in nature and can be applied in healthy volunteers or patient groups. They complement emission tomography data which can be used to confirm brain penetration and dose-occupancy relationships to specific targets and can extend or even stand in for these methods when ligands are not available or multiple targets are involved in the desired effects providing intermediate markers of functional outcome.

References and Further Reading

Abbott CC, Jaramillo A, Wilcox CE, Hamilton DA (2013) Antipsychotic drug effects in schizophrenia: a review of longitudinal fMRI investigations and neural interpretations. Curr Med Chem 20:428–437

Allen P, Azis M, Modinos G, Bossong MG, Bonoldi I, Samson C, Quinn B, Kempton MJ, Howes OD, Stone JM, Calem M, Perez J, Bhattacharayya S, Broome MR, Grace AA, Zelaya F, Mcguire P (2017) Increased resting hippocampal and basal ganglia perfusion in people at ultra high risk for psychosis: replication in a second cohort. Schizophr Bull. https://doi.org/10.1093/schbul/sbx169

Alsop DC, Detre JA, Golay X, Gunther M, Hendrikse J, Hernandez-Garcia L, Lu H, Macintosh BJ, Parkes LM, Smits M, Van Osch MJ, Wang DJ, Wong EC,

Zaharchuk G (2015) Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: a consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. Magn Reson Med 73:102–116

Arthurs OJ, Stephenson CM, Rice K, Lupson VC, Spiegelhalter DJ, Boniface SJ, Bullmore ET (2004) Dopaminergic effects on electrophysiological and functional MRI measures of human cortical stimulus-response power laws. Neuroimage 21:540–546

Breiter HC, Gollub RL, Weisskoff RM, Kennedy DN, Makris N, Berke JD, Goodman JM, Kantor HL, Gastfriend DR, Riorden JP, Mathew RT, Rosen BR, Hyman SE (1997) Acute effects of cocaine on human brain activity and emotion. Neuron 19:591–611

Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186–198

Buonocore MH, Maddock RJ (2015) Magnetic resonance spectroscopy of the brain: a review of physical principles and technical methods. Rev Neurosci 26:609–632

Cohen ER, Ugurbil K, Kim SG (2002) Effect of basal conditions on the magnitude and dynamics of the blood oxygenation level-dependent fMRI response. J Cereb Blood Flow Metab 22:1042–1053

Coimbra A, Baumgartner R, Schwarz AJ (2013) Pharmacological FMRI in drug discovery and development. In: Garrido L, Beckmann L (eds) New applications of NMR in drug discovery and development Royal Society Of Chemstry

Dai W, Garcia D, De Bazelaire C, Alsop DC (2008) Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields. Magn Reson Med 60:1488–1497

De Simoni S, Schwarz AJ, O'daly OG, Marquand AF, Brittain C, Gonzales C, Stephenson S, Williams SC, Mehta MA (2013) Test-retest reliability of the bold pharmacological MRI response to ketamine in healthy volunteers. Neuroimage 64:75–90

Deakin JF, Lees J, Mckie S, Hallak JE, Williams SR, Dursun SM (2008) Glutamate and the neural basis of the subjective effects of ketamine: a pharmaco-magnetic resonance imaging study. Arch Gen Psychiatry 65:154–164

Doyle OM, De Simoni S, Schwarz AJ, Brittain C, O'daly OG, Williams SC, Mehta MA (2013) Quantifying the attenuation of the ketamine pharmacological magnetic resonance imaging response in humans: a validation using antipsychotic and glutamatergic agents. J Pharmacol Exp Ther 345:151–160

Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, Fetcho RN, Zebley B, Oathes DJ, Etkin A, Schatzberg AF, Sudheimer K, Keller J, Mayberg HS, Gunning FM, Alexopoulos GS, Fox MD, Pascual-Leone A, Voss HU, Casey BJ, Dubin MJ, Liston C (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 23:28–38

Duff EP, Vennart W, Wise RG, Howard MA, Harris RE, Lee M, Wartolowska K, Wanigasekera V, Wilson FJ, Whitlock M, Tracey I, Woolrich MW, Smith SM (2015) Learning to identify CNS drug action and efficacy using multistudy FMRI data. Sci Transl Med 7:274ra16

Dukart J, Holiga S, Chatham C, Hawkins P, Forsyth A, Mcmillan R, Myers J, Lingford-Hughes AR, Nutt DJ, Merlo-Pich E, Risterucci C, Boak L, Umbricht D, Schobel S, Liu T, Mehta MA, Zelaya FO, Williams SC, Brown G, Paulus M, Honey GD, Muthukumaraswamy S, Hipp J, Bertolino A, Sambataro F (2018) Cerebral blood flow predicts differential neurotransmitter activity. Sci Rep 8:4074

Fonseka TM, Macqueen GM, Kennedy SH (2018) Neuroimaging biomarkers as predictors of treatment outcome in major depressive disorder. J Affect Disord 233:21–35

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229

Godlewska BR, Norbury R, Selvaraj S, Cowen PJ, Harmer CJ (2012) Short-term SSRI treatment normalises amygdala hyperactivity in depressed patients. Psychol Med 42:2609–2617

Handley R, Zelaya FO, Reinders AA, Marques TR, Mehta MA, O'gorman R, Alsop DC, Taylor H, Johnston A, Williams S, Mcguire P, Pariante CM, Kapur S, Dazzan P (2013) Acute effects of single-dose aripiprazole and haloperidol on resting cerebral blood flow (RCBF) in the human brain. Hum Brain Mapp 34:272–282

Hart H, Radua J, Nakao T, Mataix-Cols D, Rubia K (2012) Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. JAMA Psychiatry 70:185–98

Huettel SA, Song AW, Mccarthy G (2014) Functional magnetic resonance imaging, 3rd edn. Sunderland, Sinauer Associates

Insel TR, Cuthbert BN (2015) Medicine. Brain disorders? Precisely. Science 348:499–500

Javitt DC, Carter CS, Krystal JH, Kantrowitz JT, Girgis RR, Kegeles LS, Ragland JD, Maddock RJ, Lesh TA, Tanase C, Corlett PR, Rothman DL, Mason G, Qiu M, Robinson J, Potter WZ, Carlson M, Wall MM, Choo TH, Grinband J, Lieberman JA (2018) Utility of imaging-based biomarkers for glutamate-targeted drug development in psychotic disorders: a randomized clinical trial. JAMA Psychiatry 75:11–19

Jelen LA, King S, Mullins PG, Stone JM (2018) Beyond static measures: a review of functional magnetic resonance spectroscopy and its potential to investigate dynamic glutamatergic abnormalities in schizophrenia. J Psychopharmacol 32:497–508

Jonckers E, Shah D, Hamaide J, Verhoye M, Van Der Linden A (2015) The power of using functional FMRI on small rodents to study brain pharmacology and disease. Front Pharmacol 6:231

Joules R, Doyle OM, Schwarz AJ, O'daly OG, Brammer M, Williams SC, Mehta MA (2015) Ketamine induces a robust whole-brain connectivity pattern that can be differentially modulated by drugs of different mechanism and clinical profile. Psychopharmacology 232:4205–4218

Kannurpatti SS, Motes MA, Biswal BB, Rypma B (2014) Assessment of unconstrained cerebrovascular reactivity marker for large age-range FMRI studies. PLoS One 9:E88751

Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, Heninger GR, Bowers MB Jr, Charney DS (1994) Subanesthetic effects of the noncompetitive nmda antagonist, ketamine, in humans. psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch Gen Psychiatry 51:199–214

Leppanen JM (2006) Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. Curr Opin Psychiatry 19:34–39

Li Z, Vidorreta M, Katchmar N, Alsop DC, Wolf DH, Detre JA (2018) Effects of resting state condition on reliability, trait specificity, and network connectivity of brain function measured with arterial spin labeled perfusion MRI. Neuroimage 173:165–175

Mandeville JB (2012) Iron FMRI measurements of CBV and implications for BOLD signal. Neuroimage 62:1000–1008

Marquand AF, O'daly OG, De Simoni S, Alsop DC, Maguire RP, Williams SC, Zelaya FO, Mehta MA (2012) Dissociable effects of methylphenidate, atomoxetine and placebo on regional cerebral blood flow in healthy volunteers at rest: a multi-class pattern recognition approach. Neuroimage 60:1015–1024

Mathias EJ, Kenny A, Plank MJ, David T (2018) Integrated models of neurovascular coupling and BOLD signals: responses for varying neural activations. Neuroimage 174:69–86

Mccutcheon R, Beck K, Jauhar S, Howes OD (2017) Defining the locus of dopaminergic dysfunction in schizophrenia: a meta-analysis and test of the mesolimbic hypothesis. Schizophr Bull

Mehta MA, Schmechtig A, Kotoula V, Mccolm J, Jackson K, Brittain C, Tauscher-Wisniewski S, Kinon BJ, Morrison PD, Pollak T, Mant T, Williams SCR, Schwarz AJ (2018) Group II metabotropic glutamate receptor agonist prodrugs Ly2979165 and Ly2140023 attenuate the functional imaging response to ketamine in healthy subjects. Psychopharmacology 235:1875–1886

Merritt K, Egerton A, Kempton MJ, Taylor MJ, Mcguire PK (2016) Nature of glutamate alterations in schizophrenia: a meta-analysis of proton magnetic resonance spectroscopy studies. JAMA Psychiatry 73:665–674

Murphy K, Birn RM, Bandettini PA (2013) Resting-state FMRI confounds and cleanup. Neuroimage 80:349–359

Paloyelis Y, Doyle OM, Zelaya FO, Maltezos S, Williams SC, Fotopoulou A, Howard MA (2016) A spatiotemporal profile of in vivo cerebral blood flow changes following intranasal oxytocin in humans. Biol Psychiatry 79:693–705

Petzold GC, Murthy VN (2011) Role of astrocytes in neurovascular coupling. Neuron 71:782–797

Pruim RHR, Mennes M, Van Rooij D, Llera A, Buitelaar JK, Beckmann CF (2015) ICA-aroma: a robust ICA-based strategy for removing motion artifacts from FMRI data. Neuroimage 112:267–277

Roalf DR, Nanga RPR, Rupert PE, Hariharan H, Quarmley M, Calkins ME, Dress E, Prabhakaran K, Elliott MA, Moberg PJ, Gur RC, Gur RE, Reddy R, Turetsky BI (2017) Glutamate imaging (glucest) reveals lower brain glucest contrast in patients on the psychosis spectrum. Mol Psychiatry 22:1298–1305

Schulz KP, Bedard AV, Fan J, Hildebrandt TB, Stein MA, Ivanov I, Halperin JM, Newcorn JH (2017) Striatal activation predicts differential therapeutic responses to methylphenidate and atomoxetine. J Am Acad Child Adolesc Psychiatry 56:602–609. E2

Schwarz AJ, Becerra L, Upadhyay J, Anderson J, Baumgartner R, Coimbra A, Evelhoch J, Hargreaves R, Robertson B, Iyengar S, Tauscher J, Bleakman D, Borsook D (2011) A procedural framework for good imaging practice in pharmacological FMRI studies applied to drug development #1: processes and requirements. Drug Discov Today 16:583–593

Schwerk A, Alves FD, Pouwels PJ, Van Amelsvoort T (2014) Metabolic alterations associated with schizophrenia: a critical evaluation of proton magnetic resonance spectroscopy studies. J Neurochem 128:1–87

Smith SM, Vidaurre D, Beckmann CF, Glasser MF, Jenkinson M, Miller KL, Nichols TE, Robinson EC, Salimi-Khorshidi G, Woolrich MW, Barch DM, Ugurbil K, Van Essen DC (2013) Functional connectomics from resting-state FMRI. Trends Cogn Sci 17:666–682

Snitz BE, Macdonald A 3rd, Cohen JD, Cho RY, Becker T, Carter CS (2005) Lateral and medial hypofrontality in first-episode schizophrenia: functional activity in a medication-naive state and effects of short-term atypical antipsychotic treatment. Am J Psychiatry 162:2322–2329

Warbrick T, Mobascher A, Brinkmeyer J, Musso F, Stoecker T, Shah NJ, Fink GR, Winterer G (2012) Nicotine effects on brain function during a visual oddball task: a comparison between conventional and EEG-Informed FMRI analysis. J Cogn Neurosci 24:1682–1694

Williams LM, Korgaonkar MS, Song YC, Paton R, Eagles S, Goldstein-Piekarski A, Grieve SM, Harris AW, Usherwood T, Etkin A (2015) Amygdala reactivity to emotional faces in the prediction of general and medication-specific responses to antidepressant treatment in the randomized iSPOT-D trial. Neuropsychopharmacology 40:2398–2408

Wolke S (2018) Mechanisms of reward in depression: an intervention study investigating the acute effects of lurasidone on cerebral blood flow and the neural

correlates of reward and penalty processing. Phd, King's College London

Wong EC, Buxton RB, Frank LR (1997) Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. NMR Biomed 10:237–249

Zelaya FO, Fernández-Seara MA, Black KJ, Williams SCR, Mehta MA (2015) Perfusion in pharmacologic imaging. In: Bammer R (ed) MR & CT perfusion imaging: clinical applications and theoretical principles. Lippincott Williams & Wilkins, Philadelphia

Traditional Chinese Medicine and Clinical Pharmacology

23

Anita Chen Marshall

Contents

A. C. Marshall (✉)
Sequoia Healing Center for Acupuncture and Herbal Medicine, Alameda, CA, USA

American College of Traditional Chinese Medicine (ACTCM), California Institute of Integral Studies (CIIS), San Francisco, CA, USA

Bastyr University, Seattle, WA, USA

John A. Burns School of Medicine, University of Hawaii, Honolulu, HI, USA
e-mail: dranitamarshall@gmail.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_60

Abstract

Traditional Chinese Medicine (TCM) represents thousands of years in treatment protocols for health, healing, and longevity and has evolved into a complex healthcare system. The following chapter is designed to give a basic concise overview of TCM principles and practices, in relation to clinical pharmacology, and focuses on mutually relevant areas of interest. Much as society learned to harness electrical currents to empower machinery, TCM theories of energetic patterns help to empower the functionality of the human body. The major theories include the vital life force, known as Qi, Yin and Yang, Five Elements, herbal medicine, multivariate diagnostics, syndrome differentiation, and acupuncture, among others. In TCM, there is a direct relationship to the energy of nature and the energetic patterns within the human body and the human consciousness. This comprehensive approach is gaining momentum in clinical and network pharmacology, as new technologies emerge regarding TCM collaborative compounds versus singular chemical constituents in treatment modalities. Of major importance to clinical pharmacology are recent TCM evidenced-based empirical studies, correlations, and interactions between herbs and pharmaceuticals, updated data on toxicity, adverse reactions, quality assurance, and herbal medicine standardization. Global concerns over antimicrobial resistance to drugs have put a spotlight on TCM herbal medicines as alternatives with greatly reduced resistance factors. From ancient archives to the Nobel Prize to a formal endorsement by the World Health Organization as a global healthcare system, TCM has widened the lens of modern science. It offers the wisdom and experience of millennia to inspire the think tank of today's medicine and pharmacology and tomorrow's generations of integrative healthcare advancements.

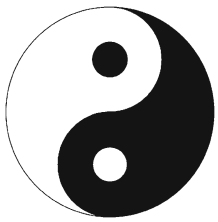
Introduction

Traditional Chinese Medicine (TCM) is renowned and respected among the world's time-honored systems of wellness and healing. Although legend and mythology allude to TCM origins as far back as 5000 years BCE, it was during China's Shang dynasty (1766–1122 BCE) TCM was systematized among the highly esteemed traditional herbalists and healing practitioners of the imperial courts. The first authenticated text to reference the concept of Yin and Yang (阴阳), which is the most important and distinctive theory of TCM, was the "Book of Changes" (Yi Jing 易經), dating back to around 700 BCE. Together with the concept of Qi (氣), vital life force, also translated as Chi (pronounced as "chee"), they provide the basic foundation for TCM's comprehensive mind, body, and spirit ideology and practice.

On May 25, 2019, TCM was officially adopted into the 11th version of the International Statistical Classification of Diseases and Related Health Problems (ICD). This document represents the formal global classification of diseases by the

World Health Assembly, the governing body of the World Health Organization (WHO). This official compendium has recognized TCM as a traditional healthcare system and endorsed it to its 129 global member states. Excerpted from the WHO website regarding the ICD-11: "ICD is the foundation for the identification of health trends and statistics globally, and the international standard for reporting diseases and health conditions. ICD-11 was submitted to the Seventy-second World Health Assembly in May, 2019 and, following endorsement, Member States will start reporting using ICD-11on1 January 2022."

This chapter has been carefully written to inform and inspire readers to evaluate better the benefits of TCM in relation to conventional medicine and clinical pharmacology. TCM has a rich lineage of traditional knowledge bases and skill sets and has survived millennia of medical and scientific evolution to be acknowledged as a major contributor to twenty-first-century global healthcare. The compelling wisdom that TCM offers modern medicine is the primary position of prevention above treatment and discovering each individual's unique blueprint of wellness.



The law of yin and yang is
the natural order of the universe,
the foundation of all things,
the mother of all changes,
the root of life and death…

The Yellow Emperor's Inner Classic
(Huang Di Nei Jing -黃帝內經)

Principles and Practices of Traditional Chinese Medicine (TCM)

One of the most important theoretical concepts of TCM is Yin (阴) and Yang (阳) that were recorded in the "Book of Change" (Yi Jing 易經), dating from 700 BCE. The iconic Yin/Yang symbol represents opposite properties and functions but has complementary qualities (see Table 1). Yin and Yang are mutually dependent and supportive of each other; one cannot survive without the other. Extreme Yin will transform itself into Yang, and extreme Yang will transform itself into Yin. There is Yin (small dark circle) within Yang (large white area above) and Yang (small white circle) within Yin (large dark area below), as illustrated in the Yin/Yang symbol.

All treatment goals are centralized around balancing the Yin and Yang through harmonizing the body, mind, and spirit. Internally, this is accomplished through a combination of the healing properties of food, herbs, acupuncture, moxibustion, Tui Na massage, cupping, meditation, and Taichi and Qigong exercises. In addition, there is correcting of harmful external environmental factors: expelling toxins, avoiding extreme temperatures, and eliminating physiological and emotional stressors. In TCM, a state of wellness and well-being can only be achieved when there is a state of equilibrium. What may be viewed in metaphysics as the harmonic convergence is expressed in conventional medicine as homeostasis and in TCM as ultimate balance.

TCM has been a champion of energetic medicine for millennia, because its underlying principles and practices stem from a foundational philosophy of the pervasive energetic force that permeates all of nature. In TCM, Qi is the key which opens the meridian channels or energy pathways and provides movement throughout the interconnected domains of health and healing. In the *Yellow Emperor's Inner Classic* (*Huang Di Nei Jing,* 黄帝内经), the most revered ancient text of TCM, there is a description of the principles of Yin and Yang, the blood vessels, and meridians, also known as channels (Jing Mai 经脈). Qi flows within the body, as well as the five Zang organs (五脏), the small intestine, gall bladder, large

Table 1 Courtesy of Anita Chen Marshall, DAOM, Pharm D, LAc

Correspondences of Yin and Yang									
Yang	Sun	Sky	Brightness	Male	Activity	Fire	Rapid	Posterior	Strength
Yin	Moon	Earth	Shade	Female	Rest	Water	Slow	Anterior	Weakness

intestine, stomach, and urinary bladder; six Fu organs (六腑), the heart, liver, lungs, spleen, kidneys, and pericardium; and associated acupuncture points (Xue Wei 穴位).

In addition to the concept of Yin/Yang in TCM, this text also introduces the Five Elements Theory (五行). This theory encompasses all natural phenomena, as they symbolically relate to human nature. As expressed in the *Yellow Emperor's Inner Classic*, the Five Element activities of Wood (liver, gallbladder), Fire (heart, small intestine), Earth (spleen, stomach), Metal (lung, large intestine), and Water (kidney, urinary bladder) are arranged in relationship sequences in which each element interacts with one another by generating, controlling, overacting, and insulting.

- **Generating Sequence (normal)**: Each element generates another and is generated by another. Wood generates Fire, Fire generates Earth, Earth generates Metal, Metal generates Water, and Water generates Wood. Often, TCM elaborates on this reality of one element generating another, as evidenced in a mother-child relationship. For example, Wood is the child of Water, while Wood is the mother of Fire.
- **Controlling Sequence (normal)**: Each element controls another and is controlled by another. For example, Wood controls Earth, Earth controls Water, Water controls Fire, Fire controls Metal, and Metal controls Wood. Another controlling example may be found in the following: Wood controls Earth, but Earth generates Metal, which in turn controls Wood. Therefore, this is a self-regulating sequence that keeps the body in balance. Excess controlling is perceived as "overacting" and can weaken other elements.
- **Insulting Sequence (abnormal)**: It is the reverse order of the Controlling Sequence. Wood insults Metal, Metal insults Fire, Fire insults Water, Water insults Earth, and Earth insults Wood. This causes the body to go out of balance. When the element becomes excessive, it will insult and weaken the corresponding element and, over time, it will result in imbalances and pathological symptoms (Fig. 1).

Applications of Five Elements Theory

The applications of the Five Elements Theory in TCM are numerous, and the systems of correspondences are linked to many explicit phenomena and qualities in nature and the human body. These correspondences are inclusive of organs, tissues, tastes, directions, climates, colors, seasons, etc. TCM also recognizes the healing power of musical tones, which correspond to each of the five elements. These tones are Jue (角Wood), Zhi (徵Fire), Gong (宫Earth), Shang (商Metal), and Yu (羽Water). Just as music is composed of harmonic and resonant layers, these unique tones are intrinsic to the interrelated vibratory balance of the human body, mind, and spirit, which TCM understands are needed to be in harmony with each other to achieve a sense of balance and health. Considering the importance of these tonal qualities, TCM recognizes music therapy as a vital healing modality. Each food and herb in TCM has a certain taste which is related to one of the elements. The five tastes are sour for Wood, bitter for Fire, sweet for Earth, pungent for Metal, and salty for Water. Each of the tastes has a unique effect on the body. For example, the sour taste generates fluids and Yin; it has an astringent quality and, therefore, can control perspiration and diarrhea. As another example, bitter taste clears heat, sedates, and hardens. Table 2 outlines these integrated corresponding aspects.

TCM Diagnostic Methods

During the Song dynasty (960–1279 CE), TCM was fully systematized to include comprehensive diagnosis, disease symptom profile analysis, and multivariate treatment protocols. Together with the theory of Yin and Yang, the Five Elements Theory forms the essential philosophy and methodology of TCM. As mentioned previously, TCM practitioners use a combination of the theories of Yin/Yang and the Five Elements to identify the cause of each disease, excess, deficiency, or debility and develop a customized treatment protocol. In the course of this

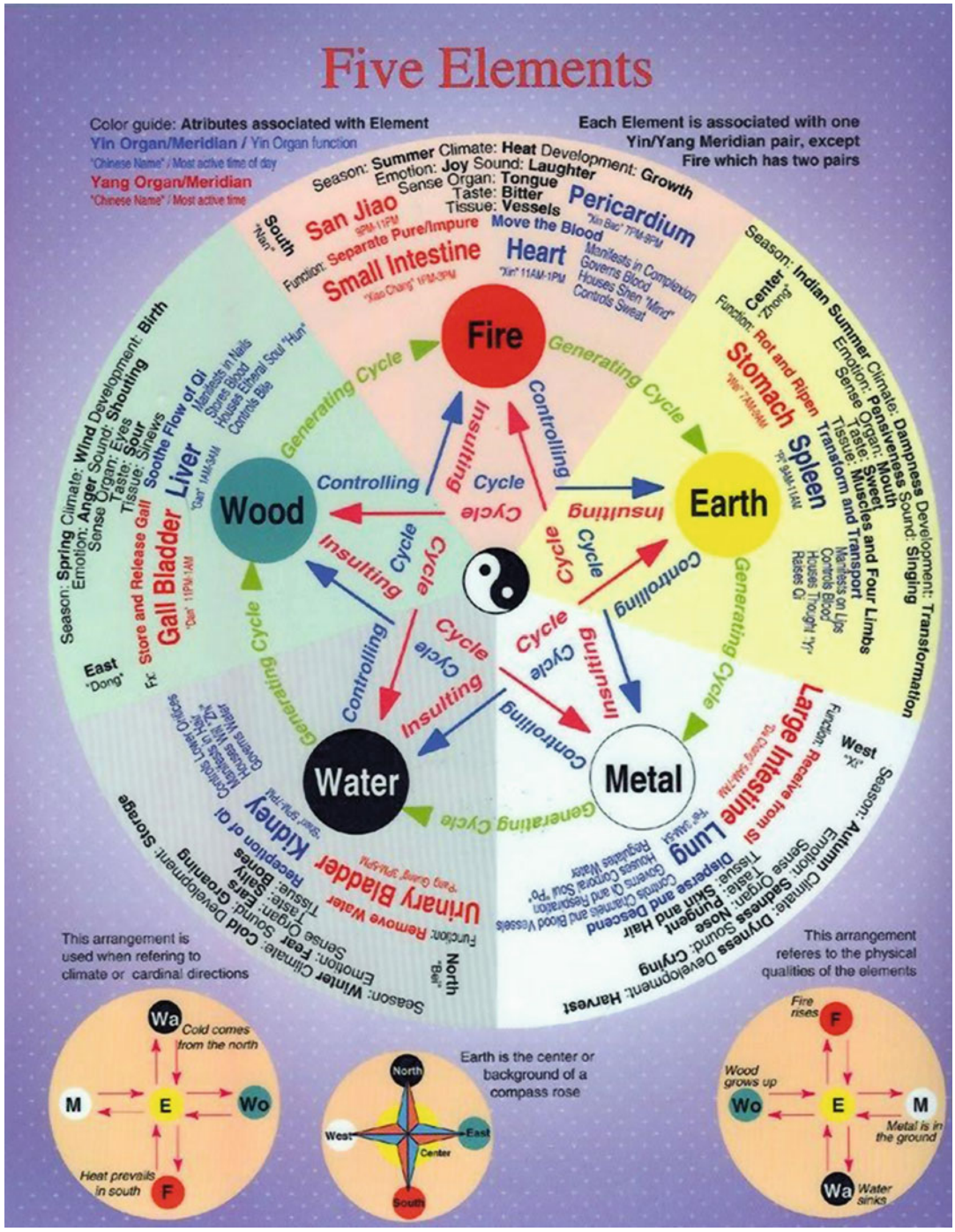


Fig. 1 Five elements chart. (Courtesy of Zdenek Zumr, MTCM/*Introduction to Oriental Medicine through Charts and Graphs*, 1994)

cohesive cycle of diagnosis from root causes to customized remedies, appropriate treatments are prescribed.

TCM uses the body's microsystems to treat the whole person effectively, mirroring a holographic perspective of each part reflecting the whole. This

Table 2 Corresponding aspects of five elements theory. (Courtesy of Anita Chen Marshall, DAOM, Pharm.D., LAc)

	Wood (木)	Fire (火)	Earth (土)	Metal (金)	Water (水)
Yin organs	Liver	Heart	Spleen	Lung	Kidney
Yang organs	Gallbladder	Small intestine	Stomach	Large intestine	Urinary bladder
Sense organs	Eyes	Tongue	Mouth	Nose	Ears
Tissues	Sinews	Vessels	Muscles	Skin	Bones
Tastes	Sour	Bitter	Sweet	Pungent	Salty
Colors	Green	Red	Yellow	White	Black
Emotions	Anger	Joy	Pensiveness	Sadness	Fear
Seasons	Spring	Summer	None	Autumn	Winter
Directions	East	South	Center	West	North
Colors	Green	Red	Yellow	White	Black
Sounds	Shouting	Laughing	Singing	Crying	Groaning
Odors	Rancid	Scorched	Fragrant	Rotten	Putrid
Spirit	Ethereal soul	Mind	Thought	Corporal soul	Will power
Climates	Wind	Heat	Dampness	Dryness	Cold
Development	Birth	Growth	Transformation	Harvest	Storage

can be seen in various parts of the body reflected in multiple areas within the face, eyes, ears, hands, and feet. On a therapeutic level, TCM addresses these through acupuncture point manipulation. As an example: auricular acupuncture, hand and foot reflexology, etc.

While western medicine utilizes a stethoscope for detecting auscultations, TCM takes note of various sounds and pitches of the voice, coughing, breathing, wheezing, hiccupping, groaning, etc. In general, a loud sound reflects a full/excess pattern, while a weak sound is indicative of an empty/deficient pattern. TCM also takes into account the various odors emanating from the body. These odors can be related to different organ imbalances within the five elements system of correspondences, for example, rancid for the liver, scorched for the heart, sweet for the spleen, rotten for the lungs, and putrid for the kidneys. Halitosis indicates heat in the stomach and so on.

Examination and consultation in TCM are similar to taking a western medical history and complete physical. This process is extensive and can help to find the causative factors of the illness by determining the timing, environment, situations with living conditions, emotional and social issues, environmental exposures, etc. The most commonly asked questions are regarding symptom profiles, such as fever, chills, sweating, painful areas, urination and bowel movement, food and taste, thirst and drink, sleep patterns, and condition of eyes, ears, throat, and other body parts. There are specific questions for women regarding menstruation, discharges, pregnancy, and childbirth, as well as, for men, regarding virility and impotence, loss of strength, hair loss, hearing loss, and other targeted male issues.

Pulses: Of particular importance to TCM, which distinguishes it from allopathy, is the attention to various pulses, unique to major organs and meridians. Chinese pulse diagnosis is an effective tool to extract detailed information on the state of the internal organs and the corresponding meridians by palpating the radial artery on both wrists, dividing them into three areas and accessing them at three different levels: superficial, middle, and deep.

As important as the rate of pulses are the qualities of pulses, which can be described as fast, slow, strong, weak, thin, wiry, tight, choppy, slippery, floating, deep, etc. These pulse indicators give the TCM practitioner a clear view of underlying conditions. For example, a strong wiry pulse may indicate liver Qi stagnation and excess of stress; a slippery pulse indicates dampness within the meridians, and a floating pulse indicates allergy condition or cold symptoms.

Tongue: Another specialized TCM diagnostic area is the microsystem of the tongue. The

practitioner takes careful note of the different shapes, colors, patterns, and coatings of the tongue, each giving specific details regarding the patient's condition. As treatments progress, multiple changes in the appearance of the tongue take place, indicating the efficacy of the chosen protocol.

A central principle of TCM diagnostic protocol is known as *syndrome differentiation*. This concept summarizes the interconnections of sensory diagnostic methods of observation, auditory, olfactory, consultative, pulse-taking, and palpation, in conjunction with pathological signs of the body, including the unique stages of disease development. Combined, these diagnostic methods are useful in identifying disease etiology, location, nature, and the relationship between the disease factors and a functional, balanced Qi. After evaluating these properties through diagnostic methods, the combined information leads to a syndrome differentiation analysis that aids in determining the most effective therapeutic protocol to eliminate pathogenic profiles.

TCM Treatments

TCM is a holistic and complex treatment system containing several treatment modalities that are utilized to promote healing and wellness. These modalities are acupuncture, herbal medicine, moxibustion, acupressure, Tui Na massage, cupping, nutrition and lifestyle consultation, meditation, and Taichi and Qigong exercises to strengthen the body and promote Qi and blood circulation (Fig. 2).

Moxibustion involves the heating of acupuncture points with smoldering mugwort herb (known as moxa). Moxibustion stimulates circulation, counteracts cold and dampness in the body, and promotes the smooth flow of blood and Qi. This safe, noninvasive technique may be used alone, but it is generally used in conjunction with acupuncture treatment.

Acupressure uses the same acupuncture points with gentle but firm pressure of the fingers and hands on the surface of the skin to increase the Qi and blood circulation, stimulating the body's natural curative abilities. Acupressure is the most effective method for treatment of tension-related ailments by using the power and sensitivity of the human hand, relaxing the muscles and relieving tension.

Tui Na massage means "pushing grasping" and is a powerful form of Chinese medical bodywork. Based on the same Oriental medical principles as acupuncture, Tui Na seeks to improve the flow of Qi through the meridian channels and is particularly effective for conditions involving muscles, tendons, and joints, such as structural misalignment, orthopedic problems, and sports injuries. It can also be used to treat internal diseases.

Cupping is an ancient technique, in which a special cup is applied to the skin and held in place by vacuum-like suction. The suction draws superficial tissue into the cup, which may be left in place or moved along the body. Cupping brings fresh blood to the area and helps improve circulation. Traditional cupping, sometimes referred to as *fire cupping*, uses heat to create a suction inside of glass cups. In modern times, plastic cups are used with a small vacuum pump to create suction, and small rubber cups that can be hand manipulated, without heat, have been introduced for convenience and safety.

Dietary therapy and nutritional consultation refers to selection of foods to help enhance patient's overall health, based on their patterns of disharmony. According to TCM, everyone has a different body constitution, temperament, and health conditions, which change at different ages. Diet and nutrition can lead the body back to balance by avoiding foods that aggravate imbalances. This concept correlates well in today's modern medicine where a large percentage of cancers and other degenerative diseases are known to be caused by improper diet and nutritional deficiencies.

Implicit in any diet is consideration of the agronomic factors in growing and cultivating plants with toxic pesticides, herbicides, and fungicides, as well as commercial farming of fish, poultry, and cattle with GMO feed, added hormones, and antibiotics, adding to toxic overload of the environment. These were nonexistent

Fig. 2 TCM treatment modalities. (Courtesy of Anita Chen Marshall, DAOM, Pharm D, LAc)

thousands of years ago when TCM was first developed but, in a modern polluted global environment, represent a very real challenge to contemporary human, animal, and environmental health.

A definitive landmark study, published in *The Lancet* in April 2019, involved a compilation of statistics (1990–2017) on morbidity and mortality related to dietary factors on a global scale. The findings indicate that 11 million deaths in 2017 were attributable to dietary risk factors showing suboptimal diet is responsible for more deaths than any other risks globally, including tobacco smoking. TCM acknowledged the vital importance of diet and nutrition thousands of years ago and continues to rank food therapy as the first line of treatment (Afshin et al. 2019).

Qigong means "life force energy cultivation" by practicing deep breathing techniques along with certain forms of movement, some of which mimic animals in nature. Medical Qigong is used as a treatment modality to promote, preserve, and restore health. The full benefits of Qigong exercises can be achieved when using correct posture and diaphragmatic breathing with conscious mental direction in a daily regimen. Qigong is often used to help reduce blood pressure, stress, and anxiety and promote restful sleep and well-being, in addition to eliminating common aging degeneration and promoting longevity.

Taichi, a form of Qigong, is a slow movement system of set forms of exercise based on Qigong principles of breathing, meditation, strength, and balance. Hua Tuo (华佗) is a reputed physician during 140–208 CE, who gained fame for his profound abilities in acupuncture, moxibustion, and medical *Da Yian* (大雁) exercises. It was his keen observation of the animal kingdom that lead him to develop the *Wuqinxi* (*Exercise of the Five Animals*, 五禽戲) from studying movements of the tiger, deer, bear, monkey, and crane.

Acupuncture

In TCM, there are 12 major meridians with more than 700 acupuncture points distributed bilaterally throughout the entire body and 2 central

meridians. The anterior meridian is Ren Mai (Conception Vessel 任脈), and the posterior meridian is Du Mai (Governing Vessel 督脈).

Acupuncture is a treatment modality performed by inserting very fine sterile needles, not much thicker than a human hair, in specific locations on the skin, known as acupuncture points. The depth of needle insertion depends on the disease being treated and the location of the acupuncture point. For example, a shallow insertion is generally used for treating disease at the superficial level, i.e., initial stage of common colds and skin and muscle conditions. By contrast, deeper needle insertion is generally required for treating internal disease, e.g., organs and bones.

Acupuncture stimulation regulates the flow of Qi within the energy pathways, which is referred to as meridian or channel that interconnects throughout the entire body. The practice of acupuncture affects the endocrine and central nervous system, including responses from both the sympathetic nervous system (Yang) and the parasympathetic nervous system (Yin) to achieve equilibrium and promote the body's innate healing ability (Fig. 3).

According to ancient Chinese texts, each organ has a two-hour ultimate activity period per day. When the energy of a meridian is not flowing well, it will exert certain symptoms, and the symptoms become exacerbated. By paying attention to the timing of the symptoms, a TCM practitioner can detect the underlying root cause of the problem.

For example, patients with asthmatic conditions often awaken coughing with an asthma attack between 3 and 5 AM (Lung meridian period) during the pre-dawn time. The TCM doctor can utilize the time of the meridian to design the treatment protocol. For example, the best time to tonify the Kidney meridian deficiency is 5–7 PM. If the Liver meridian (1–3 AM) is excessive, one can use the Small Intestine meridian period (1–3 PM) to sedate the Liver meridian (Fig. 4) (Table 3).

According to the World Health Organization publication reporting the effectiveness of acupuncture as evidenced in reviews and analysis of clinical trials (WHO 2002) and the 2017 Acupuncture Evidence Project from Australia, acupuncture has been recognized with significant clinical results for: allergic rhinitis, chemotherapy-induced nausea and vomiting (with anti-emetics), chronic low back pain, headache (tension and chronic), knee osteoarthritis, migraine prevention, postoperative nausea and vomiting and postoperative pain. These conditions show the strongest evidence for acupuncture effectiveness, while there are hundreds of more conditions that show a positive response (McDonald and Janz 2017). In addition, acupuncture has been acknowledged in the Journal of Integrative Medicine in 2017, involving thirteen professional institutes and clinical venues in the US, as having been proven to be a non-pharmacological option as a solution to the opioid crisis in America (Fan et al. 2017).

Chinese Herbal Medicine

Nature is the most prolific and elaborate laboratory on earth, and Chinese herbal formulations have held historical influence for nearly 5000 years. China has developed into the largest population in human history, over 1.4 billion as of 2019, noted by the United Nations. This grand vision on the ultimate human experience gives TCM the perfect empirical model for human trials that has proven highly effective over several millennia and now is center stage in the global arena of pharmaceutical research and development.

Central to TCM is herbology and the cultivation, combination, and application of herbal remedies. TCM makes use of herbs and herbal decoctions to strengthen organ function and rebalance the metabolism. A deep base of knowledge and a keen sensitivity to the essence of various herbal components give the TCM practitioner a precise method of generating a healing effect that reaches beyond the chemical composition and physical properties of the herbs.

The practitioner chooses the herbal formula that best complements the individual's unique symptom profile and the body's energetic vibration, as determined through various diagnostic processes.

According to the World Health Organization (WHO), herbal medicines provide the basis of

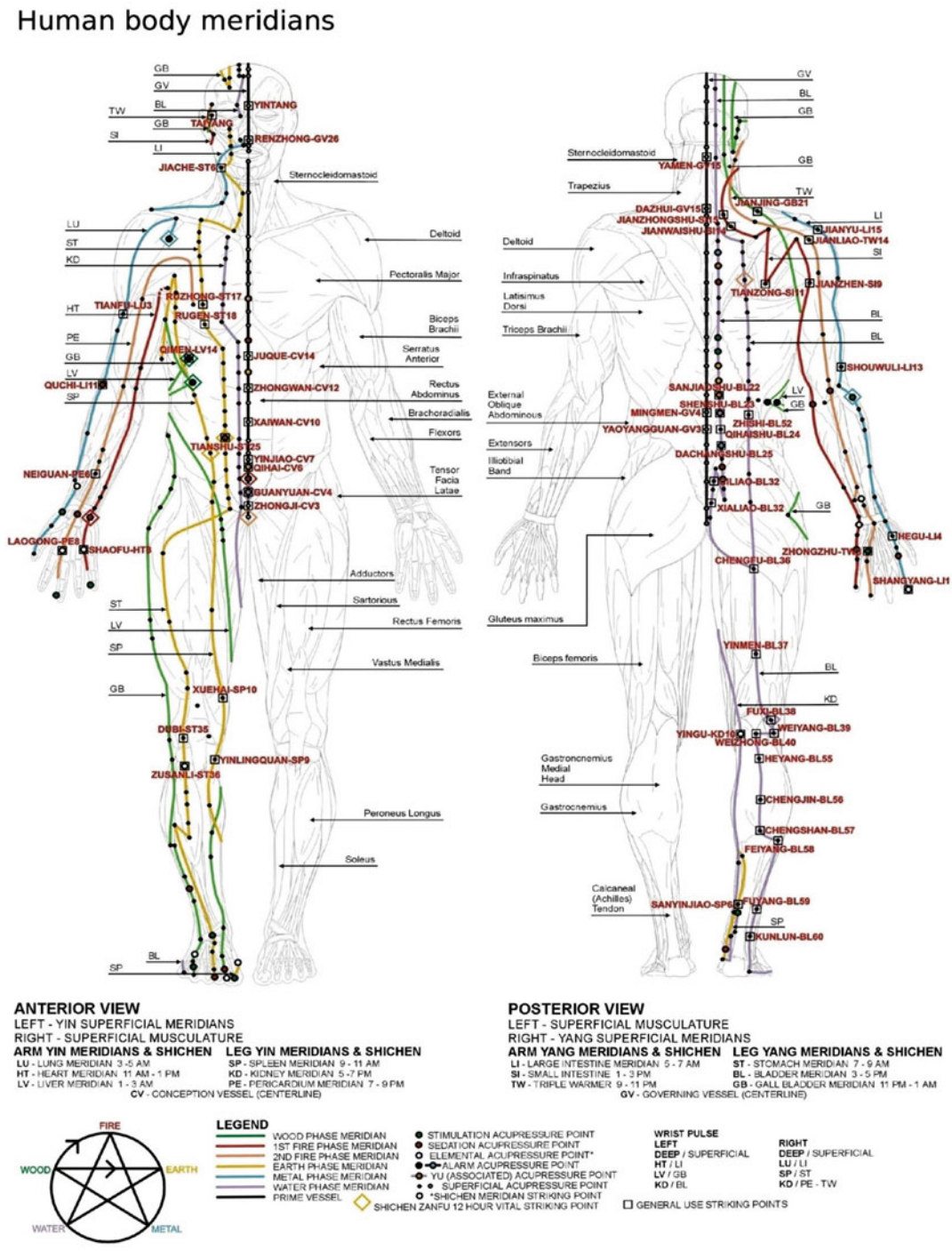


Fig. 3 Meridian channels chart –Wikimedia. (Permission to use: Creative Commons License 4.0)

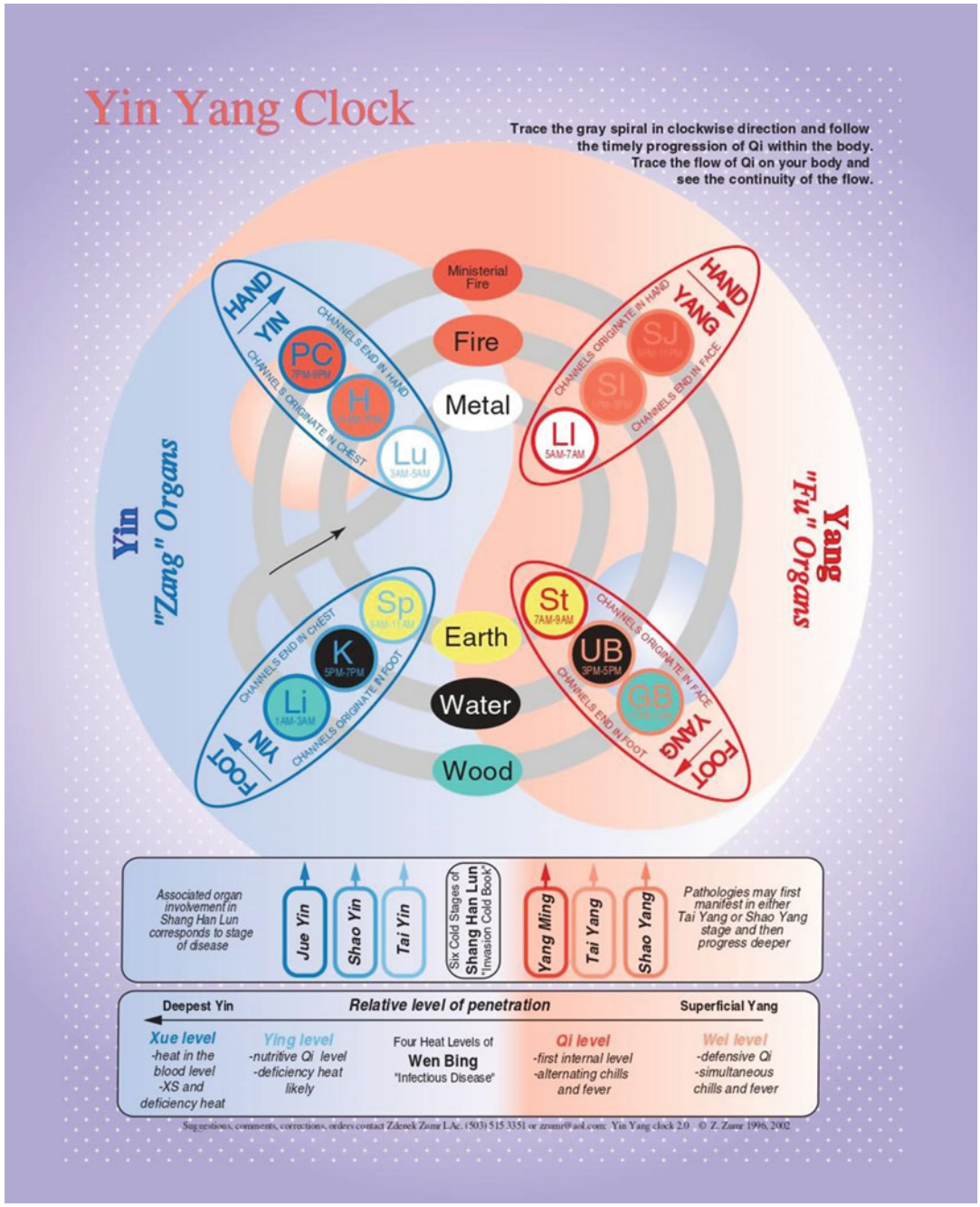


Fig. 4 Yin Yang clock. (Courtesy of Zdenek Zumr MTCM: *Introduction to Oriental Medicine through Charts and Graphs*, 1994)

Table 3 Twelve meridian time flow chart. (Courtesy of Anita Chen Marshall, DAOM, Pharm D, LAc)

Zang organs (Yin meridians)	Hours	Fu organs (Yang meridians)	Hours
Lung (Hand Tai Yin)	3–5 AM	Large Intestine (Hand Yang Ming)	5–7 AM
Spleen (Foot Tai Yin)	9–11 AM	Stomach (Foot Yang Ming)	7–9 AM
Heart (Hand Shao Yin)	11 AM–1 PM	Small intestine (Hand Tai Yang)	1–3 PM
Kidney (Foot Shao Yin)	5–7 PM	Urinary bladder (Foot Tai Yang)	3–5 PM
Pericardium (Hand Jue Yin)	7–9 PM	San Jiao (Hand Shao Yang)	9–11 PM
Liver (Foot Jue Yin)	1–3 PM	Gallbladder (Foot Shao Yang)	11 PM–1 AM

primary healthcare for more than four billion people on a global level, comprising over half of the world's population.

It is estimated that nearly 85% of traditional medicine modalities involve plant extracts that have been the foundation of modern pharmacology for centuries. There is a bounty of intrinsic and extrinsic properties that TCM recognizes as complementary to the innate healing capacity of the human metabolism. Herbs have biochemical compositions of active and buffering agents that synergistically balance each other. Some herbs are well known as food sources, while others are purely medicinal.

In contrast to the targeted medications of conventional allopathy that are largely developed and prescribed to ameliorate specific metabolic dysfunctions or specific symptoms, TCM herbal formulations work in collaboration with one another to achieve a systemic rebalancing. Each herb plays a unique role in the ultimate goal of achieving equilibrium, which is specific to each person. By combining the expertise of multiple herbal formulations, TCM can precisely address an individual patient's particular set of issues and develop a custom-designed healing protocol.

Volumes of herbal formulas, medical applications, and variables on TCM practices for individualized protocols can be found in numerous historical texts from various pivotal TCM authors and practitioners, and many still are circulated and researched today. Respect is due to the ancestral authors, whose wisdom and intuitive knowledge laid the groundwork for TCM, then and now. All TCM practitioners are grateful to pass on this heritage and offer their personal contributions for the next generation.

1. **Huang Di Nei Jing** (黃帝內經) *The Yellow Emperor's Inner Classic* (300 BCE) is the most revered of TCM's ancient texts. It is believed to have been the product of hundreds of separate writings gathered into two texts: *Suwen* (*Basic Questions* 素問) and *Lingshu* (*Spiritual Pivot* 靈樞). It is often regarded as the theoretical foundation of TCM and its diagnostic methods. It includes the dynamics between forces of nature (Yin/Yang) and the universal instinct to regain and maintain balance and harmony.
2. **Shen Nong Ben Cao Jing (**神农本草经) *Divine Husbandman's Classic of the Materia Medica* (200 CE). This manuscript detailed over 300 herbs according to thermal property, taste, toxicity, dosage, and forms. It was owing to this manuscript and this era that many widely used herbal remedies are still utilized today, such as Ma Huang (Herba Ephedrae) for asthma relief, Dang Gui (Radicis *Angelica sinensis*) for regulation of menstruation, and Huang Lian (Rhizoma coptidis) for cessation of diarrhea.
3. **Shang Han Za Bing Lun** (伤寒杂病论), *Treatise on Cold Pathogenic and Miscellaneous Diseases* (150–219 CE), was written during the Eastern Han Dynasty by physician Zhang Zhong Jing (张仲景). The main focus of the manuscript was on treatments for epidemic infectious diseases, causing fevers. In addition, he wrote *Jin Gui Yao Lue* (金贵要略) *Essential Prescriptions of the Golden Coffer*, which focuses on treatment for internal diseases.
4. **Yao Xing Ben Cao** (药性本草) *Materia Medica of Medicinal Properties* (600 CE). This manuscript elucidated the complex subjects of combination, reaction, taste, temperature, toxicity, function, primary clinical application, processing, and preparation. This is also an important text in light of it being the

first attempt to collaborate herbal empirical data, clinical application, and classical TCM theory as a cohesive compendium.

5. **Shi Liao Ben Cao (**食疗本草) *Materia Medica of Diet Therapy* (700 CE) was written by Meng Shan (孟詵;). It emphasized the importance of diet and the use of 227 herbs as therapeutic agents.
6. **Ben Cao Gang Mu** (本草纲目) *The Grand Materia Medica* (1578 CE) was compiled by Li Shi Zhen (李時珍). A formidable encyclopedia, it contains information on 1892 substances: 1173 from plants, 444 from animals, and 275 minerals. In addition, it proposed 11,000 herbal formulas for specific ailments.
7. **Ben Cao Gang Mu Shi Yi** (本草綱目拾遺) (1765 CE) was compiled by Zhao XueMin (趙學敏). This text provided a complete description of over 700 natural substances catalogued for the first time.
8. **Zhong Hua Ben Cao** (中华本草) *The Chinese Herbal* (1999 CE) was compiled by the Chinese State Administration of TCM involving over 500 scholars/experts from at least 60 academic and research institutions, and took 10 years to complete. This monumental textbook consists of 10 volumes, totaling 9,282 pages, with 8,534 illustrations and 8,980 Chinese Materia Medica (CMM) monographs. There is a condensed 2-volume version available. The modern literature covered is up to 1994. It is the most extensive documentation of CMM to date.

Characteristics in Herbal Medicine

In TCM, there is an acknowledgment of variables that are known as *characteristics*. Precise differentiation of the pattern of symptoms is correlated to the distinguished properties of each herb in herbal medicine. These characteristics reference specific value sets that are uniquely descriptive and, therefore, aid the practitioner in determining the herbal choices which correlate best to a series of symptoms and individual patient constitution.

The characteristics of Chinese Herbal medicine are divided according to taste, thermal property, direction, and channel affiliations. While different characteristics have separate evaluative criteria, each of them must be taken into account while formulating the herbal medicine. In exploring these options, for instance, choosing the correct taste, but the wrong thermal property or an inappropriate meridian, could result in an ineffective treatment. If the patient has a heat condition, they should not be given herbs with hot or warm thermal property. By doing so, it would not only be ineffective but may also worsen the heat condition. The directional affinity is used to guide the herb's function toward the treatment area. For example, *Huang Qin* (Radix Scutellariae) clears heat in the upper jiao (chest) area and, therefore, can be used to treat lung or liver diseases. *Du Huo* (Radix Angelicae Pubescentis) moves in a downward direction, useful for both treating and guiding other herbs in treatment for musculoskeletal disorders of the lower body (Fig. 5).

Contemporary Applications of TCM Herbs in Clinical Pharmacology

Historically, herbal medicine was the single most effective medicinal option available to humans. Food and plant derivatives were the centerpiece of treatment protocols before the advent of industrialization, synthesis of chemicals, and pharmaceutical development. Today, there is resurgence of interest toward potent phytochemicals and other compounds from plants to discover more effective and less toxic therapeutic modalities for infectious and degenerative diseases, as well as cancer treatments.

TCM is currently at the forefront of global medical research and is evolving exhaustive databases of the complexity of chemical compounds found within classic TCM herbal decoctions that are being currently investigated in pharmaceutical studies. Over 71,000 entries, between 1945 and 2019 (abstracts not available until the 1980s), are noted for TCM in *Pubmed.gov* from the US National Library of Medicine and the National Institutes of Health.

Additionally, there are over 3000 studies, specific to certain herbal compounds that are used in infectious and degenerative diseases, many of

Fig. 5 Herbal characteristics. (Courtesy of Anita Chen Marshall, DAOM, Pharm D, LAc)

which have developed states of resistance, despite standard antibiotics and pharmaceuticals.

Just as herbal medicine is a cornerstone of TCM, pharmaceuticals have become the mainstay of conventional medicine. Major pharmaceutical companies in the USA, Europe, and Asia are premiere research and development sources for these drugs. Much of the history of pharmaceuticals has its origin with herbal traditions, such as TCM. An examination of this history and how herbs and drugs have become synonymous for medicine in different eras can be quite enlightening. Current statistics indicate an estimated wealth of 300,000–500,000 plant species on the planet, of which only 10% has been cultivated and developed for medicinal purposes. Biodiversity efforts are being proposed to help protect these plants from extinction and to sustain their natural global habitats (Rates 2001).

Herbal compounds have a well-documented history for treating diseases. Between 35,000 and 70,000 plant species have been screened for medicinal compounds, and over 200 modern drugs owe their origins to the plant kingdom (Veeresham 2012). It was reported by Fabricant and Farnsworth (2001) that 80% of 122 plant-derived drugs were related to their original ethnopharmacological purposes.

Many of the commonly known herbal compounds used today, such as curcumin from *Jiang Huang* (Rhizoma Curcumae Longae, turmeric), *Sheng Jiang* (ginger), and *Guan Ye Lian Qiao* (*Hypericum perforatum*, aka, St. John's wort), are utilized for their anti-inflammatory, antidepressant, and antiemetic properties. Curcumin has become a major component in cancer research in the decades between 1999 and 2019 with over 21,000 studies cited in the PMC – National Library of Medicine and National Institutes of Health databases. Many of these are landmark studies indicating curcumin's significance in cancer prevention, cancer stem cell elimination, and advanced cancer therapies.

It is postulated that 60% of antitumor and anti-infectious drugs, already on the market or

under clinical trials, are of plant origin (Cragg 1997). Other plant compounds including cannabinoids, muscarine, physostigmine, yohimbine, forskolin, and colchicine have been clinically used in numerous pharmacological and biomedical studies (Williamson et al. 1996). Studies have shown a TCM herbal formula may be applied for several western classifications of pathologies.

For example, *Liu Wei Di Huang Wan* is a classic TCM formula used for the TCM pattern "Yin deficiency and tonify Kidney meridian." Recent studies have proven that this formula is effective for multiple diseases, including metabolic syndrome, diabetes, and cardiovascular disease, based on clinical experiences and pharmacological studies (Jiang 2005).

Clinical pharmacology now involves a multidisciplinary cohort of sciences including botany, chemistry, toxicology, anthropology, agronomy, and biotechnology, all of which have advanced pharmacology and improved the standards of drug development. Principles in TCM that adhere to a philosophy of interrelated networks in nature and the human physiology and psychology are now being seriously investigated in network pharmacology and systems biology.

A modern pharmacological approach, bioactivity-guided fractionation, is essential in the isolation of active substances. Because TCM often prescribes a multivariate decoction of several herbs, designed to assist a core biological response, research in standardization may not always rely on a singular compound or focus of action, but on a synergistic result, which has both pharmacodynamic and pharmacokinetic properties.

For instance, Radix Ginseng, well reputed in TCM circles, is a prime example in which the whole plant or its saponin fractions are more active than the isolated compounds (Hamburger and Hostettman 1989). As TCM and network pharmacology continue to join efforts toward cataloguing herbal compounds and designing an international platform of standardization, herbal medicines will be more acceptable to empirical review parameters of clinical study.

Herbal Medicine Formulation

Chinese herbal medicine recognizes an optimal therapeutic effect is achieved through identifying and analyzing the intricate nuances of the specific disorder in relationship to the person's unique character and constitution. Single herbs may be prescribed in the event of a clear diagnosis without additional complexities. However, as most pathological conditions are aggregations of imbalances, a system of complementary herbal remedies, custom-designed for the individual and the pattern of symptoms, has evolved from ancient days.

In Huang Di Nei Jing (*The Yellow Emperor's Inner Classic*), the basic principle of formulating an herbal medicine is described as the system of Jun Chen Zuo Shi (Chief, Deputy, Assistant, Envoy). Each herb performs an important and essential function within the formula.

This general concept has been expanded from the original, but the basic principle is still being followed in modern times.

Jun Yao (Chief Herb) is the most important component in each formula. It is used to treat the key disease or symptom, which is used at a large dose to maximize the effect. **Chen Yao (Deputy Herb)** reinforces the effects of the chief herb and also treats the associated disease or symptoms. **Zuo Yao (Assistant Herb)** reinforces the chief or deputy herb to treat the disease or directly treat the secondary symptoms. It is used to counteract the toxicity or to minimize the undesirable strong effects of the chief and deputy herb. **Shi Yao (Envoy Herb)** acts as a channel-guiding herb to direct the formula to the affected area of the body and also to harmonize all of the herbs within the formula, which is usually used in small doses.

Interactions Between Herbs

Although TCM herbal medicine has thousands of years of history attesting to its efficacy and safety, there are defined situations of toxicity that TCM practitioners are trained to understand and address. These guidelines contain six herb/herb interaction categories:

1. **Xiang Xu** (mutual accentuation), **Xiang Shi** (mutual enhancement).

   Example: combination of *Shi Gao* (gypsum) and *Zhi Mu* (Rhizoma Anemarrhenae) to "clear heat and purge" fire.
2. **Xiang Wei** (mutual counteraction) **Xiang Sha** (mutual suppression), **Xiang Wu** (mutual antagonism).

   Example: combination of *Lai Fu Zi* (Semen Raphani) and *Ren Shen* (Radix Ginseng), in which the effect of the latter herb is decreased. *Lai Fu Zi* is used as antidote for *Ren Shen* overdose (symptoms: palpitation, bloating, tremor, distress).
3. **Xiang Fan** (mutual incompatibility). Example: eighteen incompatibles (*Shi Ba Fan*) – *Gan Cao* (Radix Glycyrrhizae) is incompatible with *Gan Sui* (Radix Euphorbiae Kansui), *Da Ji* (Radix Euphorbiae seu Kaoxiae), *Yuan Hua* (Flos Genkwa), and *Hai Zao* (Herba Sargassum). *Wu Tou* (Radix Aconiti) is incompatible with *Chuan Bei Mu* (Bulbus Fritillariae Cirrhosae), *Zhe Bai Mu* (Bulbus Fritillariae Thunbergii), *Gua Lou* (Fructus Trichosanthis), *Ban Xia* (Rhizoma Pinelliae), *Bai Lian* (Radix Ampelopsis), and *Ba Ji* (Rhizoma Bletillae). *Li Lu* (Radix et Rhizoma Veratri) is incompatible with *Ren Shen* (Radix Ginseng), *Bei Sha Shen* (Radix Glehniae), *Nan Sha Shen* (Radix Adenophorae), *Ku Shen* (Radix Sophorae Flavescentis), *Dan Shen* (Radix Salviae Miltiorrhizae), *Xuan Shen* (Radix Scrophulariae), *Bai Shao* (Radix Paeoniae Alba), *Chi Shao* (Radix Paeoniae Rubra), and *Xi Xin* (Herba Asari).

In conjunction with these categories of interactions, there are two levels of cautions and contraindications that are important: safety and effectiveness factors. These include interaction of the specific herbal natures with the specific symptoms of the condition and interaction of herbs directly with other herbs. While most herbs have been proven safe individually or in combination, certain herbal combinations have created adverse side effects and reactions.

Interactions Between Herbs and Drugs

As people are increasingly taking medications and herbal supplements, especially in the geriatric population, the potential of herb and drug interactions is becoming a growing concern for healthcare professionals, as well as the general public. There are two types of interactions: *pharmacokinetic* and *pharmacodynamic*. Pharmacokinetic interactions refer to the fluctuations in bioavailability of herb and drug molecules in the body, as a result of changes in absorption, distribution, metabolism, and elimination. Pharmacodynamic interactions refer to the fluctuations in bioavailability of ingested substances, as a result of synergistic or antagonistic interaction between herb and drug molecules.

Herbs also are known to exhibit numerous interactions with drugs. Duplicating similar reactions of herbs and drugs could cause overdose or ineffectiveness of drugs in the patient's metabolism. Following are descriptions of herbal functions which may interfere with or exacerbate biochemical reactions of specific classifications of drugs.

Guan Ye Lian Qiao (Hypericum perforatum), aka St. John's wort, is commonly used as an OTC supplement for depression and it is a liver enzyme inducer, which will lower the blood level of other medications if given together, leading to decreased therapeutic effect. It may also cause serotonin syndrome, if given concurrently with Selective Serotonin Reuptake Inhibitor (SSRI) drugs. Another herb, *Chai Hu* (Radix Bupleuri), should be used with caution with the drug interferon, as this herb has shown increased risk of acute pneumonitis. The herb may over stimulate neutrophils to release granulocytes elastase and oxygen radicals which may cause damage to lung tissues.

Physicians and clinical pharmacologists may be interested in exploring possibilities of TCM herbal medicines that offer similar bioactive properties as pharmaceutical drugs, yet with fewer to no adverse reactions.

Understanding both the chemical compatibility and incompatibility between herb and drug functions can develop options in research and treatment protocols that may provide safer and more effective modalities. Especially in major cardiovascular

areas, where multiple drugs are often prescribed to achieve one effect, such as modulating hypertension or anticoagulation, TCM herbs may be more targeted to achieve the same results with a single herbal compound. Also, in areas of diabetes, depression, and anxiety, TCM herbal medicines have proven to be effective alternatives to more toxic multiple drug regimens. As with all active substances that can alter metabolic functions, caution should be used when prescribing any herb or drug that could be contraindicated. Table 4 displays certain interactions of herbs and drugs that should be considered in any combined treatment protocol.

Herbal Processing, Quality Assurance, and Standardization

The herbal extracts can be derived from the raw herbs through a solvent, such as grain alcohol, or an alcohol-free solvent, such as vegetable glycerin. However, the herbal decoction is the most

Table 4 Herb functions and drug interactions. (Courtesy of Anita Chen Marshall, DOAM, PharmD, LAc)

Herb pharmacological functions	Drug interactions
Elevate blood pressure and stimulate **heart muscle** **Sympathomimetic:** *Ma Huang* (Herba Ephedrae), *Qing Pi* (Pericarpium Citri Reticulatae Viride), *Zhi Shi* (Fructus Aurantii Immaturus), *Zhi Ke* (Fructus Aurantii) **Hypertensive**: *Gan Cao* (Radix Glycyrrhizae) **Positive inotropic**: *ChenPi* (Pericarpium Citri Reticulatae), *Fu Zi* (Radix Aconiti Lateralis Praeparata), *Ren Shen* (Radix Ginseng), *Wu Zhu Yu* (Fructus Evodiae) **Note:** **Hypotensive:** *Gou Teng* (Ramulus Uncariae cum Uncis), *Di Long* (Pheretima), and *Ginkgo biloba* (Folium Ginkgo) may cause hypotension when used with antihypertensive drugs	**Antihypertensive drugs:** Caution: may increase risk of hypertension **Calcium channel blockers**: amlodipine (Norvasc), nifedipine (Procardia, Adalat), efonidipine (Landel), felodipine (Plendil) **Beta-blockers**: metoprolol (Lopressor), labetalol (Normodyne, Trandate), propranolol (Inderal), carvedilol (Coreg) **ACE inhibitors**: lisinopril (Prinivil, Zestril), benazepril, (Lotensin), captopril (Capoten) **Angiotensin receptor blockers**: losartan (Cozaar), valsartan (Diovan)
Improve blood circulation, inhibit platelet aggregation *Dan Shen* (Radix Salviae Miltiorrhizae), *Ji Xue Teng* (Caulis Spaatolobi), *San Qi* (*Panax* Notoginseng), *Tao Ren* (Semen Persicae), *Pu Huang* (Pollen Tyhpae) ((He, LS et al. 2016)	**Anticoagulant and antiplatelet drugs** Caution: may lead to prolonged PTT and excessive bleeding. Examples: warfarin (Coumadin), heparin, apixaban (Eliquis)
Drain damp effect (increase urine output) *Fu Ling* (Poria), *.Zhu Ling* (*Polyporus*), *Che Qian Zi* (Semen Plantaginis), *Ze Xie* (Rhizoma Alismatis), *Bai Zhu* (Rhizoma Atractylodis Macrocephalae)	**Diuretic drugs** Caution: may lead to excessive fluid and electrolyte loss Examples: hydrochlorothiazide (Microzide), furosemide (Lasix), spironolactone(Aldactone)
Lower blood sugar *Zhi Mu* (Radix Anemarrhenae), *Shi Gao* (Gypsum Fibrosum), *Shan Yao* (Rhizoma Dioscorea), *Huang Qi* (Radix Astragalus), *Xuan Shen* (Scrophularia ningpoensis), *Cang Zhu* (Rhizoma Atractylodis)	**Antidiabetic drugs** Caution: may cause hypoglycemia metformin (Glucophage), tolbutamide (Orinase), glipizide (Glucotrol), glyburide (Glynase, Micronase), pioglitazone (Actos), insulin
Calm the mind (shen) sedative *Suan Zao Ren* (Semen Ziziphi Spinosae), *Yuan Zhi* (Radix Polygalae), *Chai Hu* (Radix Bupleuri), *Tien Ma* (Rhizoma Gastrodiae), *Ye Jiao Teng* (Caulis Polygoni Multiflori), *He Huang Pi* (Cortex Albiziae), *Fu Shou* (Fructus Citri Sarcodactylis)	**CNS suppressant drugs** Caution: may cause excessive drowsiness and sedation **Opioids:** morphine, codeine, hydrocodone, oxycodone **Antidepressants**: selective serotonin reuptake inhibitors (SSRI), fluoxetine (Prozac), fluvoxamine (Luvox CR), paroxetine (Paxil) **Antipsychotics**: aripiprazole (Abilify), quetiapine (Seroquel) **Antiseizure**: phenytoin(Dilantin), pregabalin (Lyrica) gabapentin(Neurontin) **Benzodiazepines:** diazepam (Valium), Lorazepam (Ativan), Alprazolam (Xanax) **OTC sleep aids:** valerian, kava, melatonin, diphenhydramine (Benadryl), doxylamine (Unisome)

common dosage form for herbs in China. The herbs are presoaked in enough water to cover them for a period of time, depending on the nature of the herbs and purpose. The herbs are then cooked until the water boils and then let simmer for a period of time, between 30 and 60 min, depending upon the particular blend of herbs, before the decoction (tea) is ready for consumption. The remainder of the herbs can be boiled the second time or discarded. In addition, a variation on the herbal decoction can be freeze-dried and condensed into granules. Herbal extracts and granules, similar to contemporary medications, are more convenient for patients to take than the customary traditional raw herbs.

The process of preparing the raw herbs can be complex in timing. Leaves and flowers take less time to decoct than barks and roots; therefore, they should be added to the decoction process last. *He Shou Wu* (Radix Polygoni Multiflori) should be decocted with black beans at least for 30 min to decrease the side effects of loose stools, diarrhea, and liver toxicity. Long-term consumption of *He Shou Wu* has been associated with liver toxicity. If discontinued, adverse reactions are reversible. However, if it is overdosed through unmonitored consistent consumption, cirrhosis or death may occur (Lei et al. 2015).

Many factors must be considered in the cultivation, concentration, toxicity/adverse reactions, and standardization of herbal medicines. These include geographical origin, varying processing and manufacturing techniques, as well as contraindications and compatibility with other herbs and pharmaceuticals (Zhang et al. 2017).

As language can be a challenging issue in identifying a Chinese herb, in the interests of international standardization, use of both the Chinese term and the binomial, including the Latin botanical name and/or pharmaceutical name and any specific genus, would help to establish a global consensus regarding the exact herb in question. Effective monitoring of herbal medicine safety will result from the collaboration of botanists, phytochemists, pharmacologists, and other major contributors to the cultivation, processing, and distribution of herbal medicines (Ekor 2014).

The accurate processing of herbal decoctions can reduce the toxicity of a few specific herbs. For example, improper preparation or overdose of *Fu Zi* (Radix Aconiti Lateralis Praeparata) is the main cause of possible adverse reactions. A person with a weak constitution may also have toxic reactions to *Fu Zi*, even when taken in low doses. Patients who have a past history of cardiovascular disease, or are taking anti-arrhythmic medications, should take *Fu Zi* with extreme caution due to its strong cardiovascular effects. *Fu Zi* should be decocted for more than an hour until the roots are very soft to decrease strong undesirable side effects and toxicity. Gross overdose can be fatal.

Proper processing, concentration, and dosages are as relevant in clinical pharmacology. Quality assurance and manufacturing integrity are related to effectiveness and toxicity, as well as less regulated sources, such as the generic drug industry. Quality assurance is now a major issue regarding the safety standards and potency of drug manufacturing for both herbal medicines and pharmaceuticals.

Prime examples of this are the industry recalls in 2018–2019 of losartan and valsartan, the widely prescribed angiotensin II receptor blockers for treating hypertension. In addition, the popular OTC medication for heartburn, Ranitidine (Zantac), was recalled in October 2019. Generic manufacturers are under the FDA microscope since it was discovered that these drugs were contaminated with ingredients containing trace amounts of *N*-nitroso-*N*-methyl-4-amino butyric acid (NMBA) and other carcinogens. The genotoxic impurities may be caused by changes made in the active pharmaceutical ingredient (API) manufacturing process but could also came from the reuse of materials, such as solvents. The growing concern over adulterated products and contamination of the API in China and India has grown into a regulatory challenge as to how to improve quality while maintaining cost-effectiveness.

Up to 70% of drug manufacturing plants in China and India are being rated as substandard: falsifying records, tainting products, and allowing unsanitary conditions (*Bottle of Lies: The Inside Story of the Generic Drug Boom*, book by

Katherine Eban 2019). It appears that the pharmaceutical industry in the USA and Europe, which obtains 80% of its raw materials from subcontracted global sources, is finding itself in a challenging situation. Inspection of the raw materials and these offshore manufacturing facilities and processes is integral to the future of safer and more effective pharmaceuticals, as well as herbal medicines and nutraceuticals (*Identifying the Root Causes of Drug Shortages and Finding Enduring Solutions* (Duke-Margolis et al. 2018) Duke University and FDA presentation-Washington, DC, November 27, 2018).

Herbal Medicine and Pharmaceuticals: Toxicity, Adverse Reactions, and Mortality

TCM herbal medicine is often scrutinized and criticized by the scientific community for nonregulated issues of safety and efficacy, as well as lack of evidence-based, peer-reviewed studies. The interesting aspect of empiricism is that it is based on observation, induction, deduction, experimentation, and evaluation. TCM has a documented history of such empirical values in the development of its multivariate diagnostics, combinative herbal therapies, and intrinsic relationship with the philosophical and physiological factors of human nature within the entirety of nature. The isolation of critical active ingredients may often be misleading in verifying how nature navigates both in the environment and within living inhabitants.

As regards toxicity and adverse reactions to TCM herbal formulations, there are several encyclopedias and collected documents, some of which are being translated into English. These include the *Pharmacopoeia of the People's Republic of China* (PPRC), *Toxic Chinese Materia Medica: Past and Present* (Du Ju Zhong Yao Gu Jin Yong), *Encyclopedia of Toxic Chinese Materia Medica* (You Du Zhong Cao Yao Da Ci Dian), *Toxic Drugs Herbal* (Du Yao Ben Cao), and the *Modern Toxicology of Chinese Materia Medica* (Xian Dai Zhong Yao Du Li Xue).

China's State Administration of Traditional Chinese Medicine requires all TCM practitioners to obtain the education and Chinese medicine degree to be officially registered and practice legally. In the Traditional Medicine Strategy 2014–2023 report, the WHO addressed concerns regarding toxicity and adverse reactions to TCM herbal medicines to strengthen safety protocols of TCM herbal products by global practitioners. The International Organization for Standardization (ISO) is in the process of developing quality standardization options for Chinese herbal medications, inclusive of raw materials, processing, and finished product quality assurance. These efforts will be beneficial in the regulation and standardization of the globalized marketing of TCM herbal medications (Zhou et al. 2019).

The concepts of toxicity and adverse reactions are no strangers to TCM, historically. In the Qing dynasty (1644–1911 CE), *Herbal Lihai* authored by Ling Huan identified a variety of species of medicinal plants in the order of three aspects: harm, benefit, and therapy. The harmful side effects and contraindications were listed in the beginning, indicating that TCM practitioners recognized possible adverse effects of herbal remedies, long before the age of clinical pharmacology (Zeng and Jiang 2010).

According to a 2014 report by Donald Light, PhD, from the Edmond J. Safra Center for Ethics at Harvard University, the following information depicts the adverse reactions and deaths from prescribed drugs in the USA and European Union as gleaned from systemic hospital reviews: serious adverse reactions result in over 1.9 million hospitalizations a year, with 840,000 hospitalized patients being given drugs that cause serious adverse reactions for a total of 2.74 million serious adverse drug reactions annually. In addition, prescribed drugs result in an estimated 128,000 deaths annually in the USA alone, accounting the use of prescription drugs being the 4thleading cause of death, along with stroke. The European Commission estimates that adverse reactions from prescription drugs cause 200,000 annual deaths (Light 2014).

The International Regulatory Cooperation for Herbal Medicines (IRCH) and the WHO Collaborating Centre for International Drug Monitoring are focusing on the regulation, education, and

standardization of traditional herbal medicine throughout the 129 member states. The 2019 adoption of TCM into the WHO ICD-11 calls for member states to begin reporting statistics and evidence-based data by January 1, 2022.

Once data is collected and analyzed, adverse reactions and morbidity and mortality rates for traditional herbal medicines may be discerned by equitable comparison of pharmaceuticals and TCM herbal medicines.

Antimicrobial Resistance and Herbal Medicine

Antimicrobial resistance is one of the most urgent health risks of our time and threatens to undo a century of medical progress, said Dr. Tedros Adhanom Ghebreyesus, WHO Director-General. Drug-resistant bacteria and viruses have adapted over several decades due to rampant use of antibiotics. There has been a monumental rise of refractory infections causing thousands of deaths annually worldwide.

The standard anti-pathogenic strategies of clinical pharmacology through modern medicine have reached a pivotal turning point in effectiveness, and a different approach needs to be investigated to reduce the current epidemic levels of antimicrobial resistance to drug therapies. There are a number of herbs that exhibit antibacterial and antiviral properties that may provide ways to decrease the development of antimicrobial resistance.

The WHO has enacted a new program to address resolution of this antimicrobial resistance dilemma entitled AWaRe (Access, Watch and Reserve). Released in October 2019, the program outlined a tool was developed by the WHO Essential Medicines List to contain the rising resistance and make antibiotic use safer and more effective. It classifies antibiotics into three groups: Access, Watch and Reserve to help emphasize their optimal uses and potential for antimicrobial resistance. (1) Access group: Have activity against common infections while have lower potential to develop resistance. (2) Watch group: Have higher resistance potential, should be prioritized as key target to be monitored. (3) Reserve group: Should be used sparingly and reserved for treatment of infections caused by multi-drug resistant organisms. It should be used only as a last resort.

The AWaRe campaign promoted by the WHO has mandated several actions areas including increasing the consumption of antibiotics in the Access group to at least 60% and reducing the use of the antibiotics in most at risk of resistance from the Watch and Reserve groups. Using Access antibiotics lowers the risk of resistance because they are “narrow-spectrum” antibiotics (that target a specific microorganism rather than several). They can be made available in generic forms and represent a less costly treatment option.

TCM and Infectious Disease

Conventional western medicine addresses infectious disease by a linear approach in suppressing and eliminating suspected viruses and bacteria that are presumed to be the primary causative factors. TCM and herbal medicine tend to offer a more comprehensive approach to infectious disease by applying the syndrome differentiation principle, a core feature of the TCM approach to individualized treatment. It has a vantage point in going beyond the viral and bacterial etiology concept to address the environment of the patient constitution in which the pathogens exist and the numerous influences affecting the patient’s natural immunity. TCM protocols not only consider inhibition of pathogens but also, simultaneously, establish balance in the body and maintain organ recovery during severe and drug-resistant infection.

TCM acknowledges two syndromes, deficiency and excess, which occur commonly in severe infectious disease situations. In these cases, distinct treatments are used to dispel the pathogens by choosing either a deficiency or excess protocol which may include detoxification, clearing heat, cooling blood, removing blood stagnation, tonifying Qi, and nourishing Yin.

By carefully appropriating the treatments to the syndrome differentiation, these methods can reduce inflammation, ameliorate circulation disorders, and aid in tissue fibrinolysis.

According to a comprehensive report on SARS (Severe Acute Respiratory Syndrome) from the WHO (2004), TCM has been shown to be an valuable adjunct therapy with western medicine for treatment of the epidemic outbreak of SARS in 2003. It should be applied at the beginning stage of the infection while considering a patient's individual differences and underlying conditions. The outbreak of Coronavirus Disease (COVID-19), originated from Wuhan, China in December 2019, spread rapidly in China and around the world, was declared pandemic by the WHO on March 11, 2020. There were 1,853,155 confirmed cases with 114,247 death reported by April 12, 2020. In February 2020, there were a few reports from China indicating that TCM and integrated therapies had shown advantages over western medicine alone in treating hospitalized patient with severe acute respiratory syndrome.

A report from Chen, JK. (2020), compiled and translated 3 different reports from China on *How COVID-19 (2019-nCoV) is Currently Treated in China with TCM*. It listed 4 different stages of treatment for COVID-19 in China. (1) Prevention phase (2) Influenza phase (3) Pneumonia phase (4) Recovery phase. Each phase presented a different clinical condition; therefore, it required different herbal formulas to treat the underlining conditions. The advantages of TCM treatments are primarily due to its effectiveness in relieving symptoms, decreasing lung inflammation, improving the degree of blood oxygen saturation, modulating immunological responses, reducing the required dosage of glucocorticoid and other Western medicines, decreasing the morbidity and mortality rate, and lowering the cost of treatment by shortening the duration of infectious stage. Since there is currently no effective treatment and no vaccine for COVID-19, the national authorities, health care workers and researchers should be inspired to explore the potential of a TCM integrative treatment approach.

The following categories of TCM herbs show antiviral, antibacterial, antifungal, and antiparasitic properties by their primary biochemical compounds and exhibit bioactivities in relation to the unique disease etiology:

1. **Antiviral herbs**: *Jin Yin Hua* ((Flos Lonicerae), *Lian Qiao* (Fructus Forsythiae), *Ban lan Gen* (RadixIsatidis), *Da Qing Ye* (Folium Isatidis), *Huang Qin* (Radix Scutellariae), Yu Xing Cao (Herba Houttuyriae), Xian Chun Ye (Folium Toonae Sinensis). Bioactivity: inhibition of viral replication and secretion of inflammatory cytokines from macrophages. May contain: tannins, flavonoids, polysaccharides, and alkaloids. *Xian Chun Ye* (Folium Toonae Sinensis) contains TSL-1 that selectively bind to SARS-CoV virus to inhibit the cellular entry of the virus into the host cells, it may be an important adjunct therapy in treating SARS-CoV (Chen CJ, et al. 2008), which is closely related to SARS-CoV2. More research should be done to investigate the effectiveness of adding this herb to the treatment for COVID-19.
2. **Antibacterial herbs**: *Jin Yin Hua* (Flos Lonicerae), *Pu Gong Ying* (Herba Taraxaci, dandelion), *Da Huang* (Radix et Rhizoma Rhei, rhubarb), *Huang Lian* (Rhizoma Coptidis, *Coptis* root), *Di Fu Zi* (Fructus Kochiae, broom cypress fruit, *Kochia*). Bioactivity: disruption of bacterial metabolism, structure, and functions; reversal of bacterial resistance; may contain organic acids, alkaloids, flavonoids, quinones, and volatile oils
3. **Antifungal herbs**: *Ku Shen Gen* (Radix Sophorae Flavescentis, sophora), *She Chuang Zi* (Fructus Cnidii, cnidium fruit), *Di Fu Zi* (Fructus Kochiae). Bioactivity: inhibition of fungal growth; may contain terpenoids and aldehydes, flavonoids, alkaloids, phenols, and volatile oils
4. **Antiparasitic herbs**: *Bing Lang* (Semen Arecae, betel nut), *Qing Hao* (Herba Artemisiae Annuae, wormwood). Bioactivity: induction of paralysis, blockage of nutrient absorption, inhibition of metabolism, and

elimination of parasites; may contain alkaloids and phenols (Ma et al. 2019)

Advancements in Cancer Research with TCM

Cancer research has gained unparalleled momentum as the world population topped 7.7 billion in 2019 with over 14 million new cases of cancer being diagnosed annually and over 9.6 million cancer-related deaths reported worldwide. New cancer cases are expected to rise, inexorably, to 23.6 million by 2030. Cancer now has become the second leading cause of death globally (WHO 2018). These staggering statistics point to the imperative for all nations to focus on more effective protocols for cancer prevention and treatment.

Meta-analyses of randomized controlled studies involving TCM as adjunctive supportive cancer therapy show improved survival rates and performance status of chemotherapy and radiation treatments. TCM in partnership with clinical pharmacology offers promise for such research and development, both as adjuvant and primary cancer treatments. The following information has been excerpted from the chapter "Principles of Supportive and Palliative Care/Chinese Herbal Medicine as Adjunct Therapy in Patients with Lung Cancer" published in the textbook *Modern Thoracic Oncology*, Vol. 1, 2018.

Extensive evaluations have been conducted evaluating the specific mechanisms by which adjunctive Chinese herbs address the needs of cancer patients. Table 5 shows the bioactivities of herbs exhibiting therapeutic and palliative action (McCulloch et al. 2018).

One remarkable adjuvant herbal medicine development getting examined since 1999 is known as PHY906. The original herbal formula in TCM is referred to as *Huang Qin Tang* and is comprised of four herbs: *Huang Qin* (Radix *Scutellaria*), *Bai Shao* (Radix Paeoniae Alba), *Gan Cao* (Radix Glycyrrhizae), and *Da Zao* (Fructus Jujubae). Huang Qin Tang has been used for centuries in Asia to treat gastrointestinal disorders, nausea, and diarrhea.

Table 5 Bioactivities of TCM herbs. (Modeled after McCulloch et al. (2018))

Bioactivity	TCM herbs
Antiproliferative	*Bai Hua She She Cao* (Herba Oldenlandia), *Huang Qin* (Scutellaria baicalensis), *Jiang Huang* (Rhizoma Curcumae Longae, turmeric), *Ling Zhi* (reishi mushroom or *Ganoderma*)
Apoptotic	*Jiang Huang* (Rhizoma Curcumae Longae, turmeric), Green tea (*Camellia sinensis*)
Inhibitive (NF-kappa B, COX-2, telomerase) Pulmonary fibrosis (radiation-induced)	*Jiang Huang* (Rhizoma Curcumae Longae, tumeric) Green tea (*Camellia sinensis*) *Dang Gui* (*Angelicae sinensis*)
Synergistic with chemotherapy agents (anthracycline, cisplatin, and irinotecan)	*Green tea* (Camellia sinensis)
Reductive and preventive (radiation-induced) Enteritis DNA damage Pneumonitis	*Ren Shen* (Radix Ginseng, Panax ginseng) *Dang Gui* (Radix Angelicae sinensis, dong quai) *Bai Shao* (Radix Paeoniae Alba)
Bone marrow and intestinal toxicity	*Shu Di Huang* (Radix Rehmanniae Preparata) and *Chuanxiong* (Rhizoma Ligustici Chuanxiong) *Dong Chong Xia Cao* (*Ophiocordyceps sinensis*)

PHY906 has now been formulated as a powder containing a spray-dried aqueous extract derived from the four principal herbs. The research was undertaken by Professor Ying Chi Cheng, who co-founded with Yale University's Office of Cooperative Research, PhytoCeutica, a biopharmaceutical company dedicated to developing cancer drugs from TCM herbs. According to the National Cancer Institute's drug trial information site: ". . .PHY906 possesses a wide range of pharmacological activities such as the enhancement of oral uptake of pharmacologically active agents, inhibition of CYP3A4, modulation of certain

cytokines, macrophages and lymphocytes, and inhibition of expression of MMP, NF-kB, beta-glucuronidase, the NK-1 receptor, and the delta-opioid receptor" (Dillon 2007).

TCM has now been shown to be a clinical competitor of pharmaceutical drugs in efficacy with recent studies proving the bio-chemotherapeutic effects of synergistic herbal compounds in oncological applications. These same herbs are also proving superior ratings over standard of care in decreased adverse effects and toxicity levels.

The result of a German study on pancreatic cancer elimination with dietary and herbal compounds was published in 2014 in the International Journal of Oncology (Li et al. 2014).

The unique aspects of this study are the bioactivities of a collection of multiple agents found as superior to a singular agent in increasing apoptosis and remission and decreasing tumorigenicity and metastasis and the targeting of cancer stem cells as primary in elimination of advanced pancreatic ductal adenocarcinoma. PDA is a highly aggressive malignancy with poor prognosis given the current therapeutic options, which do not target cancer stem cells (CSCs), considered to be the central cause for the aggressiveness.

Polyphenols from dietary agents, sulforaphane and quercetin, along with the green tea catechin EGCG, resveratrol, genistein, and curcumin, hold promise as anti-CSC agents in PDA. In vitro data and mouse experiments suggest that sulforaphane eliminates pancreatic CSCs by inhibition of bioactive agents for cancer stem cells targeting NF-κB activity and self-renewal potential and sensitizes the cells to apoptosis induction. The combination of dietary agents was found to be superior in reducing the self-renewal potential, viability, and migratory potential, along with increased induction of apoptosis.

In addition, the diterpenoid triepoxide triptolide was noted to be a promising combination partner. Triptolide has a long history in Traditional Chinese Medicine for the treatment of rheumatoid arthritis and cancer. It is an active compound of Lei Gong Teng (*Tripterygium wilfordii*), demonstrated in a prior study to effectively inhibit NF-κB activity, epithelial-mesenchymal transition, and stemlike features in PDA cells (Li et al. 2014). Previous studies have shown chemopreventive nutritional polyphenols, and isothiocyanates may neutralize genetic defects by epigenetic regulation, the potential reason why these substances attenuate the processes of tumorigenesis, progression and metastasis, and sensitization for drug treatment. Future studies will need to address an "epigenetic diet" created for the presence of selected bioactive agents for prevention and treatment of cancer (Appari et al. 2014).

The previous illustrarion (Fig. 6) details many of the study's findings regarding the links with genetics, epigenetics, and tumorigenicity and how cancer stem cells are biologically affected by the compounds found in TCM herbal medicines and dietary considerations.

The emerging sciences of nutrigenetics and nutrigenomics are embracing these concepts and developing strategies to educate healthcare professionals and patients as to the imperative of nutritional therapy in healthcare. The evidence contained in the study relating to the elimination of cancer stem cells by phytochemical influences should motivate advanced investigation for future cancer research and treatment protocols.

Network Pharmacology and Systems Biology: The Next Generation of Herbal Medicine

Aligned with the meridian concepts of TCM, network pharmacology includes regulation of the signaling pathway with multiple channels, upgraded drug efficacy, and higher success rates for clinical trials, as well as reduced adverse reactions and decreased budgetary considerations of pharmaceutical research and development. Network pharmacology also explores the parameters of how and where a singular target can inhibit or activate disease phenotypes.

This vital information can assist in developing therapies that are less vulnerable to drug resistance. These mechanisms can be accomplished by a synergistic effect, much like the TCM principles and practices that embrace a multivariate, collaborative approach. Network pharmacology also offers molecular interaction validation

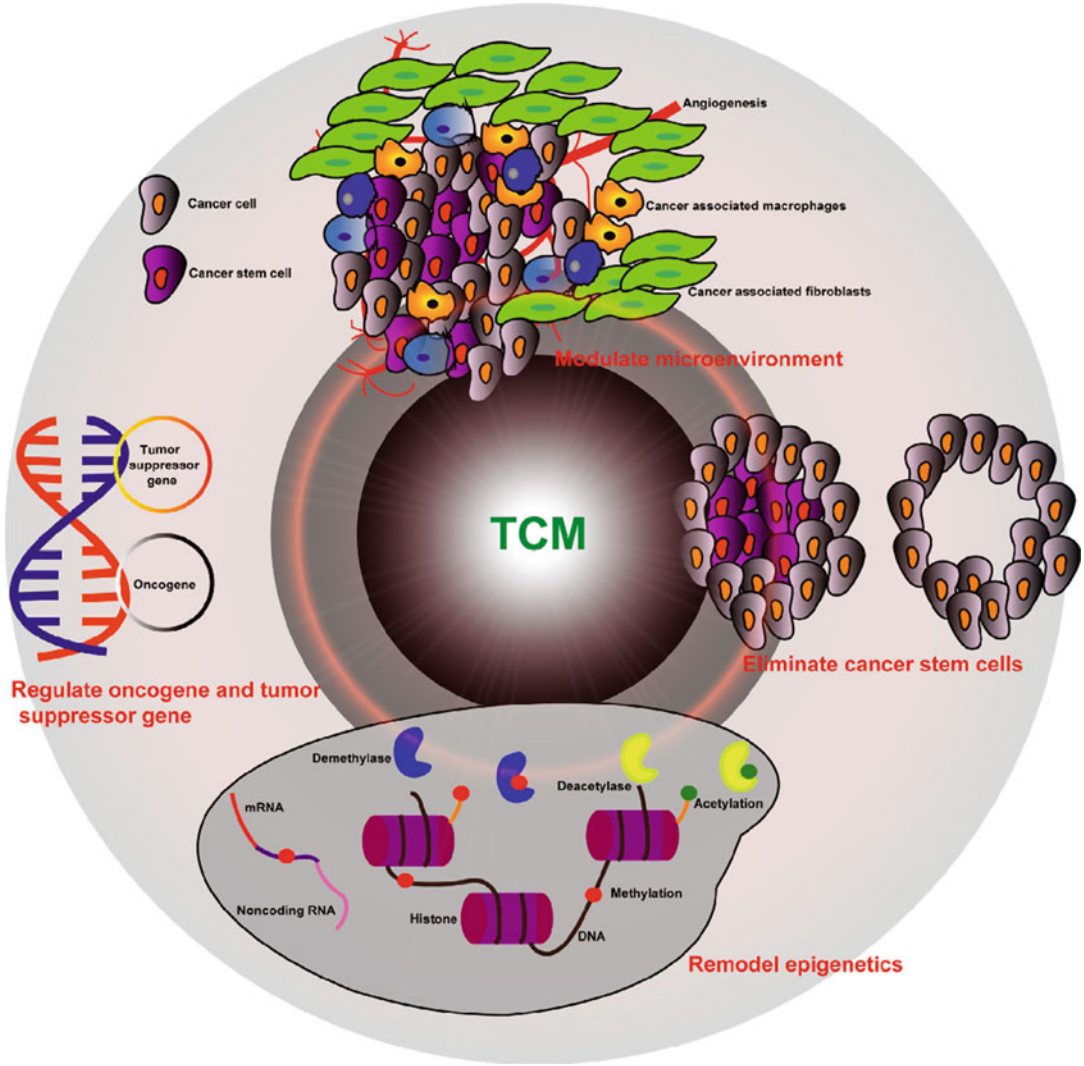


Fig. 6 Schematic diagram demonstrating the application of TCM in cancer therapy based on genetics, epigenetics, tumor microenvironment, and cancer stem cells. (Use Permission by Creative Commons Licensure 4.0: © 2019, Yuening Xiang, Zimu Guo, Pengfei Zhu, Jia Chen, Yongye Huang *Cancer Medicine* published by John Wiley & Sons Ltd)

technology as a viable tool which reveals drug activity mechanisms and verifies the predictability of a drug network.

Owing to its origins to bioscience research, systems biology is a recent trend, providing a more holistic view on the complexity of biological systems interactions. This is a more encompassing perspective than isolating and altering a singular molecular component. Network pharmacology, as a systems biology-based methodology, replaces the standard accepted drug design of a solo pharmaceutical by investigating multiple activities that drugs may exert collectively on biological networks. There is a search for combining therapeutic drug protocols that are emerging in network pharmacology.

In allegiance with systems biology, network pharmacology is challenging the single-target drug research and development model in favor of a more comprehensive and collaborative approach.

Over three decades of pharmaceutical development has produced hundreds of popularized and largely successful drugs. However, "the single-symptom, single-disease, single-drug concept" has not been as influential on certain degenerative diseases, including hypertension, depression, diabetes, inflammation, and cancer. While they may have marginal success in management, all too often, singular drug therapies are inadequate, and multiple drugs or increased dosages are recommended, causing an increase in adverse reactions and heightened toxicity, without achieving the needed metabolic therapeutic result (Kola and Landis 2004). Subsequent to these pharmaceutical pitfalls, the shift is toward a "network-target, multiple-component-therapeutics" mode (Li et al. 2013, 2014).

TCM exemplifies the natural order when approaching treatment protocols. Restoring balance is paramount, and reflecting the body's natural ability to heal is considered most valuable when choosing combinative herbal prescriptions. Network pharmacology is adopting some of these ancient concepts by utilizing TCM's wider lens on multiple action and reaction formulations and integrating the information of disease phenotypes, genetic influences, and biological molecules in a more integrated research model (Li et al. 2013).

TCM and Clinical Pharmacology In Silico

One of the frontline advantages of network pharmacology and systems biology is the intimate partnership with bioscience informatics and multiple exhaustive global databases being constructed. A number of databases have emerged over the past decade, which bring TCM herbal medicine and network pharmacology to the same dialogue through informatics, algorithms, data mining, and extensively designed digital platforms. As an example, TCM-Mesh highlights TCM network pharmacology through cataloguing diseases, herbs, compounds, genes, proteins, and drugs into a comprehensive and cross-referenced biological digital data vault (Zhang et al. 2019). Other relevant network pharmacology and TCM databases include:

1. **Asian Anticancer Materia Database, Institute of East-West Medicine, USA**. This summarizes 700 kinds of anticancer drug information from Asia, of which 80% are derived from medicinal plants. This database also includes the commonly used Chinese medicine name, Latin name, medicinal properties, the major compounds, and other information (Yi F et al. 2018).
2. **Chem-TCM, Institute of Pharmaceutical Science at King's College London, UK**. This contains more than 350 TCMs and over 9500 compounds. It records the compounds, related plants, chemical properties, common target activities, and other information.
3. **Chinese National Compound Library, National Health and Family Planning Commission of the People's Republic of China**. This is a library of small-molecule compounds consisting of core libraries and satellite libraries. It contains physical and chemical information of nearly two million small molecules.
4. **CHMIS-C (Comprehensive Herbal Medicine Information System for Cancer), University of Michigan Medical School, USA**. This provides 527 anticancer herb prescriptions, 937 components, and 9366 small-molecule structures for the clinical treatment of different types of cancer, combined with a reference database and a molecular target aided database.
5. **CNPD (Chinese Natural Products Database), Shanghai Institute of Materia Medica, Chinese Academy of Sciences, China**. The CNPD database currently collects more than 57,000 natural products from 37 categories, of which 70% of the molecules are drug-like molecules. The relevant data include the CAS number, name, molecular formula, molecular weight, melting point, and other physical and chemical properties of natural products.
6. **MAS3.0, CapitalBio, China**. MAS (molecule annotation system) is a whole data mining and function annotation solution to extract and analyze biological molecule relationships from public knowledge base of biological molecules and signification. MAS analysis platform is a web client program for interactive navigation. MAS uses relational databases of biological

networks created from millions of individually modeled relationships between genes, proteins, diseases, and tissues.

7. **PharmMapper Server, Shanghai Institute of Materia Medica, China**. PharmMapper Server is a freely accessed web server designed to identify potential target candidates for the given probe small molecules, drugs, natural products, or other newly discovered compounds with binding targets unidentified using pharmacophore mapping approach (Liu XF et al. 2010).

Nobel Prize Honors TCM

Qinghao (Herba Artemisiae Annuae), a TCM herb, provided the basis for an anti-malaria formula, since the protozoan had become resistant to quinine, an older medication used for treating the pathogen. This "new" discovery had been revealed in Ge Hong's *A Handbook of Prescriptions for Emergencies* (肘后备急方), an ancient medical and scientific document. In 1972, after screening over 2000 different Chinese herbs, phytochemist and researcher Youyou Tu and a cohort of scientists uncovered the exact formulation of artemisinin (*qinghaosu* 青蒿素) and dihydroartemisinin from Qinghao.

A graduate of the Department of Pharmaceutics at Beijing Medical College, Tu was chosen to join the China Academy of Chinese Medical Sciences and currently holds the highest ranking post of Chief Scientist in the Academy. According to Professor Tu, malaria was one of the epidemic diseases with the most comprehensive records in TCM literature. She found a landmark quote for alleviating malaria fevers that motivated the research: "A handful of Qinghao immersed in two liters of water, wring out the juice and drink it all" (青蒿一握，以水二升渍，绞取汁，尽服之). Tu experimented with lower heat during extraction using water, ethanol, and ether ester to arrive at dihydroartemisinin, which proved to be a clinically effective compound in curing malaria.

Due to the relentless efforts of Tu and her research team over several decades, more than 200 million malaria patients have received artemisinin or artemisinin combination therapies. In 2015, Youyou Tu was awarded the Nobel Prize in Medicine with William C. Campbell and Satoshi Ōmura; this was the first time a Chinese scientist had won the coveted award.

Perspective

TCM originated at a time when there were no sophisticated instruments as there are today in medicine and science. Regardless, TCM evolved as a highly advanced system able to identify energetic patterns within the human body aligned with the energetic properties of nature.

A striking relationship between TCM and clinical pharmacology has been established through numerous international databases linking a myriad of herbal, phytochemical, molecular, cellular, genetic, and chemical components into a warehouse of possibilities for advanced pharmaceutical and nutraceutical development. A clear distinction remains, however, in the cultural identity of patents and intellectual property ownership of western pharmaceutical enterprises and the eastern mentality of a shared natural resource. TCM has never claimed nature as its private domain and is eager to share the stage with its global counterparts for the betterment of healthcare for humanity.

A global spotlight has been placed on TCM with the 2019 adoption of the ICD-11 and the endorsement of TCM as a global healthcare option. TCM practitioners, like western medical practitioners are concerned about evidence-based knowledge and quality standardization of herbal medicines. TCM is the focus of increasing research at prestigious medical and scientific institutes, and the results are statistically remarkable.

From reviewing the PubMed Central (PMC) (National Library of Medicine and the National Institutes of Health), as of March 2020, 20,600+ studies were noted on TCM, 32,000+ on acupuncture, and 30,900+ on Chinese herbal medicine published in peer-reviewed journals throughout the world.

This chapter has presented TCM with evidence-based research in relation to clinical pharmacology, including addressing viable treatment options for epidemic levels of cancer and

infectious disease such as SARS and COVID-19. Equally important is TCM's unique ability to offer new ways of combating the rapidly progressing culture of drug-resistant bacteria and viruses, which is a result of rampant abuse of antibiotics over recent decades. It also provides ways to strengthen the body and increase innate immunity with TCM diet and nutrition, acupuncture, acupressure, Chinese herbal medicine, lifestyle changes and Qigong/Taichi exercises, which is ultimately the best prophylactic medicine to prevent diseases and infections.

Both TCM and clinical pharmacology are now at a vital crossroad with international medical and scientific proposals on herbal production and standardization, as well as advances in network pharmacology, phytopharmacology, systems biology, data mining, and informatics, among others. By joining forces, TCM and clinical pharmacology can accomplish together more than either can do on their own, providing preventive and holistic approaches with the goal of giving the best health care possible to people around the world.

Acknowledgments The author wishes to extend her most sincere and heartfelt appreciation to Contributing Editor, Lorna Zilba, MA and the following persons who have graciously shared their time and expertise in reviewing this chapter and offering valuable commentary:
John K. Chen, PhD, Pharm D., OMD, LAc
Elizabeth Chen Christenson, MD, LAc
Edward B. Christenson, MD, Dipl.ABA
Cheryl Schwartz, DVM
Robert M. Grant, MA

References and Further Reading

Afshin A et al (2019) Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 393:1958–1972

Appari M et al (2014) Sulforaphane, quercetin and catechins complement each other in elimination of advanced pancreatic cancer by miR-let-7 induction and K-ras inhibition. Int J Oncol 45:1391–1400

Bensky D (2015) Chinese herbal medicine: materia medica, 3rd edn. Eastland Press, Seattle

Boh B et al (2007) Ganoderma lucidum and its pharmaceutically active compounds. Biotechnol Annu Rev 13:265–301

Chen CJ, et al. (2008) Toona sinensis Roem tender leaf extract inhibits SARS coronavirus replication, J Ethnopharmacol Oct 30;120(1):108–11

Chen CY (2011) TCM database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. PLoS One 6:e15939. https://doi.org/10.1371/journal.pone.0015939

Chen JK, Chen TT (2004) Chinese medical herbology & pharmacology. Art of Medicine Press, City of Industry

Chen, JK (2020) How COVID-19 (2019-nCoV) is Currently Treated in China with TCM, https://www.elotus.org/article/how-covid-19-2019-ncov-currently-treated-china-tcm

Cragg GM, Newman DJ, Snader KM (1997) Natural products in drug discovery and development. J Nat Prod 60:52–60

Dillon J (2007) From the potion to the pill. Yale Medicine. https://medicine.yale.edu/news/yale-medicine-magazine/from-the-potion-to-the-pill/

Duke-Margolis, Center for Health Policy, FDA (2018) Identifying the root causes of drug shortages and finding enduring solutions. Duke University and FDA Presentation, Washington, DC. https://healthpolicy.duke.edu/sites/default/files/atoms/files/dukefda_drug_shortages_presentation_slides__0.pdf

Eban K (2019) Bottle of lies: the inside story of the generic drug boom. Harper Collins, New York

Ekor M (2014) The growing use of herbal medicines: issues relating to adverse reactions and challenges in monitoring safety. Front Pharmacol 4:177

Fabricant DS, Farnsworth NR (2001) The value of plants used in traditional medicine for drug discovery. Environ Health Perspect 109(Suppl 1):69–75

Fan X et al (2011) Network toxicology and its application to traditional Chinese medicine. Zhongguo Zhong Yao Za Zhi 36:2920–2922

Fan AY et al (2017) Acupuncture's role in solving the opioid epidemic: evidence, cost-effectiveness, and care availability for acupuncture as a primary, non-pharmacologic method for pain relief and management. J Integr Med 15(6):411–425. https://doi.org/10.1016/S2095-4964(17)60378-9

Hamburger M, Hostettmann K (1989) Analytical aspects of drugs of natural origin. J Pharm Biomed Anal 7:1337–1349

He LS et al (2016) Administration of traditional chinese blood circulation activating drugs for microvascular complications in patients with type 2 diabetes mellitus. J Doabetes Res 2016:1081657

Ideker T et al (2001) A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet 2:343–372

Jia DD (1979) Concise history of Chinese medicine (Zhong Guo Yi Xue Shi Lue). Shanxi People's Press, Taoyuan

Jiang WY (2005) Therapeutic wisdom in traditional Chinese medicine: a perspective from modern science. Trends Pharmacol Sci 26:558–563

Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–715

Lei X et al (2015) Liver damage associated with *Polygonum multiflorum* Thunb.: a systematic review of case reports and case series. Evid Based Complement Alternat Med 2015:459749. Epub 2015 Jan 12. Review

Li S, Zhang B (2013) Traditional Chinese medicine network pharmacology: theory, methodology and application. Chin J Nat Med 11:110–120

Li S et al (2011) Network target for screening synergistic drug combinations with application to traditional Chinese medicine. BMC Syst Biol 5 (Suppl 1):S10

Li QY et al (2013) Network pharmacology: a new approach for Chinese herbal medicine research. Evid Based Complement Alternat Med 2013:1–9

Li S et al (2014) Network pharmacology in traditional Chinese medicine. Evid Based Complement Alternat Med 2014:138460

Liang X et al (2014) A novel network pharmacology approach to analyse traditional herbal formulae: the Liu-Wei-Di-Huang pill as a case study. Mol BioSyst 10:1014–1022

Light D (2014) New prescription drugs: a major health risk with few offsetting advantages new prescription drugs: Harvard-Edmond J. Safra Center for Ethics. https://ethics.harvard.edu/blog/new-prescription-drugs-major-health-risk-few-offsetting-advantages

Liu XF, et al (2010) PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach, Nucleic Acids Research, Nucleic Acids Research, 38(suppl_2):W609–W614.

Liu L et al (2014) Triptolide reverses hypoxia-induced epithelial-mesenchymal transition and stem-like features in pancreatic cancer by NF-κB downregulation. Int J Cancer 134:2489–2503

Ma Y et al (2019) Prevention and treatment of infectious diseases by traditional Chinese medicine: a commentary. https://onlinelibrary.wiley.com. 24 May 2019

Maciocia G (2015) The foundations of Chinese medicine, 3rd edn. Hardcover, Churchill Livingstone, London. ISBN 9780702052163

McCulloch M, Marshall AC, Nachat A (2018) Modern thoracic oncology, vol 1. World Scientific, New Jersey, pp 261–267

McDonald J, Janz S (2017) Acupuncture evidence project: a comparative literature review. Australian Acupuncture and Chinese Medicine Association, Brisbane, p 181. http://www.acupuncture.org.au/OURSERVICES/Publications/AcupunctureEvidenceProject.aspx

Rates SMK (2001) Plants as source of drugs. Toxicon 39:603–613

Sperber G, Flaws B (2007) Integrated pharmacology: combining modern pharmacology with Chinese medicine. Blue Poppy Press, Boulder

Swinney DC, Anthony J (2011) How were new medicines discovered? Nat Rev Drug Discov 10(7):507–519

Veeresham C (2012) Natural products derived from plants as a source of drugs. J Adv Pharm Technol Res 3 (4):200–201

WHO (2002) Acupuncture: review and analysis of reports on controlled clinical trials. WHO, Geneva

WHO (2004) SARS: Clinical trials on treatment using a combination of Traditional Chinese Medicine and Western Medicine. http://apps.who.int/medicinedocs/en/d/Js6170e/4.html

WHO (2014) WHO traditional medicine strategy 2014–2023. World Health Organization. https://www.who.int/medicines/publications/traditional/trm_strategy14_23/en/

WHO (2018) Newsroom health topics. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/cancer

Williamson E et al (1996) Selection, preparation and pharmacological evaluation of plant material. Wiley, Chichester

Wright CW et al (2010) Ancient Chinese methods are remarkably effective for the preparation of artemisinin-rich extracts of Qing Hao with potent antimalarial activity. Molecules 15(2):804–812

Xiang Y et al (2019) Traditional Chinese medicine as a cancer treatment: modern perspectives of ancient but advanced science. Cancer Med 8:1958–1975. Published by Wiley

Yi F et al (2018) In silico approach in reveal traditional medicine plants pharmacological material basis. Chin Med 13:33

Zeng ZP Jiang JG (2010) Analysis of the adverse reactions induced by natural product-derived drugs. Br J Pharmacol 159(7):1374–1391

Zhang R et al (2017) TCM-Mesh: the database and analytical system for network pharmacology analysis for TCM preparations. Sci Rep 7:2821

Zhang R et al (2019) Network pharmacology databases for traditional Chinese medicine: review and assessment. Front Pharmacol 10:123

Zhou X et al (2019) Current status and major challenges to the safety and efficacy presented by Chinese Herbal Medicine. Medicines (Basel) 6(1):1

Pharmacodynamic Evaluation: Herbal Medicine

24

Gulam Mohammed Husain, Mohammad Ahmed Khan, Mohd Urooj, and Munawwar Husain Kazmi

Contents

Abstract

Herbal medicines are the primary therapy available to large segment of population across the globe. Globalization of herbal medicines has expanded their market to industrialized countries. Regulatory agencies have also extended their attention towards herbal medicines, although the majority of markets still remain largely unregulated. Further, important clinical information related to efficacy, effectiveness, dosage, adverse effects, and contraindications of herbal medicine needs to be generated to bring them into mainstream healthcare. Research into molecular effects and clinical efficacy of the numerous herbs are ongoing. Pharmacodynamic evaluation of herbal medi-

G. M. Husain (✉) · M. A. Khan · M. Urooj
Pharmacology Research Laboratory, Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India
e-mail: gmhusain@gmail.com; khan.ahmed1511@gmail.com; qaziurooj@gmail.com

M. H. Kazmi
Department of Ilmul Advia (Unani Pharmacology), Central Research Institute of Unani Medicine (Under CCRUM), Hyderabad, India
e-mail: drmunawwar@yahoo.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_52

cine through clinical studies can provide scientific evidence on benefit/risk against a particular disease. However, there are numerous challenges in conducting efficacy studies on herbal medicines. The conduct of placebo controlled clinical trials of herbal medicine poses ethical issues. It is difficult and sometimes not feasible to mimic the organoleptic properties of herbal medicines while using placebo arm. Randomization of trial subjects may be against the principles of practice of systems of medicines based on holistic approach of treatment. Varying doses mentioned in literature also makes it difficult to optimize clinical dose for testing in humans. Lack of stringent quality control of formulation aspects can lead to varying outcomes in clinical studies. This chapter discusses the ethical aspects, design consideration, quality issue, and pharmacokinetic and bioanalytical challenges in pharmacodynamic evaluation of herbal medicines in clinical setting.

Background

Globalization of Herbal Medicines

Plants and herbs have been used by mankind for ages and are now being used globally for the treatment and prevention of disease. Plants also contribute to a number of clinically used modern drugs comprising of approximately 25% of drugs prescribed globally. However, plants are more commonly consumed as herbs and herbal extracts often under categories of drugs (as herbal medicines and traditional medicines) and food (as dietary supplements) (Xutian et al. 2009; USFDA 2016). As per WHO, herbal medicines include herbs, herbal materials, herbal preparations, and finished herbal products that contain, as active ingredients, parts of plants, or other plant materials, or combinations. Traditional herbal medicines (or traditional medicines) are the herbal medicines which have undergone no or minimal processing and have been used as per regional or local healing practices (Tilburt and Kaptchuk 2008). Some herbs having health-promoting effects are also marketed as such in crude form.

Herbal medicines are used both as primary treatment as well as adjunct to conventional drugs by different populations. Especially in some developing countries, traditional system of medicines is the only system of treatment available to a large population. According to WHO, 90% population in Africa, 70% in India are dependent on herbal medicines for their healthcare needs (WHO 2005). Traditional medicine units are present in more than 90% general hospitals of China. Herbal medicine in Germany is part of one of the five main elements of naturopathy (also called as Kneipp therapies). The percentage of German population using herbal medicines increased from 52% in 1970 to 70% in 2010. In 2011, Germany spent one billion Euro on herbal medicines which is approximately 20% of the total expenditure for over the counter (OTC) drugs (Joos et al. 2012).

Recent decades have seen the globalization of herbal medicines leading to expanded use of ethnobotanicals in the industrialized countries. According to National Center for Complementary and Alternative Medicine in the United States, herbal medicines are the most commonly used (18.9%) alternative medicine (Barnes et al. 2008). In two different surveys, 38% of US adults have reported use of herbal medicine whereas 40% of Hong Kong population have shown faith in traditional herbal medicines (Barnes et al. 2008; Ernst et al. 2005; Chan et al. 2003). As of 2008, the global annual turnover of herbal medicines had reached nearly US$ 80 billion (Anonymous 2017) with millions of US dollars of industry investment being made in medicinal herbs with promising potential (Zamiska 2006). The global herbal medicine market was expected to reach $107 billion by the year 2017 (Anonymous 2012). However, this investment still remains small compared to huge size of global pharmaceutical market. National health authorities are also increasing their attention towards herbal medicines due to the excessive reliance of less developed countries on herbal medicines as well as due to expanding base in developed countries.

Irrespective of the reason for their choice, the consumers have right to receive herbal

medicines that are safe and effective. Further, scientific studies need to be carried out to generate science based information on efficacy, effectiveness, dosage, adverse effects, and contraindications. Research into molecular effects and clinical efficacy of the numerous herbs are ongoing. Pharmacodynamic evaluation of herbal medicine through clinical studies is an important step of this process which can provide scientific evidence on benefit/risk of any herbal medicine in the treatment of disease.

Regulatory Challenges

The diverse countries practice herbal medicines due to historical reasons as well as different holistic approaches. This makes their regulation a difficult task. Further, regulating large number of herbs is also a challenge. The WHO survey of 129 countries identified lack of research data, herbal medicine control mechanisms, safety monitoring, and methods for evaluation of safety and efficacy as some of the challenges in regulation of herbal medicines. Regulatory authorities of various countries are yet to develop strict guidelines for robust assessment of safety and efficacy. Notwithstanding, many markets of herbal medicines have remained largely unregulated leading to poor quality control during sourcing of raw materials, production, distribution, and sale. In addition, there exists an overlap in definitions of plant-based products among different countries. The same plant or plant-based product may be defined as a dietary supplement, food, or herbal medicine by different regulators. For example, EMEA defines herbal drugs as the whole, fragmented or cut, plants, parts of plants, algae, fungi, lichen in an unprocessed state usually in dried form or afresh. Herbal drug preparations are the herbal drugs which are subjected to various treatments and prepared as tinctures, extracts, essential oils, expressed juice, or process exudates. In contrary, USFDA defines all herbal products as dietary supplements and therefore as foods. Whereas, USFDA defines a Botanical Drug Product as derived from plants, algae, or macroscopic fungi and prepared from botanical raw materials by one or more of the processes such as pulverization, decoction, expression, aqueous extraction, ethanolic extraction, or other similar process, intended for use as a drug. Although, safety of herbal medicines remains a global concern, a global survey conducted by WHO revealed that till 2003, 63% of 191 member states did not have laws and regulations for herbal medicines and herbal medicines were sold as over-the-counter (OTC) products in 68% of member states. In relation to quality, only 24% countries had national pharmacopoeia of herbal drugs (WHO 2005, 2008).

The registration requirements for herbal drugs also vary from country to country. For example, herbal drugs with sufficient evidence of the medicinal use of the product throughout a period of at least 30 years, including at least 15 years in the community have been exempted from testing and trials for safety and efficacy by the Committee on Herbal Medicinal Products (HMPC) within the European Medicines Agency (EMEA). In India, Department of Ayurveda, Yoga & Naturopathy, Unani, Siddha and Homoeopathy (AYUSH) is the main regulator for herbal medicines particularly traditional system products majority of which do not require stringent safety and efficacy studies as governed by The Drugs & Cosmetics Act.

USFDA registers herbal products used in complementary and alternative medicine which, based on its use, may be regulated as food (including food additives and dietary supplements), drug, cosmetic, or biological product. Further, in such case, no exemption in regulation is granted to herbal medicines. For example, a dietary product used for promoting optimal health may be subject to the requirements for foods including all the safety and quality testing as per FDA regulations. Whereas, if the same is to be used as a treatment regimen for a particular disease, it will be subject to regulation as a drug. The product may also be categorized as “new drug” under the Act and premarket review and approval by FDA may be needed which includes evaluation of safety and effectiveness under the prescribed conditions as suggested in the labeling.

Ethical Aspects in Conducting Human Studies

Clinical trial is the foremost part of drug development process involving human subjects. Clinical evaluation of herbal medicine should be subjected to alike ethical principles as per national and international guidelines and regulations implicated for conventional medicine. Some of the prominent code of ethics and set of laws which direct ethical clinical research include Nuremberg Code (1947); Belmont Report (1979); US Common Rule (1991); Declaration of Helsinki (2000); and Council for International Organizations of Medical Sciences (2002).

The major ethical shortcoming of a clinical trial is the population which will gain is not the same which is taking the burden of risk by participation. Thus, all guidelines and regulation are stern to protection of rights, safety, and well-being of clinical trial participants. Clinical trial for herbal medicine must be conducted as per defined protocol approved by ethics committee as well as guidelines of regulatory agency. Below mentioned points in Fig. 1 are ethical concern needed to be addressed for clinical research:

Risk and benefit ratio of trial drug should be identified and assessed to minimize the risk and maximize the benefits to participant. However, safety of the subjects on herbal product trial is often overlooked, because it is believed that herbal medicine is safe as being used clinically from prolonged time and sometimes from centuries. This may cause error in risk assessment leading to serious consequences. For instance, the whole ephedra plant traditionally used to treat bronchial asthma possess toxicity potential including cardiac arrhythmia, myocardial infarction, and stroke (Koonrungsesomboon and Karbwang 2016).

Another ethical issue is using placebo arm in controlled clinical trial of herbal medicine. Depriving a patient of active treatment is unethical especially in critical disease like diabetes, cancer, etc. It is therefore recommended to conduct an add-on trial for herbal medicines which may be useful to prevent adverse effect of modern treatment until system is not able to standardize and optimize the dose, regimen, and duration of treatment. Herbal medicine trial involving geriatric population also raises an issue of ethics, due to concern of age and impaired organ function which may alter the pharmacokinetics and pharmacodynamic of trial drug. The dose which is considered to be safe in general population may become toxic in elderly population (Routledge et al. 2004). Ethical standards must be kept into account in line with national and international guidelines for herbal drug clinical trials without compromising the rights, safety, and well-being of trial participants.

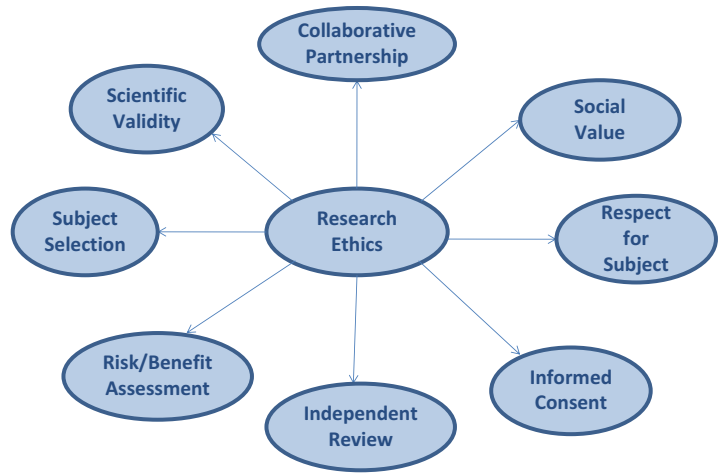


Fig. 1 Ethical consideration for clinical trial on herbal medicine

Design of Clinical Trials for Herbal Medicines

Control and Randomization

Control

A control or control group may be defined as a group of clinical trial participants which do not receive the drug or treatment being investigated as part of the trial. The control group is essential in controlled clinical trial to ascertain the effectiveness of study drug by nullifying the effect of confounding factors. There are four major controls that may be utilized while conducting clinical pharmacodynamic evaluation of a herbal medicine viz. placebo control; active treatment control; no treatment control; and different doses or regimen as control.

The foremost challenge with trial of herbal medicine is selection of appropriate control. Selection of placebo control puts forward the challenge of mimicking the exact taste, color, odor, texture, and physical state as treatment. For example, preparation of placebo for decoction, complex natural product having strong aroma and odor-like ginger or garlic, etc. The second issue is opting for placebo instead of availability of an active treatment, which is ethically unjustified to let the patient without treatment when standard therapy is available.

Similarly, in order to compare the efficacy of a test drug, an established and widely accepted drug must be chosen as active (positive) control. The active control trial has merit of ethical and practical concern. The primary choice for active control drug selection is that the drug should have indistinguishable mechanism of action or similar action as test drug. If it is not available, then drug of same therapeutic class with same indication has to be chosen. Herbal drugs may be used as an active control wherever feasible such as silymarin for hepatoprotective activity. Selection of active control is also governed by the nature and complexity of the herbal drug, for example, enriched fraction or whole extract; crude preparation or isolated active ingredients.

The selection of active control group from conventional medicine for comparison of efficacy of herbal drugs is debatable as it definitely influences comparability of groups. The outcome for both active control and test treatment could not be equally comparable due to lack of similarity in mechanism of action, dose optimization, variability in dosage regimen, and differences in formulation. Choosing no treatment as control will reduce the power of study and at the same time will be ethically unjustified. This control generally applied in experimentation of animal and clinical trials of chronic disease.

The way forward for placebo development is to move towards redesigning of dosage forms without compromising the holistic approach. The capsule as dosage form might solve the problem for herbal placebo with certain limitations like the formulation of capsule for herbal medicines with large doses. Further, in case of herbal drug clinical trial for traditional medicines, practitioner or government research organization must try to adopt robust research methodology in order to achieve the best possible active control. One possible approach is planning and conducting a pilot study with low sample size, to generate the preliminary data in support of efficacy of herbal drug. A confirmatory trial may further validate the findings in a larger, heterogeneous population.

Randomization

Randomization is a method by which study participants are assigned to a treatment group based on chance alone. Randomization generates comparable group of treatment by eliminating possible source of bias. Randomization ensures that each patient under trial have equal chance to receive any treatment which is achieved by concealment of allocation where neither patient nor researcher is aware about the treatment assignment to trial subject (Suresh 2011).

There are some statistical methods like ANCOVA (analysis of covariance) which are used to adjust imbalance among covariates between trial arms during analysis stage. There may be variation in adjustment required for each trial group, however, the major limitation with ANCOVA is that this method uses the average slope among groups to adjust the outcome

variable. Therefore, the best method to achieve balancing among covariates is to apply randomization at the stage of designing of clinical trial instead of applying statistical tool after data collection (Kabisch et al. 2011).

The randomized clinical trial (RCTs) is considered as gold standard for the assessment of treatment effect. RCTs are the highest power clinical designs to study the safety and effectiveness of new treatment or intervention and mandatory for obtaining approval by government regulatory bodies (Bothwell and Podolsky 2016). RCTs are mainly trials that entail one arm as control group where experimental group is compared with control group (see section "Control" for different types of control). In eighteenth and early nineteenth century, the trials conducted were poorly controlled to confirm the effectiveness of orthodox wide range medicine. Major development of scientific methodology in late nineteenth and twentieth century raised the demand of rigorously conducted controlled trials to prove safety and effectiveness of new medication.

There have been adequate studies conducted on herbal drugs, but well-controlled clinical trials are still lacking to prove safety and efficacy, for example, the outcome of clinical trial on extract of *Ginkgo biloba* used to treat CNS and cardiovascular disorders and *Hypericum perforatum* (St. John's wort an antidepressant) prove that both drugs are quite safe and effective. However, their clinical efficacy still needs to be supported by further well-designed controlled clinical trials. The issue of relevance of randomized clinical trial has also been raised by the medical practitioner prescribing herbal medicines since they argue that they are treating soul, mind, and body at the same time with holistic approach. These physicians are integral part of therapy and treat the patients on individual basis and therefore the outcome of the treatment cannot be generalized. Although RCTs on herbal medicines are difficult to execute, but these are the need of the hour. They require a careful planning and a team effort for establishing the efficacy of herbal medicines. Recently, the new approach of partial randomization has been proposed where first the patients are given a choice to decide the trial arm preference and allocated to the arm based on choice followed by randomization of remaining patients who do not opt for any preference. Similarly, according to the design proposed by Harvard Statistician Zelen, patients are first randomized to control or treatment arm prior to seeking informed consent followed by conditional sought of consent. As per Zelen's design, consent may not be sought from the patient receiving standard care except for privacy reasons.

The double screening recruitment model is an alternative recommended for clinical trial on herbal medicine with traditional origin where clinical trial subjects are first screened using modern diagnosis techniques and then categorized based on traditional classification system. Here the patients get equal opportunity to receive experimental treatment or standard treatment. Crossover design is possibly a better choice in case of stable disease and therapy that have short-term effects. Crossover design minimizes the variation as a patient becomes its own control, i.e., in the first phase patient will receive experimental drug and after washout period same patient will receive the placebo. Design adaptive allocation is also a suggested technique over randomization which ensures better balance between measured and unmeasured outcomes. This method also enhances statistical power.

Optimization of Clinical Dose

There are two vital transitions in clinical trials: first transition is from preclinical pharmacology study to first in human use, i.e., phase I clinical trial, and second transition is from phase IIa to IIb/III where large number of patients are exposed to the drug (Gobburu and Gopalakrishnan 2006). Both transition phases have significant influence on selection of dose levels. Selection of suboptimal dose level will impact the project timeline while higher than therapeutic dose level may produce adverse effects. In light of above facts, the major concern for herbal medicine trial is dose optimization as varied clinical doses are mentioned in the literature and there is lack of

consensus regarding the dose regimen. The objective to bring herbal medicine parallel to conventional medicine in healthcare system can only be achieved by developing proof of concepts on scientific grounds.

Reverse pharmacology is the technique to prove the experience based claims for traditional herbal agents by applying modern scientific methods. The major challenge while designing clinical trial for herbal medicines is selection of dose levels and regimen. Three scenarios come into play while applying pharmacodynamic principle to evaluate herbal medicines (Parveen et al. 2015).

(a) Herbal agents whose efficacy is demonstrated, active constituent is known, and their doses are more or less established
(b) Herbal agents with putative efficacy which needs to be demonstrated and active principle needs to be standardized
(c) Herbal agents with uncertain efficacy but used traditionally since centuries

The solution to these challenges is rigorous quality control of the herbal drug and standardization using as many as possible analytical markers. Modern analytical techniques such as hyphenated techniques like LC-MS, GC-MS needs to be adopted for standardization of herbal medicine. Formulation of herbal medicine must ensure minimum batch to batch variation. Application of pharmacodynamic principle is justified only when the identity and composition of the herbal drug is well characterized. Another challenge is pharmacokinetic profiling of herbal drugs whether it is a single or multiple constituent product. Identification and quantification of the principle biomarker(s) in biological fluids require high level of expertise.

Dose optimization can possibly be achieved either by conducting efficacy study in animals at initial levels and then by translating results for first in human studies or directly conducting phase I trial. Phase I trial must be undertaken to identify maximum tolerated dose and measurement of drug activity. Phase I trial with dose escalation method as mentioned in Fig. 2 can be conducted with low sample size. The trial subjects should be randomly allocated to three different groups with low dose group, middle range dose group, and high dose group. There should be no difference in inclusion criterion for all groups. The safety and efficacy is the primary objective for phase I trial which will be assured by changes in standardized efficacy endpoints at different dose levels. If we are able to find out maximum tolerated dose level and efficacious and safe dose level for these herbal agents then randomized controlled trials could be planned. It must be kept in mind that it is generally claimed that herbal drugs do not cause any adverse effect; however, phase I trial must be carried out under extensive medical supervision as per ICH-GCP guideline without compromising patient safety.

Selection of Study Duration

In conventional medicine dosage regimen is strictly followed to maintain uniformity; however, there is wide variation in dose and treatment duration in case of herbal or traditional medicine. Herbal medicine, either single constituent or polyherbal formulation, are frequently used as self-medication by the patients with varied duration until relieved symptomatically. Similarly, traditional practitioners prescribe same drug for same indication for different duration based on the symptomatic relief. They claim that disease and treatment is individualized to the patients in case of alternative medicine. However, for any new herbal medicine, clinical studies are needed to be designed for optimization of dose, dosage regimen, and treatment duration. This objective can be achieved by designing a pilot study with low sample size, standardized disease specific inclusion-exclusion criteria and clinical efficacy end points. Optimized dose should be given to patients with stringent follow up for uniform duration decided as per available literature and gathered experience. The follow up should be done till the patients become healthy or relieved symptomatically as well as based on the achievement of previously set objective parameters. Relapse of disease must

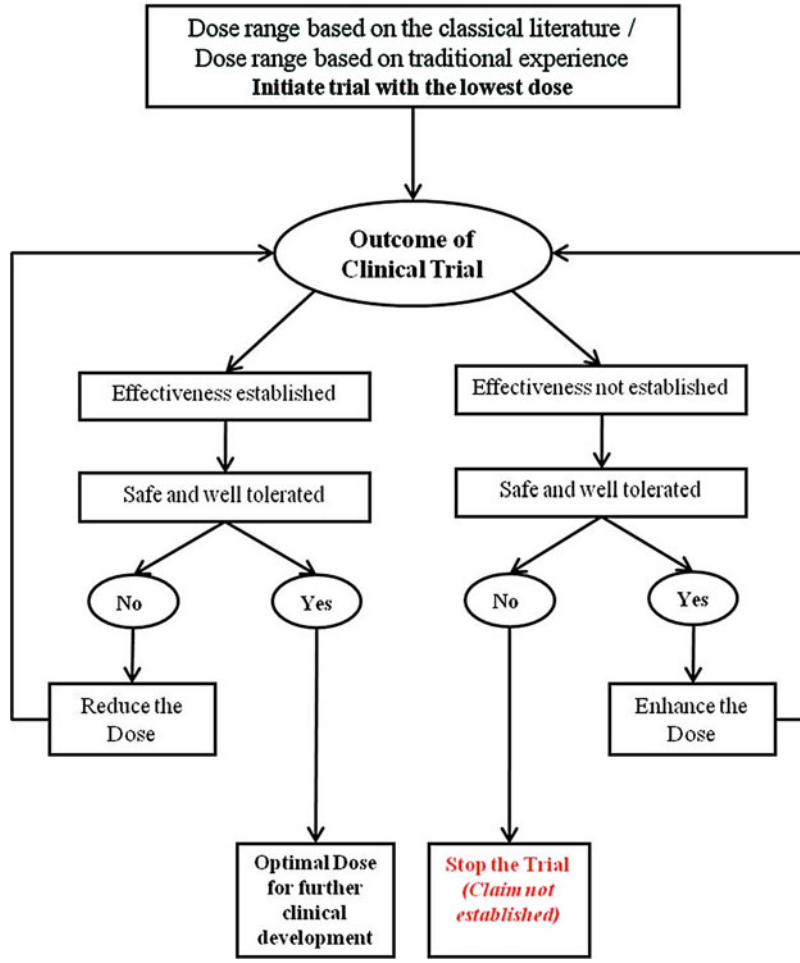


Fig. 2 Dose escalating study

also be kept in mind and duration of follow up may be extended accordingly.

Efficacy and Safety Assessment

Efficacy or effectiveness of herbal drug could be confirmed based on the clinical end points. Clinical end points (efficacy end points) are used to measure the effect of a drug by assessment of clinical events. The end points are classified into three categories viz., primary, secondary, and exploratory. Primary end point becomes main objective of clinical trial for which trial subjects are randomized and study is powered, for instance, chest pain becomes primary end point for trial of drug preventing heart attack. Primary end point is the main outcome of the study to address the research question and main hypothesis.

Secondary end points are additional clinical outcome/events for which neither subjects are randomized nor is the study powered for. Secondary end points do not possess similar statistical authority as primary end points, for example, drug for osteoporosis using fracture as primary end point and improvement in bone density as secondary end point. All other end points recorded in a study are known as exploratory end points (USFDA 2017).

The practitioners of traditional herbal medicines use different methods of diagnosis compared to conventional practitioners. Herbal practitioner uses subjective end points generally

focused on relieving the symptoms and most of the time relates a disease with under-functioning of a biological system for instance the cause given by traditional herbal practitioners for arthritis is accumulation of waste products in joints for which they prescribe the combination of herbs such as diuretics, choleretics, laxatives, and additional anti-inflammatory herbs (Vickers and Zollman 1999). Conversely, modern diagnosis is based on hard end points which can be accurately measured.

The scenario is rapidly changing now and therefore, randomized controlled trials are needed to be conducted to optimize herbal medicine outcome in line with modern clinical end points. For example, herbal medicines have promising outcome in case of renal and hepatic diseases but there is a need to confirm the results based on the biochemical and molecular evidences of renal and liver function.

Herbal drugs are perceived of being safe and devoid of any side effects because these medicines have long history of use. However, the concept of safety is relative and all medicinal products are associated with some risk in addition to their beneficial effects. The benefit and risks are always addressed with caution prior to marketing authorization of pharmaceutical product. Randomized controlled clinical trials are the best way to provide evidence for safety of medicinal product before marketing approval. There is a need of extensive clinical research to ascertain the safety of herbal medicine by conducting randomized controlled clinical trials. Further, there are some aspects of safety which could not be addressed by clinical trials hence, prior to clinical studies, nonclinical assays like preclinical carcinogenic study, developmental toxicity study, reproductive toxicity study, and repeated dose toxicity studies must be performed. There is a concern of safety in instances where herbal drugs are to be coadministered with modern medicine or with another herb or nutrients. Therefore, drug interaction studies are recommended in such cases to ensure safety. Considering the huge exposure potential of herbal drugs, researchers must also try to generate the data for post marketing vigilance for safety of herbal drugs.

Chemistry Manufacturing and Control (CMC) of Herbal Medicines

Materials of herbal origin used in herbal medicines are uniquely distinct from conventional modern medicinal products. Unlike conventional pharmaceuticals which are manufactured using synthetic molecules and reproducible processes, ingredients of herbal medicine are obtained from diverse sources. Thus, it becomes difficult to ascertain the past conditions to which these were subjected to. Further, one particular herbal material obtained from more than one cultivation source varies in quality and composition. Moreover, there are seasonal variations in the constituents of plants from the same origin. The heterogenous nature of botanicals also makes it difficult to identify active constituent, ascertain uniform quality and rule out contamination with other plant materials. Due to these reasons, the conventional methodologies cannot be used for manufacture as well as quality control of herbal medicines. Since, the herbal medicinal products are affected by processing and production techniques, quality assurance has become an important prerequisite for conducting clinical trials on herbal medicines.

Standardization and Quality Control

Primary objective of standardization and quality control is to ensure that the product batch/batches are consistent in terms of their composition, strength, and pharmacological effect. Quality control of herbal medicines involves a battery of tests including identification, assay of chemical constituents, inorganic impurities (including toxic metals), microbial load, pesticides, etc. (Ong 2004). “Phytoequivalence” is used to ensure consistency of herbal medicine by comparing the chromatographic fingerprint with the profile of a reference product (Liang et al. 2004). In summary,

quality control of herbal medicine requires a totality of evidence approach.

USFDA recommends that the quality control of botanical drugs should extend to raw material(s) and may additionally require biological assays and data on outcome variations from multiple batch clinical study. Following three have been identified by USFDA for quality control of botanical drugs (USFDA 2016):

1. Botanical raw material control
2. Quality control by chemical tests
3. Biological assay

1. *Raw Material*
   Raw material from at least three cultivation sites should be collected for the assessment of quality. The same material should be used as representative materials for evaluating therapeutic consistency and the same should be used for the production of raw material batches for multiple batch phase III studies. There may be need for additional characterization of raw material using spectroscopy/chromatography, DNA fingerprinting, etc.
2. *Herbal Medicinal Product*
   Complete pharmaceutical development is required for herbal medicinal products to avoid any changes in either raw material or manufacturing processes for clinical studies. Post clinical development, robust manufacturing is required to demonstrate that the herbal product to be marketed is equivalent to the one used in clinical study. The key factor in ensuring consistency as well as quality of batches is the reproducibility of the production process. This requires validation of all process including qualification of equipments and establishment of a formal change control system to identify effects of changes on the quality.
3. *Bioassays*
   While the raw material and drug product control are essential for establishing identity and ensuring quality, in some instances correlations between the quality and biological activity may be required using bioassay. The bioassay can ensure therapeutic consistency. The bioassay should preferably be closely related to the test drug's proposed mechanism of action.

Pharmacokinetic and Bioanalytical Challenges

Herbal medicines have become popular as complementary therapy especially against a number of chronic conditions like cancer and metabolic diseases (Hollander and Mechanick 2008). The mechanism of action of herbal medicine is different from the modern drugs having single chemical moiety. It involves a "network" approach, wherein multiple compounds act together through interaction with multiple *in vivo* targets with interdependent actions leading to optimal effect (Chan 1995). The main objective of carrying out pharmacokinetics (PK) study is to understand the *in vivo* process that a drug undergoes. PK is an integral part of drug discovery process (Wu et al. 2000). Thus, the extensive PK study of herbal medicines can play critical role in improving the understanding of clinical effects and generation of evidence for design of a reasonable dosage regimen. However, due to lack of proper understanding of PK of herbal medicine (especially multicomponent and traditional system medicine), currently the design of dosage regimen of almost all such products is based on ancient empirical therapy rather than *in vivo* profiles. This rationality in clinical practice has not been accepted by the modern medical system. In addition to the design of dosage regiment, PK studies of herbal medicines can also aid in developing better understanding of their interactions with other exogenous systems. Interactions, between phytoconstituents and prescription drugs have recently been in focus, especially due to increasing awareness among medical practitioners about the widespread adverse effects of herbal treatments undisclosed by the patients (Mukherjee et al. 2015). The Committee on Herbal Medicinal Products (HMPC) has mandated that interaction of herbal medicine with other medicinal products (EMEA 2007a, b) must be evaluated. Similarly,

FDA guidance for botanical drug products has also described that assessment of interactions between herbal medicines and other commonly used drugs and/or dietary supplements must be carried out (USFDA 2016).

Challenges in the Assessment of Pharmacokinetics

Generally, the pharmacokinetic principles used for development of single molecule drugs have been employed to understand the efficacy and toxicity of herbal medicines. Some of these principles include (Na 2010):

1. Assessment of rate and extent of absorption of components of herbal medicines
2. Understanding of metabolic fate of component (s) of herbal medicines
3. Elucidation of route(s) of elimination and elimination kinetics
4. Understanding herbal medicine-synthetic drug interaction

However, this strategy has not been able to achieve much success because of complexity of extracts especially in case of multicomponent mixtures. The high number and wide range of metabolites present in natural products are inextricable obstacles in the development of PK. Secondly due to lack of information on the active principle, it is difficult to select any particular compound or establish pharmacological basis for efficacy.

Further, the influence of dosage form, formulation aspect and/or type of extract of herbal medicines is also prominent on pharmacokinetics. For example, a recent open, single dose, crossover study was carried out to compare the pharmacokinetics of two *Ginkgo biloba* products from US market. Results showed that there was significant difference in all pharmacokinetic parameters between the test and reference products upon administration of equivalent doses. The reference product lead to higher plasma concentrations despite having less ginkgolides and bilobalide concentrations than test product (Kressmann 2001). Therefore, pharmaceutical equivalent products may not always be bioequivalent in case of herbal medicines. The multiple compounds present in different formulations lead to different kinds of metabolic and pharmacokinetic interactions leading to variable PK profiles.

While analyzing biological samples containing herbal medicines, one must keep in mind the extensive metabolic biotransformation that compounds may undergo. The evaluation of integral metabolism profile of herbal medicines has been an important breakthrough in predicting and explaining their pharmacokinetics, efficacy, herb-drug interactions as well as toxicity (Raskin et al. 2002; Li et al. 2009; Tang et al. 2009). The metabolism profile can reveal changes occurring in vivo due to herbal exposure. However, integral metabolism of herbal medicines is still largely unexplored, requiring exhaustive research.

Another challenge which is beyond the scope of traditional research is selection of analytical tools for PK elucidation of integral metabolism profile of complex herbal medicine present in a complex biological sample (e.g., blood, tissues, or urine). In a herbal medicine containing, for example, up to 300 compounds, the concentration of a particular compound in the single dose of a finished product can be in the lower mg range and the consequent plasma concentrations may go as low as μg to pg per liter range. This makes not only quantitative but global qualitative analysis difficult. The evaluation of integral metabolism profile requires advances bioanalytical strategy and tools having sufficient sensitivity as well as selectivity. The absence of reliable analytical tools remained a challenge for a long time. Currently available chromatography and mass spectrometry (MS) based technologies like liquid chromatography-MS (LC–MS), gas chromatography-MS (GC–MS), and capillary electrophoresis-MS (CE–MS) have become important facilitators of PK studies of herbal medicine. LC–MS or LC–MS/MS based assays have been commonly employed for analyzing multiple constituents of herbal medicines (Wang et al. 2008; Xin et al. 2011).

However, despite the technological advancement in the study of metabolites of single

component herbal medicines, fate of multicomponent herbal medicines still remains a challenge. The obstacles in the development of metabolic profile of multicomponent herbal medicines are many. Firstly, it is difficult to standardize the chemical composition of multicomponent herbal medicines or to neutralize the effect of batch to batch variation. Secondly, in some instances there is an overlap between the chemical composition of normal diet and herbal medicines which can lead to pseudo results. Thirdly, extensive microbial-mammalian cometabolism of multiple components can occur in the gut which varies from species to species. Fourthly, multicomponent products further complicate the analytical differentiation between exogenous metabolites from the endogenous ones. Lastly, there are many overlapping metabolic pathways for chemically similar components. Sometimes, such pathways may intercross each other as well.

Methods for Pharmacokinetic Evaluation

A few methods have been proposed for the pharmacokinetic evaluation of multicomponent herbal medicines. The classical strategy has been to study the pharmacokinetics of single constituent of a multicomponent agent on individual biological reactions, genes, enzymes, etc. followed by assembling the findings to develop the comprehensive profile. However, the PK of any constituent in isolation may differ significantly from its behavior in a multicomponent system due to drug-drug interactions. It is practically not possible to identify each metabolite in the global pool of metabolites and assess their effect on biochemical pathways. Further, implementation of this reductionist approach is not valid and does not capture the complex behavior from systems biology perspective (Xue and Roy 2003).

Hao et al. proposed the method for integrated PK study of multicomponent herbal medicines (originally proposed for traditional Chinese medicine) called the "AUC weighting integrated method" (Hao et al. 2009). The method includes following steps:

- Assessment of PK properties and therapeutic effects of each ingredient in multicomponent herbal medicine
- Selection of powerful ingredients with suitable PK characteristics as the PK markers
- Determination of blood drug concentration of the selected PK markers and plotting the blood drug concentration–time curves of the multiple components
- Integration of blood drug concentration–time profiles of each PK marker to represent the whole in vivo process of the multicomponent herbal medicine based on possible contribution weight of each marker to the pharmacological effect

Pharmacokinetic study of a herbal medicine preparation (Xuesaitong injection) is reported based on this scheme. Elucidation of identified five main PK markers viz. notoginsenoside R1, and ginsenosides Rg1, Re, Rb1, and Rd, and therapeutic effect of the same was carried out. The blood drug concentration–time profiles of these markers were subsequently investigated. The ratio of $AUC_{(0-\infty)}$ of each ingredient and sum of the $AUC_{(0-\infty)}$ of five markers was considered as the PK weight coefficient of each ingredient. Finally, sum of the value of the blood drug concentration X $AUC_{(0-\infty)}$ weight coefficient of each marker was regarded as the integrated blood drug concentration of the five markers at each sampling time point. This value was used to draw integrated drug concentration–time curve and estimate the integrated PK parameters as an indicator of whole PK properties of the preparation (Li et al. 2008).

Integrated PK profile of *Saussurea laniceps* have been reported using similar method (Yi et al. 2014). Although this method has been reported in more than one type of studies, there is argument that PK markers selected using this method might not comprehensively indicate therapeutic response of multicomponent herbal medicine (Liu et al. 2009). This is so mainly due to the fact that contribution "weight" of few of the multiple constituents to the overall therapeutic property is less known.

An ideal technique should not only indicate the integrated blood drug concentration profile but should also indicate the dose-dependent therapeutic effect of multicomponent herbal medicine. Such a method can help in carrying out simultaneous PK and PD evaluation. This can also help in utilization of PK profile in design of dosage regimen for herbal medicines. The currently available advanced technologies like liquid/gas chromatography coupled with mass spectrometry (LC-MS and GC-MS) as well as capillary electrophoresis coupled with mass spectrometry (CE-MS) and nuclear magnetic resonance (NMR) have made simultaneous detection of various metabolites of herbal medicines possible (Tolonen et al. 2009) through study of complete set of metabolome (Metabolomics). Metabolomics is a unique platform for poly PK studies which not only identifies and measures multiple constituents of herbal medicine *in vivo* but also characterizes any metabolic alterations upon human exposure (Sumner et al. 2007; Wang et al. 2005). Metabolomics has been recently utilized for the investigation of herbal medicine efficacy and toxicity (Xie et al. 2008; Ni et al. 2008; Chen et al. 2007). A proof-of-concept PK study of Pu-erh tea intervention has been reported based on metabolomics approach using tandem mass spectrometry (MS/MS) (Lan and Jia 2010). Metabolomics combined with statistical tools can address the challenges encountered in PK of multicomponent herbal medicines and in the assessment of their efficacy. Complete panel of dynamic pharmacokinetic profile for dosage regimens of multicomponent herbal medicines may also help in minimizing overdosing and consequently reduction of toxicity.

Acknowledgments The views expressed by the authors in this book chapter are personal and not necessarily convey the view point of Central Research Institute of Unani Medicine or CCRUM.

References

Anonymous (2012) Global herbal supplement market to reach $107 billion by 2017. Nutraceuticals World

Anonymous (2017) New trends & effective strategies for medicinal & natural products. 5th international conference and exhibition on pharmacognosy, phytochemistry & natural products. http://pharmacognosy-phytochemistry-natural-products.pharmaceuticalconferences.com/events-list/herbal-drugs-and-formulations

Barnes PM, Bloom B, Nahin RL (2008) Complementary and alternative medicine use among adults and children: United States, 2007. Natl Health Stat Rep 12:1–23

Bothwell LE, Podolsky SH (2016) The emergence of the randomized, controlled trial. N Engl J Med 375 (6):501–504. https://doi.org/10.1056/NEJMp1604635

Chan K (1995) Progress in traditional Chinese medicine. Trends Pharmacol Sci 16(6):182–187

Chan MF, Mok E, Wong YS, Tong TF, Day MC, Tang CK, Wong DH (2003) Attitudes of Hong Kong Chinese to traditional Chinese medicine and western medicine: survey and cluster analysis. Complement Ther Med 11(2):103–109

Chen X, Ye Y, Cheng H, Yin J, Sun C (2007) Comparative analysis of theanine, caffein and polyphenolic constituents in green tea, black tea and puer tea. Food Res Dev 12:046

EMEA (2007a) Procedure for the preparation of community monographs for herbal medicinal products with well-established medicinal use. European Medicine Agency, London

EMEA (2007b) Procedure for the preparation of community monographs for herbal medicinal products with well-established medicinal use. Committee on Herbal Medicinal Products (HMPC, London)

Ernst E, Schmidt K, Wider B (2005) CAM research in Britain: the last 10 years. Complement Ther Clin Pract 11(1):17–20

Gobburu JS, Gopalakrishnan M (2006) Optimal dose finding in drug development. In: Dose optimization in drug development, Drugs and the pharmaceutical sciences. Informa Healthcare, pp 195–215. https://doi.org/10.1201/9781420018585.ch9

Hao HP, Zheng CN, Wang GJ (2009) Thoughts and experimental exploration on pharmacokinetic study of herbal medicines with multiple-components and targets. Yao xue xue bao 44(3):270–275

Hollander JM, Mechanick JI (2008) Complementary and alternative medicine and the management of the metabolic syndrome. J Am Diet Assoc 108(3):495–509. https://doi.org/10.1016/j.jada.2007.12.007

Joos S, Glassen K, Musselmann B (2012) Herbal medicine in primary healthcare in Germany: the Patient's perspective. Evid Based Complement Alternat Med 2012:294638. https://doi.org/10.1155/2012/294638

Kabisch M, Ruckes C, Seibert-Grafe M, Blettner M (2011) Randomized controlled trials: part 17 of a series on evaluation of scientific publications. Dtsch Arztebl Int 108(39):663–668. https://doi.org/10.3238/arztebl.2011.0663

Koonrungsesomboon N, Karbwang J (2016) Ethical considerations in clinical research on herbal medicine for prevention of cardiovascular disease in the ageing. Phytomedicine 23(11):1090–1094. https://doi.org/10.1016/j.phymed.2015.10.017

Kressmann S (2001) Comparison of herbal medicinal products from Ginkgo biloba from the U.S. market. University of Frankfurt, Frankfurt

Lan K, Jia W (2010) An integrated metabolomics and pharmacokinetics strategy for multi-component drugs evaluation. Curr Drug Metab 11(1):105–114

Li X, Hao H, Wang G, Sun J, Liang Y, Xie L, Zhang Y, Yan B (2008) Integrated pharmacokinetic study of multiple effective components contained in total Panax notoginsenosides. Chin J Nat Med 6(5):377–381

Li Y, Duan J, Guo T, Xie W, Yan S, Li B, Zhou Y, Chen Y (2009) In vivo pharmacokinetics comparisons of icariin, emodin and psoralen from gan-kang granules and extracts of herba Epimedii, Nepal dock root, Ficus hirta yahl. J Ethnopharmacol 124(3):522–529. https://doi.org/10.1016/j.jep.2009.05.008

Liang YZ, Xie P, Chan K (2004) Quality control of herbal medicines. J Chromatogr Analyt Technol Biomed Life Sci 812(1–2):53–70. https://doi.org/10.1016/j.jchromb.2004.08.041

Liu H, Yang J, Du F, Gao X, Ma X, Huang Y, Xu F, Niu W, Wang F, Mao Y, Sun Y, Lu T, Liu C, Zhang B, Li C (2009) Absorption and disposition of ginsenosides after oral administration of Panax notoginseng extract to rats. Drug Metab Dispos 37(12):2290–2298. https://doi.org/10.1124/dmd.109.029819

Mukherjee PK, Bahadur S, Chaudhary SK, Kar A, Mukherjee K (2015) Chapter 1 – Quality related safety issue-evidence-based validation of herbal medicine farm to pharma. In: Evidence-based validation of herbal medicine. Elsevier, Boston, pp 1–28

Na DH (2010) Metabolism study of botanical drugs. Arch Pharm Res 33(12):1877–1879. https://doi.org/10.1007/s12272-010-1200-7

Ni Y, Su M, Lin J, Wang X, Qiu Y, Zhao A, Chen T, Jia W (2008) Metabolic profiling reveals disorder of amino acid metabolism in four brain regions from a rat model of chronic unpredictable mild stress. FEBS Lett 582(17):2627–2636. https://doi.org/10.1016/j.febslet.2008.06.040

Ong ES (2004) Extraction methods and chemical standardization of botanicals and herbal preparations. J Chromatog Analyt Technol Biomed Life Sci 812(1–2):23–33. https://doi.org/10.1016/j.jchromb.2004.07.041

Parveen A, Parveen B, Parveen R, Ahmad S (2015) Challenges and guidelines for clinical trial of herbal drugs. J Pharm Bioallied Sci 7(4):329–333. https://doi.org/10.4103/0975-7406.168035

Raskin I, Ribnicky DM, Komarnytsky S, Ilic N, Poulev A, Borisjuk N, Brinker A, Moreno DA, Ripoll C, Yakoby N, O'Neal JM, Cornwell T, Pastor I, Fridlender B (2002) Plants and human health in the twenty-first century. Trends Biotechnol 20(12):522–531

Routledge PA, O'Mahony MS, Woodhouse KW (2004) Adverse drug reactions in elderly patients. Br J Clin Pharmacol 57(2):121–126

Sumner LW, Urbanczyk-Wochniak E, Broeckling CD (2007) Metabolomics data analysis, visualization, and integration. Methods Mol Biol 406:409–436

Suresh K (2011) An overview of randomization techniques: an unbiased assessment of outcome in clinical research. J Hum Reprod Sci 4(1):8–11. https://doi.org/10.4103/0974-1208.82352

Tang J, Song X, Zhu M, Zhang J (2009) Study on the pharmacokinetics drug-drug interaction potential of Glycyrrhiza uralensis, a traditional Chinese medicine, with lidocaine in rats. Phytother Res 23(5):603–607. https://doi.org/10.1002/ptr.2450

Tilburt JC, Kaptchuk TJ (2008) Herbal medicine research and global health: an ethical analysis. Bull World Health Organ 86(8):594–599

Tolonen A, Turpeinen M, Pelkonen O (2009) Liquid chromatography-mass spectrometry in in vitro drug metabolite screening. Drug Discov Today 14(3–4):120–133. https://doi.org/10.1016/j.drudis.2008.11.002

USFDA (2016) Botanical drug development – guidance for industry. USFDA, Silver Spring

USFDA (2017) Multiple endpoints in clinical trials guidance for industry. USFDA, Sliver Spring

Vickers A, Zollman C (1999) ABC of complementary medicine: herbal medicine. BMJ 319(7216):1050–1053

Wang BL, JP H, Tan W, Sheng L, Chen H, Li Y (2008) Simultaneous quantification of four active schisandra lignans from a traditional Chinese medicine Schisandra chinensis (Wuweizi) in rat plasma using liquid chromatography/mass spectrometry. J Chromatog Analyt Technol Biomed Life Sci 865(1–2):114–120. https://doi.org/10.1016/j.jchromb.2008.02.016

Wang Y, Tang H, Nicholson JK, Hylands PJ, Sampson J, Holmes E (2005) A metabonomic strategy for the detection of the metabolic effects of chamomile (Matricaria recutita L.) ingestion. J Agric Food Chem 53(2):191–196. https://doi.org/10.1021/jf0403282

WHO (2005) National Policy on traditional medicine and regulation of herbal medicines. World Health Organisation, Geneva

WHO (2008) Essential Medicines – Biennial Report. World Health Organisation, Geneva

Wu JT, Zeng H, Qian M, Brogdon BL, Unger SE (2000) Direct plasma sample injection in multiple-component LC-MS-MS assays for high-throughput pharmacokinetic screening. Anal Chem 72(1):61–67

Xie G, Plumb R, Su M, Xu Z, Zhao A, Qiu M, Long X, Liu Z, Jia W (2008) Ultra-performance LC/TOF MS analysis of medicinal Panax herbs for metabolomic research. J Sep Sci 31(6–7):1015–1026. https://doi.org/10.1002/jssc.200700650

Xin GZ, Qi LW, Shi ZQ, Li P, Hao HP, Wang GJ, Shang J (2011) Strategies for integral metabolism profile of multiple compounds in herbal medicines: pharmacokinetics,

metabolites characterization and metabolic interactions. Curr Drug Metab 12(9):809–817

Xue T, Roy R (2003) Studying traditional Chinese medicine. Science 300(5620):740–741. https://doi.org/10.1126/science.300.5620.740

Xutian S, Zhang J, Louise W (2009) New exploration and understanding of traditional Chinese medicine. Am J Chin Med 37(3):411–426. https://doi.org/10.1142/s0192415x09006941

Yi T, Zhu L, Tang YN, Zhang JY, Liang ZT, Xu J, Zhao ZZ, ZL Y, Bian ZX, Yang ZJ, Chen HB (2014) An integrated strategy based on UPLC-DAD-QTOF-MS for metabolism and pharmacokinetic studies of herbal medicines: Tibetan "Snow Lotus" herb (Saussurea laniceps), a case study. J Ethnopharmacol 153(3):701–713. https://doi.org/10.1016/j.jep.2014.03.031

Zamiska N (2006) On the Trail of Ancient Cures. https://www.wsj.com/articles/SB116354914377323235

Chronopharmacology in Drug Development

25

Björn Lemmer

Contents

Abstract

The human body is highly organized in time and most functions display significant daily [circadian] and seasonal rhythms. Gene expression in the brain as well in the peripheral tissues has been convincingly demonstrated to be circadian phase dependent for many genes. In addition, seasonal gene expression – with inverse rhythms in the Northern and Southern Hemisphere – has been evidenced as well. Taking into account these rhythmic changes in the body, it is conceivable that also the pharmacokinetics and/or the pharmacodynamics can be rhythmic in experimental animals as well in humans. In this review chronopharmacological findings are compiled demonstrating to what extent rhythmic changes in the composition of the body in health and disease can contribute to variations in time-dependent drug treatments. The results of clinical chronopharmacological studies in hypertension, asthma, and ulcer disease and treatment of antihyperlipidemia are reviewed.

B. Lemmer (✉)
Institute of Experimental and Clinical Pharmacology and Toxicology, Medical Faculty, Mannheim Ruprecht-Karls-University of Heidelberg, Mannheim, Germany
e-mail: bjoern.lemmer@medma.uni-heidelberg.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_40

Introduction

Adequate drug dosing in a given disease/indication is the most important goal in clinical pharmacology. Drug reference books are designed to give advice to the best of scientific knowledge. Both pharmacokinetics and pharmacodynamics of a compound are well-accepted tools to achieve best dosing of a drug. Additional targets in drug dosing are involved: age, disease, gender, food, physical fitness, and family genetics. However, in recent years additional targets were shown to influence dosing, social aspects and living behavior such as shift work and transmeridian flight as well as social jet lag, and most important genetic variation in normal and diseased organ tissues – e.g., in cancer – came into focus in clinical medicine. Finally, the awareness of biological rhythms in organ and tissue functions leads to the introduction of chronopharmacology in clinical medicine. In this review the abovementioned topics will be touched in order to demonstrate the important new "environment" of clinical pharmacology.

The Biological Clock

Rhythmicity is the most ubiquitous feature of nature. Rhythms are found from unicellular to complex multicellular organisms in plants, animals, and men. Living organisms are continuously influenced by external stimuli, many of which have rhythmic patterns. Environmental rhythms in daily and seasonal patterns of light, food availability and temperature, etc. are predictable, and animals – including humans – have the ability to anticipate these environmental events with periodically and predictably changing internal conditions. These rhythmic patterns of anticipation have clear advantages and survival value. The frequencies of rhythms in nature cover nearly every division of time. There are rhythms which oscillate once per second (e.g., in the electroencephalogram), once per several seconds (respiratory rhythm, heart rate), and once per year (circannual rhythm).

The most evident environmental change which results from the regular spin of the earth around its central axis, resulting in the alternation between day and night, seems to have induced the predominant oscillation, the circadian rhythm (the about-24-h rhythm; circa = about, dies = day) described by Jürgen Aschoff (1947). There is sound evidence that living systems including humans are not only organized in space but are also highly organized in time (Albrecht 2012; Lemmer 2009, 2012a). One of the first observations on a rhythmic pattern in man was presented by the famous physiologist Sanctorius Sanctorius in 1664 when he described in a self-experiment daily variation in body weight due to transpiration. A review on biological rhythms described within the last several hundred years is compiled in a review (Lemmer 2009).

Circadian rhythms have been documented throughout the plant and animal kingdom at every level of eukaryotic organization. Circadian rhythms by definition are endogenous in nature, driven by oscillators or clocks, and persist under free-running conditions. In various species (e.g., *Drosophila melanogaster, Neurospora, mouse, golden hamster, rhesus macaque, man*), the genes controlling circadian rhythms have been identified (genes: *per, frq, clock, tau, Rev-erbalpha*). In 1971 Konopka and Benzer (1971) were able to identify on the X chromosome of *Drosophila* a region which controlled the period in the eclosion rhythm of three mutants (*per* clock gene). Circadian clocks are believed to have evolved in parallel with the geological history of the earth and have undergone selection pressures imposed by cyclic factors in the environment. These clocks regulate a wide variety of behavioral and metabolic processes in many life forms. They enhance the fitness of organisms by improving their ability to efficiently anticipate periodic events in their external environments, especially periodic changes in light, temperature, and humidity.

The mammalian circadian clocks, located in the neurons of suprachiasmatic nuclei (SCN) composed of about 10,000 heterogeneous neurons in the brain and in cells of peripheral tissues, are driven by a self-sustained molecular oscillator, which generates rhythmic gene expression with

a periodicity of about 24 h. This molecular oscillator is composed of interacting positive and negative transcription/translation feedback loops in which the heterodimeric transcription activator CLOCK/BMAL1 promotes the transcription of E-box containing Cryptochrome (Cry1 and Cry2) and Period (Per1 and Per2) genes, as well as clock-controlled output genes. After being synthesized in the cytoplasm, CRY and PER proteins feedback in the nucleus to inhibit the transactivation mediated by positive regulators. The mPER2 protein acts at the interphase between positive and negative feedback loops by indirectly promoting the circadian transcription of the Bmal1 gene and by interacting with mCRY proteins However, this is a simplified scheme with additional clock genes and transcription factors involved.

In general, the human endogenous clock does not run at a frequency of exactly 24 h but somewhat slower. The rhythm in human body temperature which is timed by the biological clock has a period of about 24.5-h under free-running conditions, i.e., without environmental time cues or zeitgebers (e.g., light, temperature). The term "zeitgeber" introduced by Jürgen Aschoff at the Max Planck Institute (Aschoff 1947) is now part of the international scientific language. Mammals such as rodents or humans can entrain their activity to regular light cycles not shorter than 22 or longer than 26 h. Zeitgebers entrain the circadian rhythm to a precise 24-h period. Zeitgebers are, therefore, necessary to entrain a living subject to a "normal" period of 24 h! This is of great importance having in mind that seasonal variations not only in the light-dark cycle but also in temperature and other environmental conditions have impact on availability of nutrition, on the onset of diseases, and on time of birth and death, the latter was demonstrated in a single family residing at different latitudes, in Europe (latitude 51°N, 1500–2013 A.C.) and South Africa (30°S, 1750–2013 A.C.) (Lemmer 2014). In this study no seasonal variation was found in birth and death in the data set from South Africa ($n = 1.284$), whereas a significant peak in winter was detected both in birth and death in the European data ($n = 2.361$) between 1500 and 1950, which was then lost from 1950 up to 2013, obviously due to changed socioeconomic and environmental conditions (Lemmer 2012b, 2014).

Chronopharmacology: Pharmacokinetics-Pharmacodynamics

The principles of the pharmacokinetic-pharmacodynamic interactions are well known and the basis for the development of new drugs. In short, pharmacokinetics deals with absorption, distribution, metabolism, and elimination of drugs. The different steps in pharmacokinetics are determined and influenced by physiological functions of the body. Pharmacokinetic parameters such as peak drug concentration [C_{max}], time to C_{max} [t_{max}], volume of distribution [V_d], area under the curve [AUC], bioavailability, plasma protein binding, and elimination half-life ([$t_{½}$] are evaluated during drug development but conventionally *not* considered to be influenced by the time of day at which a drug is administered. However, an increasing number of recently published studies convincingly gave evidence that this paradigm cannot be maintained any longer (Lemmer and Bruguerolle 1994).

The main reason to skip this paradigm is the demonstration that bodily functions, including those which are known to influence the pharmacokinetics, are not constant in time, even not within 24 h of a day. Figure 1 summarizes which parameters of the LADME System (Liberation-Absorption-Distribution-Metabolism) after oral drug administration have been demonstrated to be influenced by biological rhythms (Lemmer 1991b).

Thus, gastric emptying time of solids is faster in the morning than in the afternoon (Fig. 2) (Goo et al. 1987). In addition, the perfusion of the GI tract varies with time of day, being more pronounced ad midnight and the early morning hours than around noon or in the late afternoon (Fig. 3) (Lemmer and Nold 1991).

In the following chronokinetics are demonstrated in selected groups of compounds which are of importance in drug treatment.

Liberation	Absorption GI-Tract	Distribution	Metabolism Liver	Elimination Kidney
{Time- specified release programmable}	Perfusion Gastric pH Acid Secretion Motility Gastric Emptying Rest - Activity	Perfusion Blood Distribution Periph. Resistance Blood Cells Serum Proteins Protein Binding Rest - Activity	Perfusion First-Pass-Effect (Enzyme Activity)	Perfusion Renal Plasmaflow Glom. Filtration Urine Excretion Urine pH Electrolytes

Fig. 1 Effect of biological rhythms on pharmacokinetics of drugs (after oral administration) Copyright © SpringerNature with permission from SpringerNature

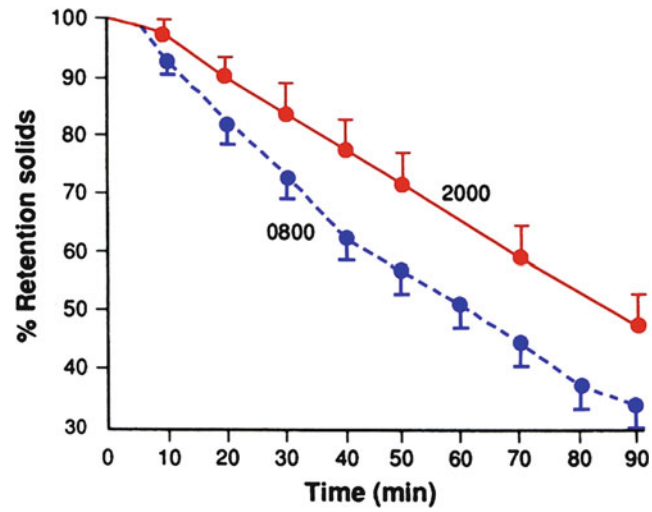


Fig. 2 Gastric emptying of solids in relation of time of day (From Goo et al. 1987). Copyright © Elsevier With permission from Elsevier

Chronopharmacology of Antiasthmatic Drugs

Since nocturnal asthma is a common event in asthmatic disease, it is not surprising that antiasthmatic drugs have also been studied in relation to time of day (for review see Lemmer 1991a, 2012a; Smolensky et al. 2007b). Theophylline was one of the first drugs for which daily variations in its pharmacokinetics were reported (for review see Smolensky et al. 2007b), though theophylline is of less importance nowadays. More than 50 studies with different theophylline preparations in different galenic formulations were published demonstrating that in general C_{max} was lower and/or t_{max} was longer after evening than after morning application of theophylline. This observation was supported by studies in which simultaneously the pharmacokinetics and the pulmonary effects of theophylline were compared in asthmatics, demonstrating that the drug might be dosed higher during the night than during daytime hours or even given as a single

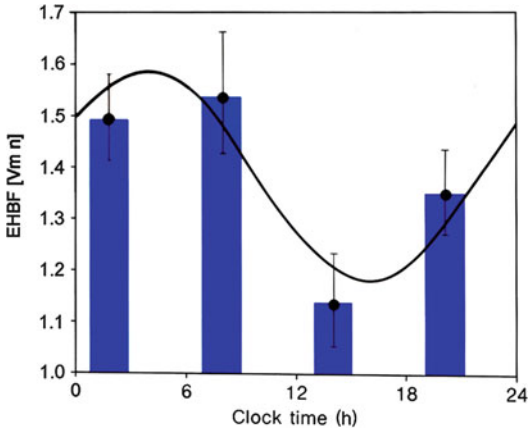


Fig. 3 Rhythm in gastrointestinal perfusion as estimated by hepatic blood flow (Lemmer and Nold 1991). Copyright © John Wiley and Sons. With permission from Wiley and Sons

evening dose in order to adequately overcome the nocturnal symptoms (see Smolensky et al. 2007b). Thus, in contrast to the general belief concerning drug concentration profiles, "the flatter the better," it seems, therefore, to be advantageous to accept greater fluctuations in drug concentrations throughout 24 h of a day. As a consequence in 1989 for the first time a pharmaceutical company was granted permission from the Food and Drug Administration in the USA and from the Bundesgesundheitsamt in Germany to market a sustained-release theophylline product for once-daily evening administration (see Smolensky et al. 2007b).

$Beta_2$ agonists are also drugs of first choice in the treatment of asthmatic patients, though inhaled beta agonists are preferable to the oral application in most cases. For oral terbutaline the pharmacokinetics but also the effects on peak expiratory flow were shown to be circadian phase dependent (Lemmer and Bruguerolle 1994) with higher C_{max} after morning than evening drug application and with t_{max} being 3.5 and 6.2 h, resp., thus resembling the daily variations observed with theophylline. A further study with oral terbutaline indicated that doubling the dose in the evening, i.e., unequally dosing during 24 h, can better control the nocturnal fall in peak flow. These studies give further support to the notion that the dose-response relationship of a given drug can be circadian phase dependent, as already demonstrated for theophylline and other compounds (see Lemmer and Bruguerolle 1994; Smolensky et al. 2007b).

Recently, the Commission on Drugs of the German Medical Association included "time of day" into their recommendations ("Arzneiverordnungen," 22th edition 2009 Allwinn et al. 2009) as an important variable influencing drug efficacy.

Chronopharmacology of H_2 Blockers in Peptic Ulcer Disease

This group of compounds are still the drugs of choice in the treatment of peptic ulcer. The chronobiologic finding on a circadian rhythm in gastric pH and acid secretion unanimously led to the recommendation that H_2 blockers (ranitidine, cimetidine, famotidine, roxatidine, nizatidine) should be taken once a day in the afternoon when acid secretion is increasing, independently of whether or not the compounds have a short or a long half-life (Lemmer 1991a; Moore 1989). In consequence of this strategy, chronopharmacology helped to improve drug treatment as well as the patient's compliance.

For both the H_2 blocker cimetidine and the proton pump inhibitor omeprazole, significant daily variations in their pharmacokinetics were shown with C_{max} being higher and t_{max} being shorter after morning than evening dosing. However, this does not seem to have an impact on drug efficacy. Daily variation in sensitivity to H_2 receptor blockade seems to be of more importance: Recently it has been shown that a continuous infusion of ranitidine over a period of 24 h does *not* lead to a constant effect, because the increase in gastric pH by ranitidine was less during the nightly than during the daytime hours of drug infusion (Sanders et al. 1988). This may indicate a partial nocturnal resistance to H_2 blockade. This interesting finding calls not only for further investigations but could also indicate that drugs with a different mechanism of action may be added to drug treatment with H_2 blockers during the nightly hours.

Chronopharmacology of Cardiovascular Active Drugs

The cardiovascular system displays pronounced daily variations in its functions as well as in its hormonal and biochemical regulatory mechanisms (Lemmer 2007b, 2017; Manfredini et al. 2012, 2017). Nearly all groups of cardiovascular active drugs were shown to exert a circadian phase dependency in their effects (Hermida et al. 2007b; Lemmer 2007a; Lemmer and Portaluppi 1997), for a less number of compounds; however, daily variation in pharmacokinetics was reported (Lemmer 2007a).

In the last years an increasing number of reports were published on the effects of cardiovascular active drugs on the 24-h blood pressure profile. It is not possible in this short overview to go into detail. In general, drugs belonging to different classes differently affected the 24-h blood pressure profile.

Beta blockers and calcium channel blockers reduced high blood pressure more pronouncedly during the daytime than during the night. However, only very few studies addressed this question in cross-over studies, i.e., evaluating morning versus evening drug dosing, as well as taking a possible variation in the drugs' pharmacokinetics or different galenic formulations into account (Lemmer 1991b, 2007c).

Both the pharmacokinetics and the cardiovascular effects on blood pressure and heart rate were simultaneously studied after oral application of propranolol (Langner and Lemmer 1988), isosorbide-5-mononitrate (IS-5-MN)) (Lemmer et al. 1991b; Scheidel and Lemmer 1991), nifedipine (Lemmer et al. 1990, 1991a), and enalapril (Witte et al. 1993) in a cross-over design (morning vs. evening) in healthy volunteers or hypertensive patients. Moreover, immediate-release and sustained-release preparations of IS-5-MN (Table 1) and nifedipine were investigated.

For propranolol and the i.r. preparations, significant daily variations in the pharmacokinetics were found with C_{max} being higher and/or t_{max} being shorter after morning than evening dosing (Table 1); results are very similar to those already described for antiasthmatics and antiulcer drugs (see above). The stereospecific metabolism of propranolol was not circadian phase dependent (Langner and Lemmer 1988). Interestingly, peak effects of propranolol in lowering heart rate coincided with peak drug concentrations only after propranolol intake at 08.00 and 14.00 h, and being delayed after drug dosing at 20.00 h and at 02.00 h (Fig. 4, Table 2), indicating a circadian time dependency in the dose-response relationship of propranolol. This clearly indicates that the chronopharmacokinetics of propranolol cannot mainly be responsible for the daily variations in the drug's hemodynamic effects; this must be related to the circadian variation in the sympathetic activity.

Table 1 Pharmacokinetics and pharmacodynamics of immediate-release and sustained-release formulations of isosorbide mononitrate (IS-5-MN), morning vs. evening: $^{*}p < 0.05$, $^{**}p < 0.01$ (Data from Lemmer et al. 1990; Scheidel and Lemmer 1991)

	IS-5-MN: i.r.		IS-5-MN: s.r.	
	06.30 h	18.30 h	08.00 h	20.00 h
Pharmacokinetic				
C_{max} (ng/ml)	1,605 ± 175	1,588 ± 173	509 ± 31	530 ± 26
t_{max} (h)	0.9 ± 0.3	2.1 ± 0.4**	5.2 ± 0.7	4.9 ± 0.3
AUC (ng/ml/h)	9,539 ± 827	10,959 ± 707	6,729 ± 375	6,418 ± 199
$t_{1/2\ ß}$ (h)	4.6 ± 0.4	4.2 ± 0.4	6.4 ± 0.6	6.1 ± 0.5
Hemodynamic				
T_{max} SBP decrease (mmHg)	0.7 ± 0.1	1.1 ± 0.1	5.0 ± 0.6	2.8 ± 0.5*
T_{max} DBP decrease (mmHg)	0.4 ± 0.1	0.6 ± 0.2	6.0 ± 0.7	2.9 ± 0.5**
T_{max} HR decrease (S/min)	0.8 ± 0.3	0.9 ± 0.2	5.2 ± 1.0	3.8 ± 0.6

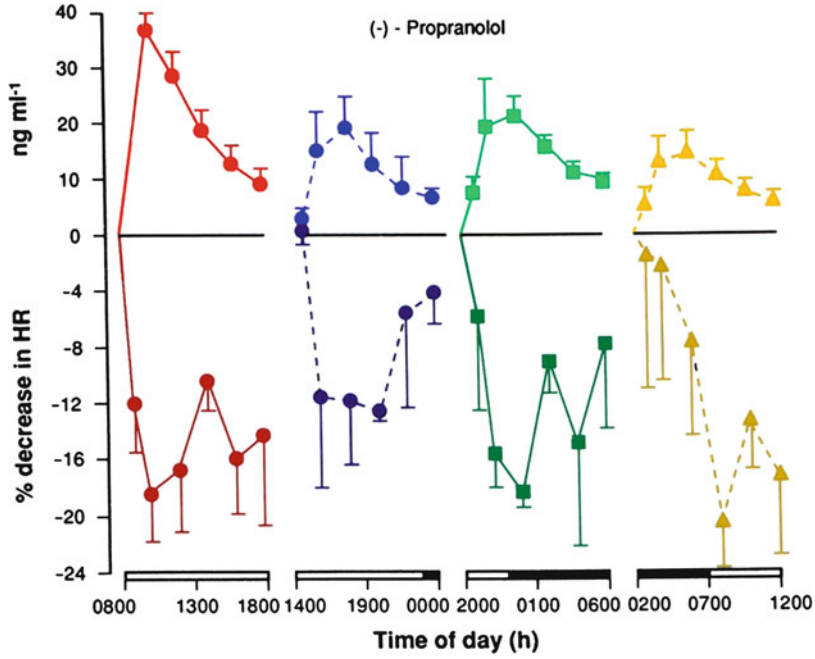


Fig. 4 Chronopharmacology of (±)-propranolol (80 mg p. o.) in healthy subjects given at four different times of day. Shown are the plasma concentration oh (-)-propranolol and the decrease in heart rate in relation to control values. (From Langner and Lemmer 1988). Copyright © SpringerNature. With permission of SpringerNature

Conventionally, different galenic formulations are only considered in that way that the duration of drug action is modified. For two drugs, however, the pharmacokinetics of both an immediate-release and a sustained-release formulation have been studied at different times of day giving evidence that circadian time can influence the kinetics depending on the formulation (Table 2): In contrast to the i.r. nifedipine mentioned above, a s.r. formulation did not display significant daily variations in its pharmacokinetics in hypertensive patients. A sustained-release preparation of the calcium channel blocker verapamil, on the other hand, did show a significant longer t_{max} after evening than morning dosing in healthy subjects.

Very similar to the findings with different formulations of nifedipine were results obtained with two formulations of IS-5-MN in healthy subjects: With the immediate-release preparation of IS-5-MN, clear-cut daily variations were found in regard to t_{max} after morning (0.9 ± 0.3 h) or after evening (2.1 ± 0.4 h) application (Table 3). Most interestingly, time to peak drug effects in decreasing blood pressure and reflexly increasing heart rate coincided with t_{max} in pharmacokinetics in the morning, but were in advance by about 1 h in the evening. No daily variations were observed in the pharmacokinetics of the sustained-release formulation of IS-5-MN. Nevertheless, peak drug effects in lowering blood pressure and increasing heart rate again coincided with t_{max} but occurred about 2 h earlier after drug application in the evening. This again indicates daily variations in the dose-response relationship of drugs as already mentioned before for digoxin, propranolol, and nifedipine. Moreover, the data also demonstrate that the kind of drug formulation may be of importance whether or not chronokinetics can be observed.

In the following cross-over studies (morning vs. evening) are presented which demonstrate that the time of day of drug application can have an effect on the drug response on the 24-h blood

Table 2 Chronokinetics of cardiovascular active drugs, * at least <0.05 (Data from and references in Lemmer 2006)

Drug	Dose (mg) and duration	Cmax (ng/ml)		t_{max} (h)		References
		Morning	Evening	Morning	Evening	
Digoxin	0.5, single dose	3.6*	1.8	1.2	3.2	Bruguerolle et al. (1988)
Enalapril	10					Witte et al. (1993)
Enalaprilat	Single dose	33.8	41.9	4.4	4.5	
Enalaprilat	3 weeks	46.7	53.5	3.5*	5.6	
IS-5-MN I.R.	60, single dose	1605.0	1588.0	0.9*	2.1	Scheidel and Lemmer (1991) and Lemmer et al. (1989)
IS-5-MN S.R.	60, single dose	509.0	530.0	5.2	4.9	Lemmer et al. (1991a)
Molsidomine	8, single dose	27.0	23.5	1.7	1.7	Nold and Lemmer (1998)
Nifedipine I.R.	10, single dose	82.0*	45.7	0.4*	0.6	Lemmer et al. (1991a)
Nifedipine S.R.	2 × 20, 1 week	48.5	50.1	2.3	2.8	Lemmer et al. (1991a, b)
Atenolol	50, Single dose	440.0	391.8	3.2	4.0	Shiga et al. (1993)
Oxprenolol	80, single dose	507.0	375.0	1.0	1.1	Koopmans et al. (1993)
Propranolol (±)	80, single dose					Langner and Lemmer (1988)
Propranolol (-)		38.6*	26.2	2.5	3.0	
Propranolol (±)	80, single dose	68.0	60.0	2.3	2.7	Semenowicz-Siuda et al. (1984)
Verapamil S.R.	360, 2 weeks	389.0	386.0	7.2*	10.6	Jespersen et al. (1989)
Verapamil	80, single dose	59.4*	25.6	1.3	2.0	Hla et al. (1992)

pressure profile. Fortunately, those studies were performed for those drugs which are drugs of choice in treating hypertension, i.e., calcium channel blockers, ACE inhibitors, and AT_1 receptor blockers. These studies clearly demonstrate that those clinical studies are of importance for drug development by the pharmaceutical industry in order to better treat chronic diseased patients with hypertension.

Cross-Over Studies with Calcium Channel Blockers

Calcium channel blockers are also not a homogenous group of drugs. Vasodilatation by calcium channel blockers occurs at lower concentrations than the cardiodepressant effects. However, the difference between vasodilating and cardiodepressant effects is greater with the 1,4-dihydropyridines (e.g., nifedipine, nitrendipine, isradipine, amlodipine) than with the verapamil- and diltiazem-like compounds. Moreover, these drugs differ in their kinetics with amlodipine having a long half-life per se.

A single morning dose of a sustained-release verapamil showed a good 24-h blood pressure control, whereas a sustained-release formulation of diltiazem was less effective at night. Dihydropyridine derivatives, differing in pharmacokinetics, seem to reduce blood pressure to a varying degree during day and night; drug formulation and dosing interval may play an additional role.

Up to now, 12 studies using a cross-over design [morning vs. evening] have been published (Table 3). In primary essential hypertensives

Table 3 Cross-over studies (morning vs. evening) with calcium channel blockers, *EH* essential hypertension, *ND* non-dipping, *NT* normotension, *RH* renal hypertension (Data from and references in Lemmer (2005, 2012a))

Drug	Dose, Duration	Patients, (n)	Timing of Treatment	24h-Profile	Reference
Amlodipine	5, 4 wks	EH, (20)	- a.m. - p.m.	preserved preserved	Mengden et al. 1992
Amlodipine	5, 3 wks	EH, (12)	- 08.00 - 20.00	preserved preserved	Nold et al. 1998
Isradipine	5, 4 wks	EH, (18)	- 07.00 - 19.00	preserved preserved	Fogari et al. 1993
Lacidipine	4, 6 wks	EH, (33)	- 07-09.00 - 22-24.00	preserved preserved ?	van Montfrans et al. 1998
Nifedipine GITS	30, 1 or 2 wks	EH, (10)	- 10.00 - 22.00	preserved preserved	Greminger et al. 1994
Nisoldipine ER	20, 4 wks	EH, (39)	- 09-11.00 - 21-23.00	preserved preserved	White et al. 1999
Nitrendipine	20, 4 wks	EH, (41)	- 07.00 - 19.00	preserved preserved	Meilhac et al. 1992
Nitrendipine	10, 3 days	EH, (6)	- 06.00 - 18.00	preserved (changed)	Umeda et al. 1994
Amlodipine	5, 3 wks	ND, (39)	- 08.00 - 20.00	dipping dipping	Lemmer et al., 2002
Isradipine	5, 4 wks	RH, (16)	- 08.00 - 20.00	non-dipping dipping	Portaluppi et al. 1995
Nisoldipine	20, 4 wks	ND, (36)	- 09-11.00 - 21-23.00	dipping dipping	White et al. 1999
Nifedipine i.r.	10, s.d.	NT, (12)	- 08.00 - 19.00	preserved preserved	Lemmer et al. 1991

with a dipper profile, amlodipine, isradipine, nifedipine GITS, and nisoldipine did not differently affect the 24-h blood pressure profile after once morning or once-evening dosing, whereas with nitrendipine and lacidipine, the profile remained unaffected or slightly changed after evening dosing. Most interestingly, the greatly disturbed blood pressure profile in secondary hypertensives [non-dippers] due to renal failure was only normalized after evening but not after morning dosing of isradipine. Similarly, amlodipine and nisoldipine ER transformed non-dippers into dippers, but both after evening and morning dosing (Table 3), which might be due to the longer "apparent" half-life of these drugs. These findings demonstrate that time of drug dosing of a dihydropyridine calcium channel blocker can be advantageous in not only reducing the elevated blood pressure but also normalizing the disturbed blood pressure profile.

A time-of-day effect was also described for the kinetics of various calcium channel blockers (Table 2). The bioavailability of an immediate-release formulation of nifedipine was found to be reduced by about 40% after evening compared to morning dosing with C_{max} being higher and t_{max} being shorter after morning dosing. No such circadian time-dependent kinetics were observed with a sustained-release formulation of nifedipine. Also regular as well as sustained-release verapamil displayed higher C_{max} and/or shorter t_{max} values after morning dosing. Similar chronokinetics have been reported after oral dosing of other cardiovascular active drugs such as enalapril, propranolol, and others. Conversely, intravenously infused nifedipine did not display daily variations in its pharmacokinetics indicating that gastrointestinal mechanisms must be involved in the drug's chronokinetics after oral application.

Cross-Over Studies with Converting Enzyme Inhibitors

Captopril is the representative of a short-acting converting enzyme inhibitor; longer acting ones

Table 4 Cross-over studies with ACE inhibitors, references in Lemmer (2006)

Drug	Patients (n)	Dose (mg)	Duration (days, wks)	Dosing time	Lowering of Blood Pressure Day	Night	24h-Profile	References
ACE-Inhibitors (Dipper)								
Benazepril	10	10	1 d	09.00h	↓↓↓	↓↓	preserved	Palatini et al., 1993
				21.00h	↓	↓↓	changed	
Enalapril	10	10	1 d	08.00h	↓↓	↓	preserved	Witte et al., 1993
				19.00h	↓↓	↓↓↓	changed	
Enalapril	10	10	3 wks	08.00h	↓↓	↓	preserved	Witte et al., 1993
				19.00h	↓	↓↓	changed	
Perindopril	18	2	4 wks	19.00h	↓↓	↓	preserved	Morgan et al., 1997
				21.00h	↓	↓↓	changed	
Quinapril	18	20	4 wks	08.00h	↓↓	↓	preserved	Palatini et al., 1992
				22.00h	↓↓	↓↓	preserved	
Ramipril	33	2.5	4 wks	08.00h	↓	(↓)	preserved	Myburgh et al., 1995
				20.00h	(↓)	↓	preserved	
Trandolapril	30	1 - 2	8 wks	morning	↓↓	↓↓	no profiles shown at both time points	Koroda et al., 2004
				bedtime	↓↓	↓↓		

Table 5 Cross-over studies with AT_1 receptor blockers. Acc. (Lemmer 2006)

Drug	Patients (n)	Dose (mg)	Duration (days, wks)	Dosing time	Lowering of Blood Pressure Day	Night	24h-Profile	References
AT_1-Receptor-Blockers (Dipper)								
Valsartan [1]	90	160	3 months	awaking	↓↓	↓↓	preserved	Hermida et al., 2003
				bedtime	↓↓	↓↓	preserved	
Irbesartan	20	100	6 wks	morning	↓↓	↓↓	Preserved	Pechère-Bertschi et al., 1998
				evening	↓↓	↓↓	preserved	
Olmesartan	18	20 (40)	12 wks	08.00	↓↓	↓↓	preserved	Smolensky et al., 2007
				20.00	↓↓	↓↓	preserved	
Telmisartan [1]	215	80	12 wks	wakening	↓↓	↓↓	preserved	Hermida et al., 2007
				bedtime	↓↓	↓↓(↓)	preserved	
AT_1-Receptor-Blockers (Non-Dipper [2])								
Valsartan [1,2]	148	160	3 months	wakening	↓↓	↓↓	preserved	Hermida et al., 2005
				bedtime	↓↓	↓↓↓	preserved	

[1] randomly assigned to morning and evening dosing; [2] non-dippers with nightly fall slightly < 10%

Table 6 Effects of CSH inhibitors on lipids (% decrease) after morning or evening administration (Illingworth 1986; Saito et al. 1991; Wallace et al. 2003)

Drug	Dose (mg/day)	Cholesterol	LDL	HDL	Triglycerides
Lovastatin					
a.m. ($n = 12$)	20–40	−21.4	−26.9	+1.1	−15.0
p.m. ($n = 12$)	20–40	−27.0	−32.2	±0.0	−15.0
Simvastatin					
a.m. ($n = 32$)	5	−13.7	−19.3	+2.8	+1.7
p.m. ($n = 29$)	5	−20.7	−28.5	+5.4	−4.9
Atorvastatin					
a.m. ($n = 15$)	40	−33.1	−47.2	+1.3	−22.8
p.m. ($n = 15$)	40	−34.3	−48.2	+2.3	−26.4

Table 7 Chronokinetics of atorvastatin after morning and evening application (Cilla et al. 1996)

Atorvastatin	Morning	Evening
C_{max} (ng/ml)	95.0	65.9
t_{max} (h)	1.9	2.9

Table 8 Chronokinetics of pravastatin after morning and evening application (Triscari et al. 1995)

Pravastatin	Morning	Evening
C_{max} (ng/ml)	28.0	11.0*
t_{max} (h)	1.0	1.5*

are enalapril, benazepril, quinapril, and ramipril, and these prodrugs are hydrolyzed to active metabolites. Converting enzyme inhibitors are not only effective antihypertensive drugs but can also increase the life expectancy in congestive heart failure.

Several studies with converting enzyme inhibitors dosed once in the morning or twice daily showed that these drugs did not greatly modify the 24-h blood pressure pattern. However, intra-arterial studies with enalapril or ramipril have shown that while causing sustained daytime reduction in blood pressure, these drugs had only marginal effects on nighttime pressures. Thus, the findings obtained with converting enzyme inhibitors in conventional, i.e., not time specific, clinical studies are controversial.

Seven cross-over studies (morning vs. evening dosing) with converting enzyme inhibitors in essential hypertensive patients were published (Table 4). They demonstrate that evening dosing in contrast to morning dosing resulted in a more pronounced nightly drop (super dipping) and the 24-h blood pressure profile was distorted by evening enalapril (Table 4). Evening dosing of quinapril and perindopril also resulted in a more pronounced effect than morning dosing. In the light of a reduced cardiac reserve of patients at risk of hypertension, a too pronounced nightly drop in blood pressure (super dipping) after evening dosing might be a potential risk factor for the occurrence of ischemic events such as cerebral infarction (Table 5).

Cross-Over Studies with AT_1 Receptor Blockers

There are four studies published with the AT_1 receptor blockers in dippers, valsartan, irbesartan, telmisartan, and olmesartan, which similarly reduced the blood pressure after both morning and bedtime dosing (Hermida et al. 2003, 2007a; Smolensky et al. 2007a; Pechere-Bertschi et al. 1998). In patients with chronic renal disease, olmesartan restored the nightly

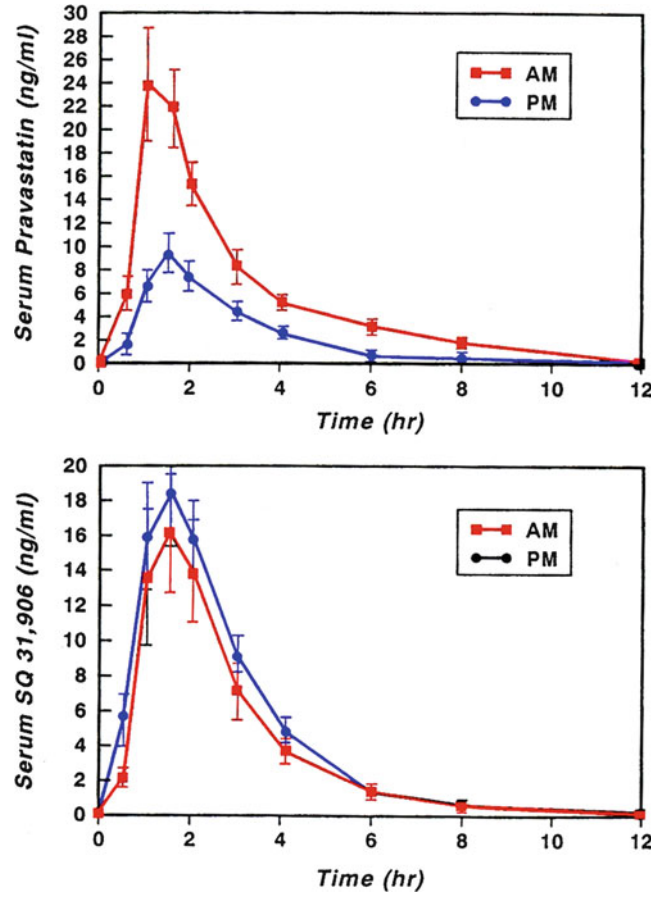


Fig. 5 Chronokinetics of pravastatin and its major metabolite SQ 31,906 after morning vs. evening application of pravastatin in 20 healthy subjects (Triscari et al. 1995). Copyright © Wolters Kluwer Health, Inc. With permission of Wolters Kluwer Health, Inc.

decline in blood pressure (Fukuda et al. 2008). In non-dippers valsartan had only a slightly more pronounced effect at night (Hermida et al. 2003).

Diuretics and Other Antihypertensive Drugs

Antihypertensives of other classes have rarely been studied in relation to possible circadian variation. Once-daily morning dosing of the diuretics such as indapamide or xipamide reduced blood pressure in essential hypertensives without changing the 24-h blood pressure pattern. In salt-sensitive hypertensive patients (dippers and non-dippers), an interesting study was performed with diuretics: Uzu and Kimura (1999) could demonstrate that diuretics did not affect the circadian blood pressure profile in dippers but transformed the non-dippers into dippers.

In conclusion, there is sound evidence that the treatment of hypertension by various groups of antihypertensive drugs is dependent on the underlying rhythmic organization of the cardiovascular system. Since we are not inbred strains of rats, individualization of drug therapy is the choice of treatment; fortunately, the ability of a great number of drugs from various antihypertensive groups allows us to treat a patient according to his individual setting. As a rule non-dipping hypertensive patients seem to be best treated when the drug is given at bedtime. In dipping hypertensives morning dosing seems to be best, and it seems not justified to treat all hypertensive patients with an antihypertensive drug given in the evening. Thus, the underlying rhythmic pattern of blood

Table 9 Summary of drugs for which chronokinetic and/or chronodynamic studies were performed

Cardiovascular Active Drugs

Beta-Blockers
- Acebutolol
- Atenolol
- Bevantolol
- Bopindolol
- Labetolol
- Mepindolol
- Bisoprolol
- Nebivolol
- Metoprolol
- Nadolol
- Oxprenolol
- Pindolol
- Propranolol
- Sotalol
- Carvedilol
- Timolol (IOP)

Beta-Agonists
- Xamoterol
- Terbutaline (IOP)
- Midodrine
- Adrenaline (IOP)

Calcium Channel Blockers
- Amlodipine
- Nifedipine
- Nisoldipine
- Diltiazem
- Nicardipine
- Clentiazem
- Nitrendipine
- Verapamil
- Lacidipine
- Isradipine
- Nivaldipine
- Cilnidipine

ACE Inhibitors
- Captopril
- Quinapril
- Perindopril
- Benazepril
- Sampatrilat
- Enalapril
- Lisinopril
- Spirapril
- Trandolapril

Diuretics
- Hydrochlorothiazide
- Indapamide
- Indapamide SR
- Furosemide
- Xipamide
- Piretanide
- Torasemide

AT_1-Receptor-Blocker
- Losartan
- Irbesartan

Organic Nitrates
- Glyceryl-trinitrate
- Isosorbide-dinitrate
- Isosorbide-5-mononitrate

Others
- Clonidine
- Phentolamine
- Indoramine
- Potassium chloride
- Sodium nitroprusside
- Prazosin

Anticancer Drugs
- Cisplatin
- THP
- Folinic acid
- Doxorubicin
- Busulphane
- Oxaliplatin
- FUDR
- Methotrexate
- Combinations

Miscellaneous
- Tuberculine
- Ethanol
- Heparin
- Placebo
- Endotoxin
- Salmon Calcitonin
- Nadroparine
- Bright light

Antiasthmatic Drugs
- Theophylline
- Orciprenaline
- Terbutaline
- Metacholine
- Methylprednisolone
- Dexamethasone
- Budesonide
- Adrenaline
- Isoprenaline
- Terbutaline + Budesonide
- Aminophylline
- Bambuterol

Psychotropic Drugs
- Diazepam
- Haloperidol
- Phenylpropanolamine
- Caffeine
- Clomipramine

H_1-Antihistamines
- Clemastine
- Cyproheptadine
- Mequitazine
- Terfenadine

Ophthalmology
- Terbutaline
- Adrenaline
- Isoprenaline
- Timolol

NSAIDs, General and Local Anesthetics and Opioids
- Acetylsalicylic acid
- Flurbiprofen
- Ketoprofen
- Paracetamol
- Tenoxicam
- Metamizole
- Carticaine
- Mepivacaine
- Morphine
- Narcotic analgesics
- Halothane
- Ibuprofen
- Pranoprofene
- Piroxicam
- Indomethacin
- Lidocaine
- Fentanyl
- Propofol

Endocrinology/Gastroenterology
- Prednisone
- Methylprednisolone
- Insulin
- Glucose
- Bezafibrate
- Simvastatin
- Omeprazole
- HMG CoA-Red. Inhib.
- Mevinolin
- ACTH
- Tolbutamide
- Clofibrate
- Lansoprazole

H_2-Blockers
- Cimetidine
- Nizatidine
- Roxatidine
- Famotidine
- Ranitidine

pressure regulation has impact on drug treatment. In order to achieve this goal, it is essential that a 24-h blood pressure profile by ABPM is used in each patient.

However, unfortunately only a small group of patients by about 25% is really controlled by ABPM as shown in epidemiological studies in Germany and worldwide (Lemmer et al. 2008).

Table 10 Recommendation of the German Medical Association to take into account chronopharmacological findings in the prescription of drugs (Allwinn et al. 2009), composition in (Lemmer 2012a)

	Rheumatic disease
p 299	Morning stiffness: evening dosing of NSAR
p 300	Taking into account rhythm in pain perception
p 307	Morning dosing of glucocorticosteroids
	Depression
p 433	Lithium increases circadian rhythm in physiological functions
p 415	Sleep deprivation, light therapy
	Hypertension
p 601	24-h blood pressure monitoring [ABPM],white coat effect
p 604	Antihypertensives: morning dosing
p 625	Altered BP rhythm in pregnancy:
	High BP evening, at night
	Bronchial asthma
p 776	Asthma with nightly symptoms
p 778	Symptoms mainly at night and morning
p 778	Day-night rhythm in FEV_1
p 785	Long-acting beta agonists in nightly asthma
p 790	Theophylline evening dosing in nightly asthma
	Peptic ulcer
p 828	Proton pump inhibitors: morning dosing
	Hyperlipidemia
p 1053–1055	Statins evening dosing
	M. Addison
p 1137	High morning glucocorticoid concentration: Evening dosing of slow-release preparations
p 1140	Prednisolone single dose in the morning
p 1138	Circadian dosing
	Pituitary insufficiency/male sexual disturbances
p 1119	Somatotropin evening dosing
p 1178	Testosterone-TTS evening dosing
	Melatonin and sleep-wake rhythm
p 1223	Melatonin in blind persons
	Melatonin in jet-lag symptoms

Chronopharmacology of CSE Inhibitors

The HMG CoA reductase inhibitors (CSE inhibitors) are widely used to treat hypercholesteria in patients. They have also been studied in relation to time of day of drug dosing (Table 6). All studies demonstrate that evening dosing should be preferred since efficacy is greater and side effects are less than after morning dosing. Moreover, for atorvastatin and pravastatin also daily variation in their pharmacokinetics was demonstrated (Tables 7 and 8), which, however, does not seem to play a role in efficacy. Fortunately, the evening dosing has been realized by the pharmaceutical industry and the medical association (Table 8).

Figure 5 shows the chronokinetics of pravastatin in healthy subjects. The lower AUC of pravastatin following PM dosing does not diminish its efficacy, possibly because PM dosing immediately precedes the diurnal peak period of hepatic cholesterol synthesis. Lower blood levels of pravastatin following PM dosing may contribute to its favorable safety profile (Triscari et al. 1995) (Tables 9 and 10).

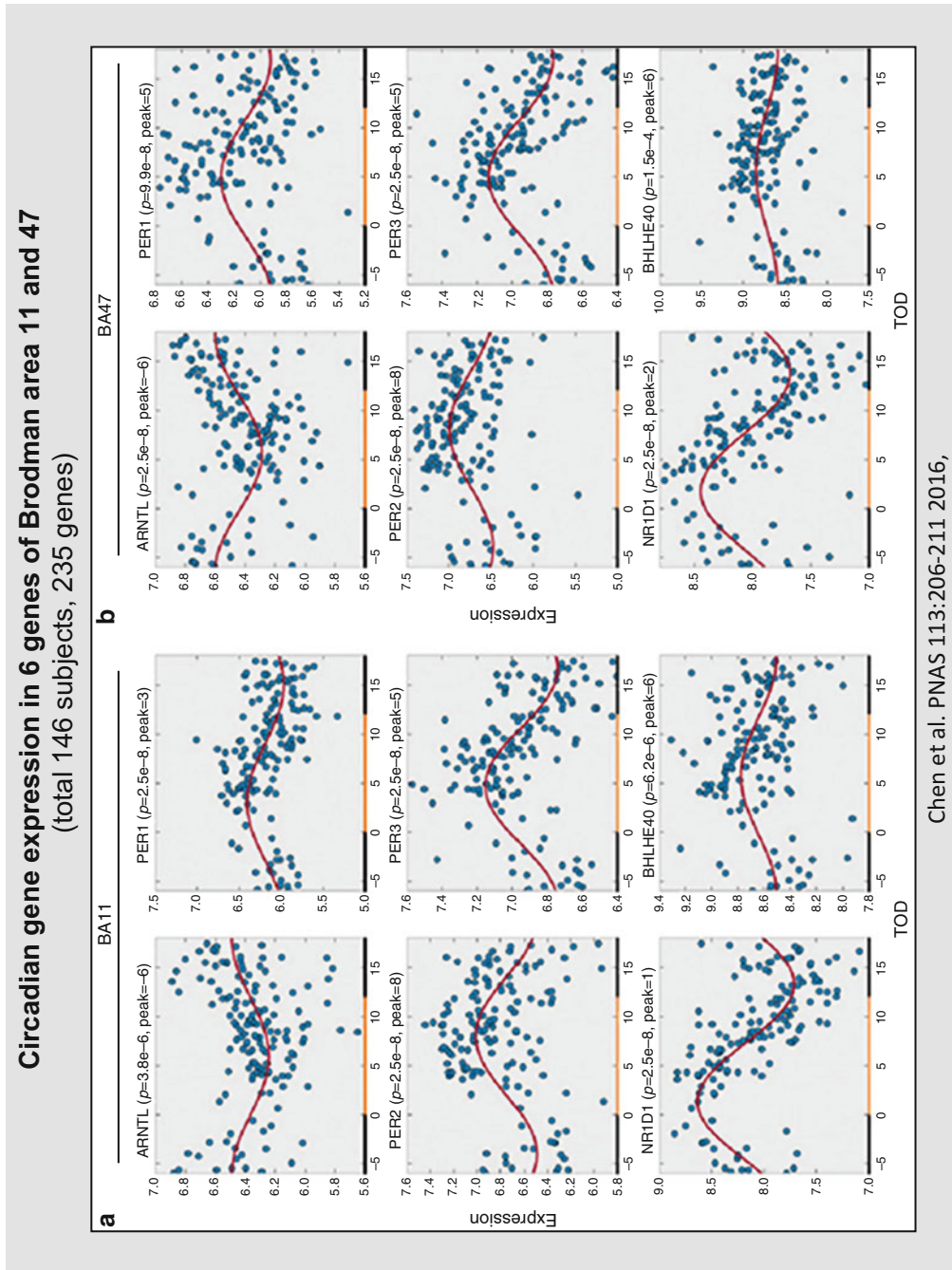


Fig. 6 Circadian gene expression in man (Chen et al. 2016). With permission from PNAS

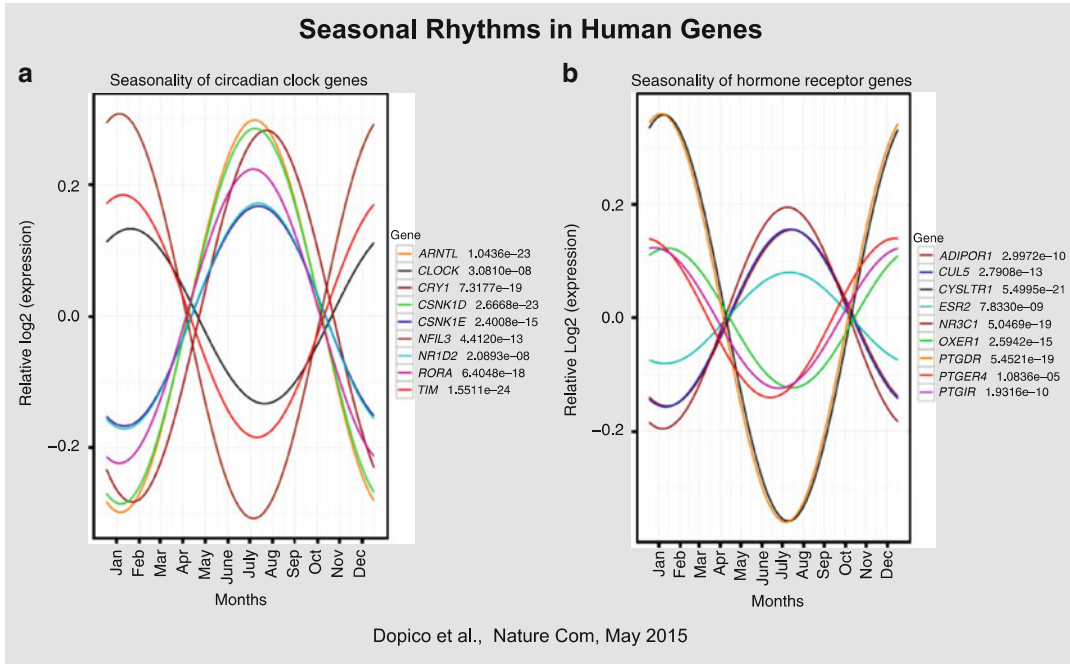


Fig. 7 Seasonal rhythms in human clock and hormone receptor genes (Dopico et al. 2015). Copyright © 2015, Rights Managed by Nature Publishing Group with creative commons license

Conclusion

In conclusion, this review sheds a light on the importance of the time of day at which a drug is given to a patient and how the pharmacokinetics and/or the effects can vary with time of day. Thus, it is of utmost importance to include this chronopharmacological paradigm into the design of a clinical study. In the last years chronopharmacological findings were included in the recommendations of the German Medical Association to the doctors when prescribing drugs in certain entities of diseases (Table 8).

In this review it was only possible to outline the contribution of circadian rhythms to the pharmacokinetics and the effects of drugs used in the treatment of various diseases. There is no doubt that the rhythmic organization of the human body has an impact on drug treatment.

This review finally draws attention to the increasing number of reports demonstrating daily and seasonal variations in central and peripheral gene expression, e.g., Chen et al. 2016, Dallmann et al. 2012, and Dopico et al. 2015 (Figs. 6, 7, and 8).

It is plausible to assume that this time-dependent expression of genes can/must be involved in daily as well as seasonal variation in disease entities as well as in metabolism and effects of drug.

In the light of the tendency for an increased individualization of drug therapy, drug development should consider an additional aspect of drug efficacy, i.e., the contribution of physiological-biological rhythms – down to the expression of genes – to the kinetics and effects of drugs. As demonstrated in this review, this is possible when the underlying rhythmic pattern in health and disease is regarded as a tool for improving drug development.

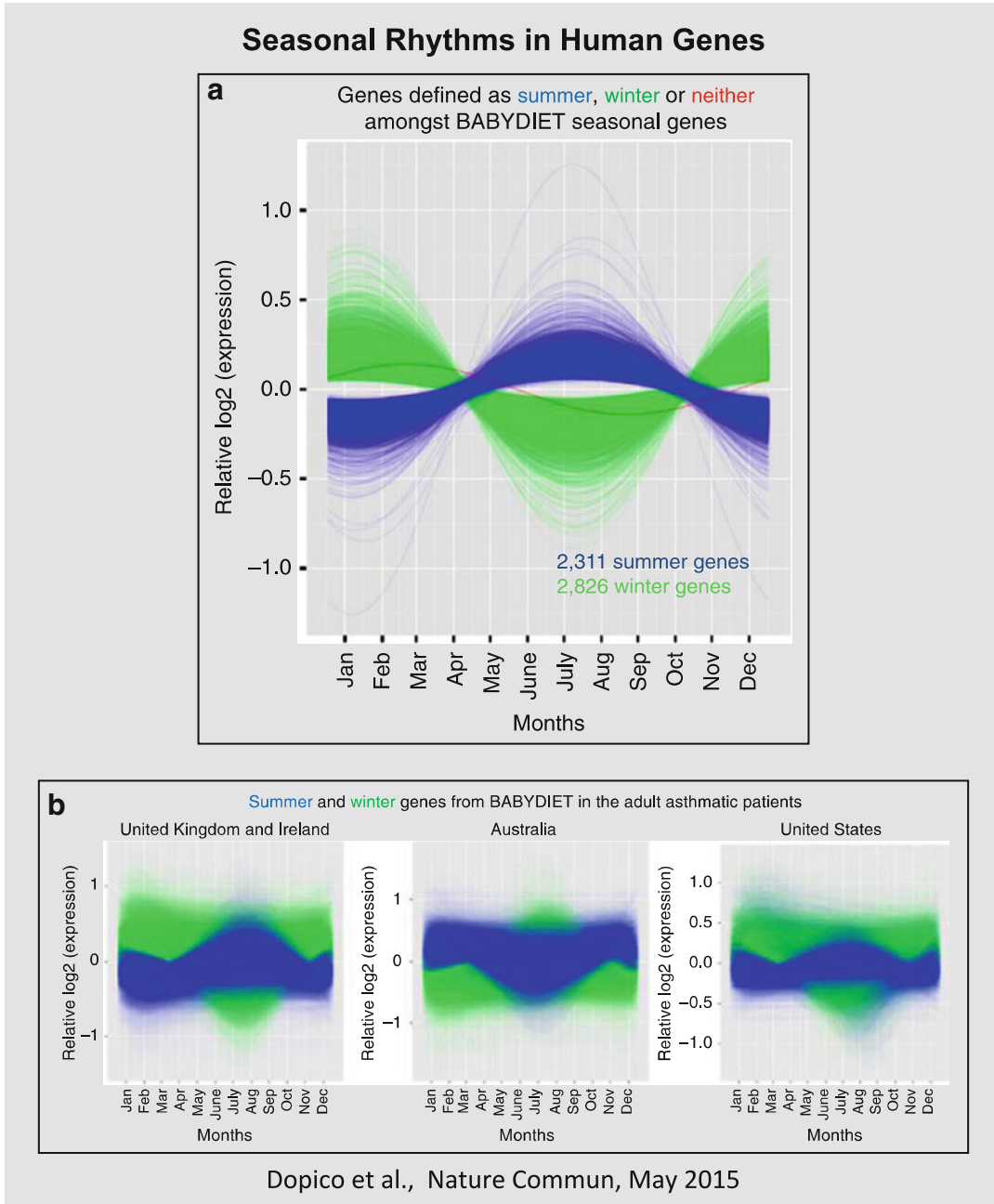


Fig. 8 Seasonal rhythms in human in the Northern and Southern Hemisphere (Dopico et al. 2015). Copyright © 2015, Rights Managed by Nature Publishing Group with creative commons license

The Nobel Prize award in 2017 for the molecular biology of the body clock underlines also the importance of biological rhythms for medicine and this award can further stimulate the incorporation of this branch of science into clinical research.

Disclosure The author has no conflict of interest.

References and Further Reading

Albrecht U (2012) Timing to perfection: the biology of central and peripheral circadian clocks. Neuron 74(2):246–260. https://doi.org/10.1016/j.neuron.2012.04.006

Allwinn R et al (2009) Arzneiverordnungen. In: Arzneimittelkommission, der, deutschen, Ärzteschaft (eds). 22 edn. Medizinische Medien Informations GmbH, Neu Isenburg

Aschoff J (1947) Einige allgemeine Gesetzmäßigkeiten physikalischer Temperaturregulation. Pflügers Arch 249:125–136

Bruguerolle B, Bouvenot G, Bartolin R, Manolis J (1988) Chronopharmacocinetique de la digoxine chez le sujet de plus de soixante-dix ans. Therapie 43:251–253

Chen CY et al (2016) Effects of aging on circadian patterns of gene expression in the human prefrontal cortex. Proc Natl Acad Sci U S A 113(1):206–211. https://doi.org/10.1073/pnas.1508249112

Cilla DD, Gibson DM, Whitfield LR, Sedman AJ (1996) Pharmacodynamic effects and pharmacokinetics of atorvastatin after administration to normocholesterolemic subjects in the morning and evening. J Clin Pharmacol 36:604–609

Dallmann R, Viola AU, Tarokh L, Cajochen C, Brown SA (2012) The human circadian metabolome. Proc Natl Acad Sci U S A 109(7):2625–2629. https://doi.org/10.1073/pnas.1114410109

Dopico XC et al (2015) Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat Commun 6:7000. https://doi.org/10.1038/ncomms8000

Fukuda M et al (2008) Angiotensin II type 1 receptor blocker, olmesartan, restores nocturnal blood pressure decline by enhancing daytime natriuresis. J Hypertens 26(3):583–588

Goo RH, Moore JG, Greenberg E, Alazraki NP (1987) Circadian variation in gastric emptying of meals in humans. Gastroenterology 93:515–518

Hermida RC et al (2003) Administration time-dependent effects of valsartan on ambulatory blood pressure in hypertensive subjects. Hypertension 42(3):283–290

Hermida RC, Ayala DE, Fernandez JR, Calvo C (2007a) Comparison of the efficacy of morning versus evening administration of telmisartan in essential hypertension. Hypertension 50(4):715–722

Hermida RC, Ayala DE, Smolensky MH, Portaluppi F (2007b) Chronotherapy in hypertensive patients: administration-time dependent effects of treatment on blood pressure regulation. Expert Rev Cardiovasc Ther 5(3):463–475

Hla KK, Latham AN, Henry JA (1992) Influence of time of administration on verapamil pharmacokinetics. Clin Pharmacol Ther 51(4):366–370

Illingworth DR (1986) Comparative efficacy of once versus twice daily mevinolin in the therapy of familial hypercholesterolemia. Clin Pharmacol Ther 40:338–343

Jespersen CM, Frederiksen M, Hansen JF, Klitgaard NA, Sorum C (1989) Circadian variation in the pharmacokinetics of verapamil. Eur J Clin Pharmacol 37(6):613–615

Koopmans R, Oosterhuis B, Karemaker JM, Wemer J, van Boxtel CJ (1993) The effect of oxprenolol dosage time on its pharmacokinetics and haemodynamic effects during exercise in man. Eur J Clin Pharmacol 44:171–176.

Konopka RJ, Benzer S (1971) Clock mutants of Drosophila melanogaster. Proc Natl Acad Sci U S A 68(9):2112–2116

Langner B, Lemmer B (1988) Circadian changes in the pharmacokinetics and cardiovascular effects of oral propranolol in healthy subjects. Eur J Clin Pharmacol 33:619–624

Lemmer B, Behne S, Becker HJ (1989) Chronopharmacology of oral nifedipine in healthy subjects. Eur J Clin Pharmacol 36 Suppl:A 177

Lemmer B (1991a) Circadian rhythms and drug delivery. J Control Release 16:63–74

Lemmer B (1991b) Implications of chronopharmacokinetics for drug delivery: antiasthmatics, H2-blockers and cardiovascular active drugs. Adv Drug Deliv Rev 6:83–100

Lemmer B (2005) Chronopharmacology and controlled drug release. Expert Opin Drug Deliv 2(4):667–681

Lemmer B (2006) The importance of circadian rhythms on drug response in hypertension and coronary heart disease – from mice and man. Pharmacol Ther 111(3):629–651

Lemmer B (2007a) Chronobiology and chronopharmacology of hypertension. In: White WB (ed) Clinical hypertension and vascular disease: blood pressure monitoring in cardiovascular medicine and therapeutics, 2nd edn. Humana Press, Totowa, pp 413–435

Lemmer B (2007b) Chronobiology and chronopharmacology of hypertension: importance of timing of dosing. In: White WB (ed) Blood pressure monitoring in cardiovascular medicine and therapeutics, 2nd edn. Humana Press, Totowa, pp 410–435

Lemmer B (2007c) Hypertension: do we need to consider the biological clock in drug dosing? Expert Rev Cardiovasc Ther 5(3):375–379

Lemmer B (2009) Discoveries of rhythms in human biological functions: a historical review. Chronobiol Int 26 (6):1019–1068

Lemmer B (2012a) Chronopharmakologie, 4th edn. Wiss. Verlagsges, Stuttgart

Lemmer B (2012b) Masking effect of modern technology on annual rhythms: variation in births and deaths in a single family history over five centuries. Chronobiol Int 26:1019–1068

Lemmer B (2014) Seasonal variation in nativity and mortality and in socioeconomic conditions in a single family residing at different latitudes: Germany versus South Africa. ChronoPhysiol Ther 4:59–66

Lemmer B (2017) Signal transduction and chronopharmacology of regulation of circadian cardiovascular rhythms in animal models of human hypertension. Heart Fail Clin. https://doi.org/10.1016/j.hfc.2017.05.009

Lemmer B, Bruguerolle B (1994) Chronopharmacokinetics. Are they clinically relevant? Clin Pharmacokinet 26(6):419–427

Lemmer B, Nold G (1991) Circadian changes in estimated hepatic blood flow in healthy subjects. Br J Clin Pharmacol 32:627–629

Lemmer B, Portaluppi F (1997) Chronopharmacology of cardiovascular diseases. In: Redfern PH, Lemmer B (eds) Physiology and pharmacology of biological rhythms, Handbook of experimental pharmacology, vol 125. Springer, Berlin, pp 251–297

Lemmer B, Nold G, Behne S, Becker HJ, Liefhold J, Kaiser R (1990) Chronopharmacology of oral nifedipine in healthy subjects and in hypertensive patients. Eur J Pharmacol 183:521

Lemmer B, Nold G, Behne S, Kaiser R (1991a) Chronopharmacokinetics and cardiovascular effects of nifedipine. Chronobiol Int 8(6):485–494

Lemmer B, Scheidel B, Blume H, Becker HJ (1991b) Clinical chronopharmacology of oral sustained-release isosorbide-5-mononitrate in healthy subjects. Eur J Clin Pharmacol 40:71–75

Lemmer B, Middeke M, Schaaf B, Eckes L (2008) Prescribing practices and morning blood pressure control: results of a large-scale, primary-care study conducted in Germany. J Hum Hypertens 22(4):295–297

Manfredini R et al (2012) Twenty-four-hour patterns in occurrence and pathophysiology of acute cardiovascular events and ischemic heart disease. Chronobiol Int. https://doi.org/10.3109/07420528.2012.715843

Manfredini R, Salmi R, Cappadona R, Signani F, Basili S, Katsiki N (2017) Sex and circadian periodicity of cardiovascular diseases. Heart Fail Clin. https://doi.org/10.1016/j.hfc.2017.05.008

Moore JG (1989) Circadian rhythmicity in gastric emptying, acid secretion, and mucosal damage by drugs: implications for drug therapy. In: Lemmer B (ed) Chronopharmacology – biochemical and cellular interactions, Chronopharmacology – cellular and biochemical interactions. Marcel Dekker, New York, pp 631–654

Nold G, Lemmer B (1998) Pharmacokinetics of sustained-release molsidomine after morning versus evening application in healthy subjects. Naunyn Schmiedebergs Arch Pharmacol 357:R173.

Nold G, Strobel G, Lemmer B (1998) Morning versus evening amlodipine: effect on circadian blood pressure profile in essential hypertensive patients. Blood Press Monit 3:17–25.

Pechere-Bertschi A et al (1998) Renal response to the angiotensin II receptor subtype 1 antagonist irbesartan versus enalapril in hypertensive patients. J Hypertens 16:385–393

Saito Y, Yoshida S, Nakaya N, Hata Y, Goto Y (1991) Comparison between morning and evening doses of simvastatin in hyperlipidemic subjects a double-blind comparative study. Arterioscler Thromb 11(4):816–826

Sanders SW, Moore JG, Buci KN, Bishop AL (1988) Circadian variation in the pharmacodynamic effects of intravenous ranitidine. Annu Rev Chronopharmacol 5:335–338

Semenowicz-Siuda K, Markiewicz A, Korczynska-Wardecka J (1984) Circadian bioavailability and some effects of propranolol in healthy subjects and in liver cirrhosis. Int J Clin Pharmacol Ther Toxicol 22:653–658.

Scheidel B, Lemmer B (1991) Chronopharmacology of oral nitrates in healthy subjects. Chronobiol Int 8:409–419

Smolensky MH, Hermida RC, Portaluppi F (2007a) Comparison of the efficacy of morning versus evening administration of olmesartan in uncomplicated essential hypertension. Chronobiol Int 24(1):171–181

Smolensky MH, Lemmer B, Reinberg AE (2007b) Chronobiology and chronotherapy of allergic rhinitis and bronchial asthma. Adv Drug Deliv Rev 59(9–10):852–882

Triscari J, Rossi L, Pan HY (1995) Chronokinetics of pravastatin administered in the PM compared with AM dosing. Am J Ther 2(4):265–268

Uzu T, Kimura G (1999) Diuretics shift circadian rhythm of blood pressure from nondipper to dipper in essential hypertension. Circulation 100(15):1635–1638

Wallace A, Chinn D, Rubin G (2003) Taking simvastatin in the morning compared with in the evening: randomised controlled trial. BMJ 327(7418):788

Witte K et al (1993) Cardiovascular effects, pharmacokinetics, and converting enzyme inhibition of enalapril after morning versus evening administration. Clin Pharmacol Ther 54:177–186

Space Pharmacology: How Space Affects Pharmacology

26

Virginia Wotring

Contents

Abstract

The spaceflight environment is at best very different from Earth and at worst (without the protection offered by the systems of a space suit or a spacecraft) deadly. Candidates for NASA's astronaut program must meet many requirements to be selected, but even the healthiest people in the best circumstances use medications from time to time. During their spaceflight missions, astronauts are exposed to perhaps the most extreme of all extreme environments, and there are features of the spaceflight environment that drive uses of particular medications. This chapter will provide an overview of these mission-related medication needs and will discuss the special considerations that must be given to stocking the medication supplies for use on spaceflight missions.

V. Wotring (✉)
Center for Space Medicine and Department of Pharmacology and Chemical Biology, Baylor College of Medicine, Houston, TX, USA
e-mail: Virginia.Wotring@bcm.edu

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_68

Fig. 1 Astronaut Peggy Whitson demonstrates some effects of microgravity. (Image: NASA)

The spaceflight environment is at best very different from Earth and at worst (without the protection offered by the systems of a space suit or a spacecraft) deadly. The Earth and its magnetic field offer us some protection from galactic cosmic rays, but this protection is reduced at higher altitudes. At the International Space Station in low orbit around the Earth, crew are exposed to ~ 50 times the radiation that we are on Earth, and the impacts of this exposure are not yet fully understood. In contrast, some effects of reduced gravity are more readily observable (Fig. 1). Deleterious effects of the spaceflight environment on bone and muscle first became apparent during Skylab missions, and others, like spaceflight-associated neuro-ocular syndrome, came to light recently and are still under study. The spaceflight food system is designed to withstand long-duration ambient storage and meet nutritional requirements on space flight missions, but in recent years, we've learned the importance of the microbiome for human health. It now seems possible that the preserved diet prepared for space missions coupled with the absence of items containing healthy microorganisms (like yogurt) could have negative impacts on crew health. And psychologically, the demanding workload, distance from friends and family, reduction in privacy, and potential for danger could each become stress points and may be additive (Table 1).

Table 1 Features of the space flight environment that can have physiological or psychological impacts on crew

Microgravity	Loss of bone mineral density, muscle atrophy (including heart), vestibular dysfunction
Elevated exposure to space radiation	Oxidative stress, DNA damage, possibly others
Closed environment	Preserved diet (little fresh food), elevated CO_2, little social variety, unvarying sensory input
Isolation	Removed from family, friends, natural surroundings on earth
Workload	Demanding, high visibility

Candidates for NASA's astronaut program must meet many requirements to be selected, and this includes medical and psychological evaluation. Even so, the healthiest of people in the best circumstances use medications from time to time. Even for adults who are fit and healthy, it is typical to use pain relievers to treat occasional joint or muscle pain, hypnotics to treat occasional insomnia, or remedies for periodic minor gastrointestinal disturbances like heartburn (Kaufman et al. 2002). During their spaceflight missions,

astronauts are exposed to perhaps the most extreme of all extreme environments (Table 1), and there are features of the spaceflight environment that drive uses of particular medications. This chapter will provide an overview of these mission-related medication needs (over the counter (OTC) and prescription) and will discuss the special considerations that must be given to stocking the medication supplies for use on spaceflight missions.

Medication Usage on Space Flights

On the first space flights, no one really knew what being in microgravity would feel like or what physiological disturbances would be experienced. In order to prepare, flight surgeons and pilots drew inferences from the experiences of airplane pilots, especially those who flew the highest and fastest planes, military test pilots. Flight surgeons provided early astronauts with medications for motion sickness and for pain (Fig. 2), but we have little information about when or if these were used.

Space shuttle missions were 7–16 days and packed with activities. Crew reported using every available spare minute to observe and photograph Earth and to perform their own experiments with being in microgravity (Fig. 3).

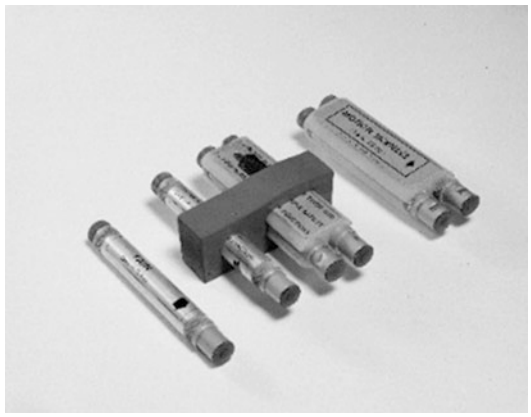


Fig. 2 Gordon Cooper carried these pre-loaded injectors for motion sickness and pain relief on his Mercury 9 flight. (Image: NASA)

After shuttle missions, astronauts were asked to complete questionnaires about the use of any medications during their mission as part of a voluntary research study. These data highlight the needs of people (even healthy ones) in the spaceflight environment. Trouble sleeping Sleep difficulty was the main driver of medication use, followed by headache, congestion, and space motion sickness (Table 2) (Putcha et al. 1999). These missions were short and intense, with a heavy workload for each crewmember. It was assumed that these medication uses could be explained mostly by the intense nature of these missions. However, a similar study conducted from medical records of International Space Station (ISS) crewmembers showed that medication usage was very similar (Table 2), even though the missions were much longer, with schedules planned to be more like a typical workdays on Earth. Taken together, these findings highlight

Fig. 3 Gerald Carr and William Pogue on Skylab 4 in 1974. (Image: NASA)

medication needs induced by the spaceflight environment itself. This environment includes not just microgravity but elevated exposure to radiation, living in a closed environmental system where all air and water are recycled, preserved food supplies, unvarying social variety, little privacy, confinement inside the spacecraft, constant exposure to noise (fans to circulate air and other machinery required to operate ISS systems), isolation from friends and family, and demanding duties in a highly visible situation.

Table 2 Indications for medication use on spaceflight missions

Space shuttle (<16 d)	ISS (> 30 d)
Sleep	Sleep
Headache	Body pain
Congestion/allergy	Congestion/allergy
Space adaptation syndrome (SAS)	Skin rash
Body pain	Space adaptation syndrome (SAS)
Gastrointestinal	Extravehicular activity (EVA)
Extravehicular activity (EVA)	Headache
	Alertness
	Gastrointestinal
	Eye
	Infection

Spaceflight Effects on Pharmacokinetics

Even after >50 years of medication use on spaceflight missions, it is not known if pharmacokinetics (how the body absorbs, distributes, metabolizes, and excretes medications) is altered by the spaceflight environment. Many astronauts have taken many medications on their missions and reported no adverse events. However, crewmembers have reported lack of desired efficacy, particularly associated with treatments for rash and space motion sickness (Fig. 4) (Wotring 2015b; Armaghani et al. 2014; Crucian et al. 2016; Putcha 2009). The use of multiple doses of sleep aids in a single night would also seem to indicate less than desired efficacy (Barger et al. 2014). Small in-flight research studies have had varying conclusions, with some showing delayed absorption or reduced peak concentrations of administered acetaminophen (Fig. 5) (Cintrón et al. 1987; Kovachevich et al. 2009), but

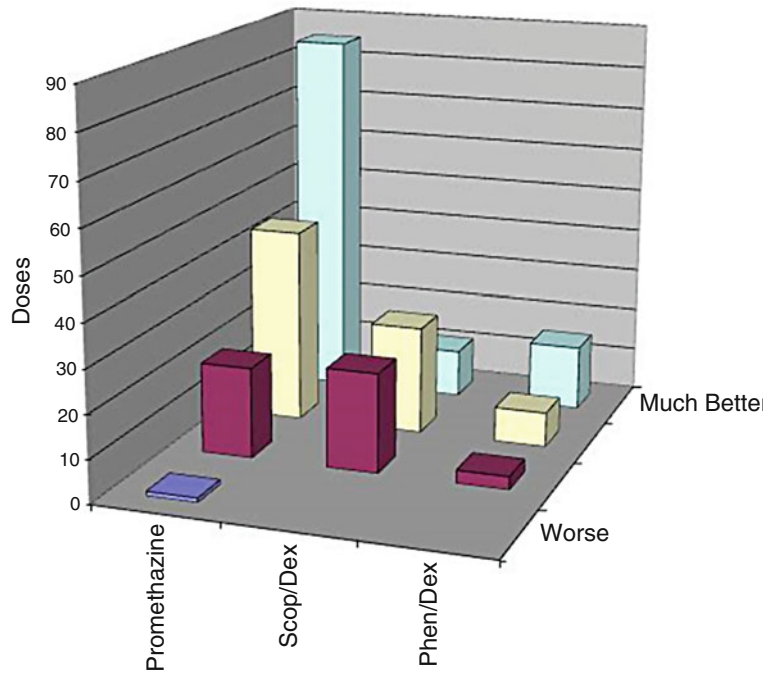


Fig. 4 Self-report relief with different nausea treatments varied considerably, with promethazine (alone, or with dexedrine (PhenDex) providing the best relief. Crewmembers reported less relief from scopolamine with dexedrine (ScopDex). (Image: NASA)

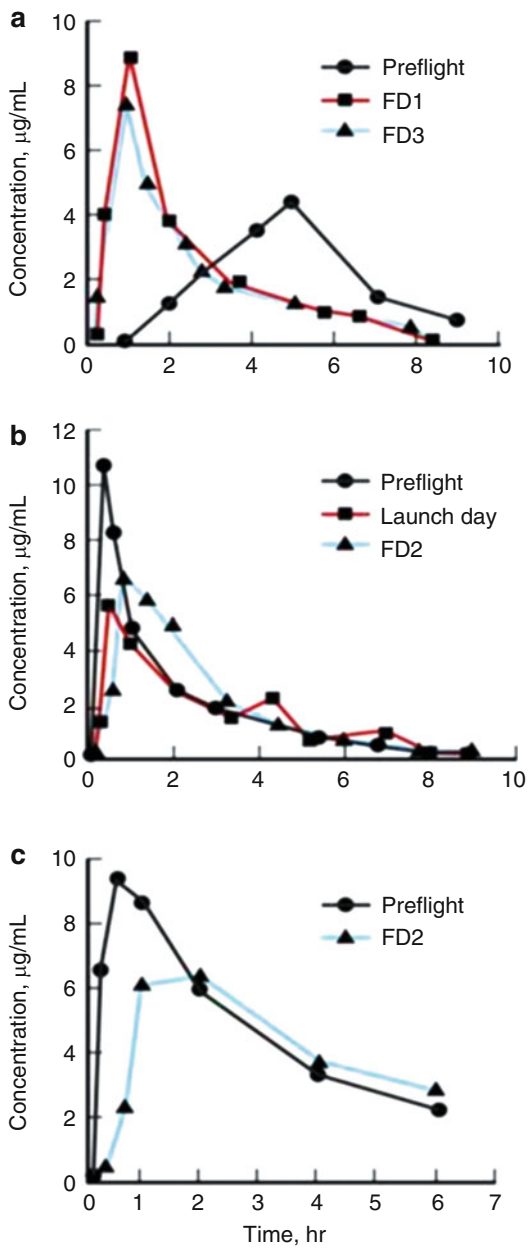


Fig. 5 Salivary concentration over time after acetaminophen dose in three different individuals (Panels A–C). (Image: NASA)

the same findings could be explained by the use of other medications, like those for nausea that can alter gastrointestinal motility. Clear evidence regarding effects (or lack thereof) of the spaceflight environment on pharmacokinetics is required for confidence in dosing strategies.

Mission-Related Needs for Medication Use

Sleep

The conditions onboard spacecraft have been less-than-ideal for sleeping. Crewmembers have noted a variety of impediments to their sleep, including noise, ambient light, and activity by nearby crewmembers (Stuster 1996). While these are certainly contributing factors, the lack of circadian cues and subsequent circadian misalignment may be an overarching cause. Humans evolved in an environment with a roughly 12 h day and a 12 h night. The hallmarks of night include darkness, cooler temperatures, and often quietness, while day usually has bright light, warmer temperatures, and sounds from diurnal animals and their activities. Space shuttles and the ISS have had constant temperatures, constant equipment noise, and relatively dim lighting. These engineering-driven features of the environment put crewmembers into a living situation almost entirely lacking the circadian cues required to set their circadian rhythms (Dijk et al. 2001; Flynn-Evans et al. 2016; Lockley et al. 2003). Recent research on light has even shown that for humans on Earth, morning light is slightly shifted toward blue wavelengths, while evening light is red-shifted. This research has led to the adoption of variable-wavelength lighting on the ISS to provide bluer light in the mornings and redder in the evenings (Fucci et al. 2005). It is expected that this will reduce the need for sleep aids on future spaceflight missions.

Currently both zaleplon and zolpidem are available for crew to use to help them fall asleep and/or stay asleep all night (Stingl et al. 2015; Barger et al. 2014). Reliance on these medications is much higher (about tenfold) than it is for healthy adults on Earth, and sometimes crewmembers have used multiple doses over the course of a night (Barger et al. 2014). Use of hypnotics with long half-lives (like the sustained release zolpidem formulation) or repeated dosing during a single night could prove dangerous in the event of a mission emergency that requires crew to awaken and perform important tasks. Because of this NASA began a study to measure the effects

of these different sleep aids on individual crewmembers before their flights. The clever experimental design allowed each crewmember to observe their reactions to the common sleep aids, from which they could make informed choices regarding use of medications in the future (Johnston 2010). There may also be a future increased reliance on administered melatonin to help set circadian rhythms, but determination of an optimal formulation, dose, and dose timing have proven difficult, possibly related to variable quality and dose in US melatonin products due to its status as a supplement rather than a medication (Arendt et al. 2008).

Bone

Loss of body weight was one of the first significant physiological findings associated with spaceflight (Zwart et al. 2014; Carpentier et al. 2018). NASA studies have determined that bones demineralized during spaceflight and muscles atrophied, both likely associated with lack of use and loading. These findings were coupled with the fact that crews tended to eat less than their bodies required (Smith and Zwart 2008). NASA has provided crew with effective countermeasures in the form of adequate calories, calcium, and vitamin D along with exercise equipment and protocols designed to prevent muscle atrophy and bone demineralization, and these countermeasures have proven effective in many crewmembers (Smith et al. 2012). As a potential alternative or supplemental treatment, a bisphosphonate osteoporosis drug has also been tested and found efficacious in maintenance of bone mineral density during spaceflight missions, particularly when combined with exercise (LeBlanc et al. 2013) (Fig. 6).

Space Adaptation Syndrome

Almost 70% of astronauts report feeling of nausea, disorientation, malaise, and headache during the first 1–3 days of a mission and again the first few days after landing. This has been described in the literature as space motion sickness (SMS) or space adaptation syndrome (SAS) (Homick

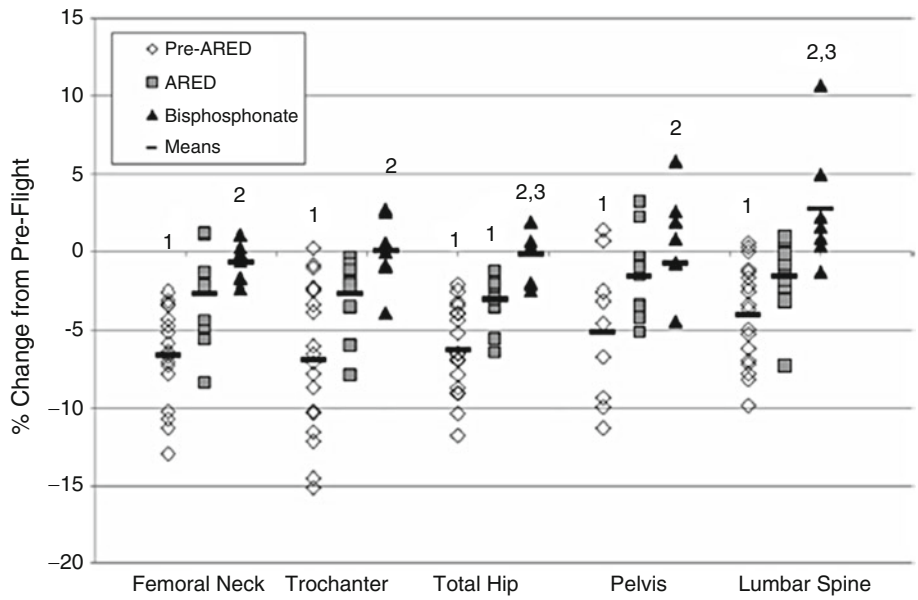


Fig. 6 Change in bone mineral density after long-duration space flight. 1 $p < 0.05$, pre versus post; 2 $p < 0.05$ (bisphosphonate group significantly different from pre-ARED); 3 $p < 0.05$ (bisphosphonate group significantly different from ARED). Pre-ARED ($n = 18$); ARED ($n = 11$); bisphosphonate ($n = 7$) (LeBlanc et al. 2013)

1979). The underlying cause is thought to be a mismatch between sensory input from the visual system (unaffected by microgravity) and the vestibular and proprioceptive systems, which are greatly affected by microgravity (Oman 1998). A variety of medications have been used to treat or prevent nausea and vomiting including various antihistamines and anticholinergics, proven treatments for motion-induced illnesses on Earth (Davis et al. 1988). Scopolamine and promethazine have been the most commonly used antinausea treatments on space mission, sometimes accompanied by dextroamphetamine to reduce lethargy (Jennings 1998). Modafinil, with its improved safety, has now replaced dextroamphetamine as the stimulant of choice in many situations (Estrada et al. 2012).

Antinausea therapy was recently revolutionized by the development of $5HT_3$ receptor antagonists like ondansetron, but even though these drugs are remarkably effective at relief of chemically induced nausea and vomiting, the $5HT_3$ receptor antagonists have proven ineffective in the treatment of motion-induced nausea (Hershkovitz et al. 2009). Evidently, the mechanisms underlying the nausea and vomiting induced by motion and that induced by chemicals (cancer chemotherapy or pregnancy) are sufficiently different that different classes of drugs are best for treatment. Numerous studies have also attempted to identify characteristics of crew who are most susceptible, but no useful correlations have been identified (Davis et al. 1988; Thornton 2011). Currently promethazine is the most commonly used treatment on ISS missions, and its use is generally limited to the first few days of a mission.

Pain

Crewmembers report several reasons for using pain relievers, including headaches, muscle pain, and joint pain (Wotring 2015b). Headaches may be associated with headward fluid shifts, especially during the microgravity adaptation period or with elevated carbon dioxide levels on the ISS (Law et al. 2014). Muscle and joint pain are sometimes reported after strenuous workouts (Fig. 7) or after extravehicular activity (EVA) in a pressurized suit (Wotring 2015b).

Ibuprofen is the most commonly chosen pain reliever among crewmembers, but aspirin, acetaminophen, and prescription pain relievers are also available. (Stingl et al. 2015).

Developing Needs for Longer Duration Missions

As spaceflight missions change in purpose, destination, or equipment, new needs are identified.

Fig. 7 Sunita Williams on the ISS treadmill. Aerobic exercise is used to maintain aerobic fitness and muscle. A harness must be used with the treadmill when in microgravity. (Image: NASA)

Several physiological effects of spaceflight are linked to duration of exposure. Several physiological systems seem to adapt to normal or possibly a new "space normal" in the early part of a mission; these include sensorimotor and renal (fluid shifts). Other needs are emerging.

Spaceflight-associated neuro-ocular syndrome (SANS) was identified fairly recently, as crewmembers routinely stayed on the ISS for 6 months or more (Mader et al. 2011). Its etiology is not yet understood, but it may include alterations in intraocular or cerebrospinal fluid pressures. The syndrome shares some features with benign intracranial hypertension, for which acetazolamide is a treatment (Lawley et al. 2017; Lee et al. 2018; Zhang and Hargens 2018). However, since acetazolamide use increases the risk of renal stone by 3% (Ahlstrand and Tiselius 1987), the use of this medication in persons at risk of renal stones (due to bone remodeling) is not recommended. This syndrome therefore needs to be better defined before safe and sensible treatments can be suggested. Regular eye exams are now part of ISS missions (Fig. 8).

Crewmembers on ISS missions are exposed to radiation 0.3 mGy/day (Wilson 2000; Hu et al. 2009), and it is expected that those on a mission to Mars will be exposed to about three times as much or 0.5–1 Gy total on a 2.5 year mission (Wilson 2000) (Zeitlin et al. 2013). Radiation can cause physiological damage, although the body has mechanisms to repair some damage (Zeitlin et al. 2013). Currently, NASA limits crew exposure to radiation to safeguard astronaut health (Cucinotta et al. 2010), but studies to optimize the body's repair mechanisms or to provide pharmacological support are underway. It is anticipated that pharmacological countermeasures to protect the body from damage or to repair radiation-induced damage will be available by the time long-distance missions into deep space begin (Wotring 2012).

Astronauts train on Earth before their missions for the duties that are planned for them to perform during their missions. For long-duration missions, there could be a time delay of years in between training and eventual in-flight need for that training. It might be optimal to maximize crewmembers' ability to recall learned material or otherwise enhance their cognitive abilities. There are drugs in development and testing for the treatment of memory impairment in dementia patients or others suffering from a loss of cognitive ability (Fond et al. 2015). Cognition enhancers, nootropics, and other agents that act on the central nervous system are in testing for dementia patients, although the use of these drugs in healthy individuals has not been tested. It is not yet known if any of these agents would improve memory or cognitive function in healthy, high-performing individuals.

Fig. 8 Karen Nyberg performs a self-examination of the eye during her ISS mission. Additional data regarding the health and function of the eye is required since spaceflight-associated neuro-ocular syndrome has been reported. (Image: NASA)

Mission-Related Requirements of the Formulary

There are multiple factors that need to be considered when selecting medications for a space mission formulary, and sometimes they are conflict with one another or with other mission needs. Availability of storage (mass and volume) is of concern on nearly every space mission but will likely be even more constrained on missions with more crew, longer duration, or significant equipment to carry. Missions may have different anticipated medical needs, mostly based on the tasks planned for the mission. For example, missions that involve landing on a planetary body with partial gravity carries added risks of falls and bone fractures, while missions where crew stay onboard a space craft have very low risk of falls or fractures. For each mission, the anticipated needs must be evaluated, and the desired drugs chosen to fit the mass and volume allotted for medication storage.

Selection of Medications for Missions

Ideal medications for space missions are well-tolerated by most people, have few, minor untoward effects, have multiple therapeutic indications, and have shelf lives that exceed the mission duration in the storage conditions available on the mission. Drug selections for the first spaceflight missions were based on knowledge from military flight surgeons and included items feasible for self-administration by persons without medical training. Now, many crew requirements are assessed based on knowledge from earlier space missions; in some cases, this has included pre- and post-mission inventories that showed the actual numbers of unit doses consumed. Going forward, accurate provisioning may be aided by two new in-flight activities: real-time inventory tracking of medical consumables (Zoldak 2016) and crew self-reports of medication use (Wotring 2015a).

The World Health Organization's Model List of Essential Medicines is an expertly vetted list of medications meant to supply every possible medical need and is updated regularly to reflect newly developed drugs (WHO 2017). This can serve as an excellent model upon which to base formulary development for provisioning any clinic. Crewmembers don't require medications to treat tropical diseases or serious conditions like AIDS, and current mission durations obviate the need for cancer treatments. Therefore, entire sections of WHO Essential Medicines List are unnecessary for spaceflight crew. For indications like pain or infection, the list identifies a selection of useful drugs while categorizing those drugs that are members of a related class. This permits selection of a single member of most medication classes for use in a small spaceflight formulary. Multiple anti-infectives are carried on spaceflight missions, each with a different mechanism of action, and multiple drugs are carried to treat the most common indications, like pain relief and sleep aids. Small supplies of medications to treat unpredictable emergencies are also included.

Predicted Medical Needs

NASA has employed modeling methods to predict which medical conditions might occur on space missions and at what frequency. The models use known medical events from past space missions and those from scenarios that are considered to be reasonable analogs, like Antarctic exploration missions and submarine or other military missions. This Integrated Medical Model (IMM) can be run for different types of spaceflight missions, varying the duration, crew composition and numbers, and activities like planetary landing or EVA (Minard et al. 2011). The output of the IMM can then be used to generate a list of medical conditions anticipated for a specific mission (Watkins et al. 2011). The mission-specific medical condition list can, in turn, be used to guide selections of medications and other medical supplies for the mission.

Stability of Medications in the Unusual Environment of Spaceflight Missions

Medications can become ineffective as they degrade over time, which reduces their efficacy.

Fig. 9 NASA pharmacists prepare streamlined medication kits based on intended use, separating commonly used medications from medications that are part of research studies, as well as from specialized packs for emergencies. The outer fabric offers good protection from light, but not from oxygen or moisture. (Image: NASA)

Additionally, a few medications have been shown to accumulate toxic compounds as they degrade. On Earth, expiration dates can drive replacement of old stock and limit negative outcomes associated with aging medications, but on spaceflight missions to distant destinations, resupply of consumables may not be feasible. NASA estimates that a journey to Mars would require 30–36 months, which is longer than the shelf life of most FDA-approved medications (Wotring 2012) (Fig. 9).

Environmental factors like humidity, light, and oxygen drive most of the chemical reactions that degrade medications. Medication manufacturers use packaging to limit exposure of their products to the environmental factors to which each product is sensitive. Given the limitations on medical kit mass and volume, sometimes medications in solid dosage forms are repackaged from stock bottles into zip-top polyethylene bags, which offer little protection from the environment but significantly reduce volume (Putcha et al. 2016).

Currently, it is not known if space radiation alters medication degradation. Preliminary efforts to evaluate this possibility have each been limited, and it is not possible to make a firm conclusion with the current data. In a survey of ~ 20% of the medications in the ISS medical kit, NASA researchers reported that six exhibited significant degradation of active ingredient during storage on the ISS for up to 28 months, and others showed changes in physical appearance indicating degradation. It important to note, however, that only a single sample of each medication was analyzed (albeit in triplicate); therefore, definitive conclusions cannot be drawn (Du et al. 2011). A later study took advantage of 550-day-old medical supplies onboard the ISS that were expired and scheduled for destruction. Samples were instead returned to Earth for analysis. Nine medications were chemically analyzed, and three failed to meet current standards for purity. Unfortunately, this opportunistic study suffers from the absence of lot-matched controls, aged for the same time period on Earth (Wotring 2016). Given the lack of clear evidence, no conclusions can yet be drawn regarding the possible effects of the spaceflight environment on medication degradation. Additional research is required before long-duration spaceflight missions commence.

Possibilities for *In Situ* Drug Analysis and Production

NASA has sponsored development of Raman spectroscopy-based analysis of finished drug products, so that astronauts may one day test unit doses immediately before use on long-

duration space missions (Shende et al. 2014). The device relies on a library of stored spectra from safe medications and from those that have degraded. It compares the spectrum it measures to this library to evaluate safety and active ingredient content.

It may be hard to imagine producing a unit dose of a medication (possibly containing only milligrams of active ingredient) in an environment where a balance can't be used to weigh ingredients. Multiple research and development groups are actively pursuing the possibility of medication production *in situ*; on our planet, this could alleviate global drug shortages and help provide adequate medication supplies in places where distribution or storage is problematic. With the advent of the first 3D-printed medication (FDA 2017), the possibility of manufacturing medication doses during missions seems more practical than it once did (Alhnan et al. 2016). Genetic engineering techniques may also permit the use of microorganisms as manufacturing plants for active ingredients (Keasling 2010). These new innovations may one day permit crew to manufacture fresh medications as needed during their missions, addressing any issues with medication stability over time as well as ensuring that medication supply meets medical needs.

Personalized Medicine for Astronauts

Personalized or precision medicine has made great improvements in efficiently matching patients with optimal treatment regimens for them, particularly in the treatment of many different cancers. As more is learned regarding genetic sequences and corresponding responses to particular medications, we expect that many more treatment regimens can be personalized based on genetic testing. On Earth, cancer treatments have driven pharmacogenomics, and astronauts have not needed these types of medications on missions. However, it is expected that genetic tests will be soon be developed to optimize treatment plans for many other indications; some of these are likely to include indications of interest to spaceflight crew (Chancellor et al. 2014; Goel and Dinges 2012).

References and Further Reading

Ahlstrand C, Tiselius HG (1987) Urine composition and stone formation during treatment with acetazolamide. Scand J Urol Nephrol 21(3):225–228

Alhnan MA, Okwuosa TC, Sadia M, Wan KW, Ahmed W, Arafat B (2016) Emergence of 3D printed dosage forms: opportunities and challenges. Pharm Res 33(8):1817–1832. https://doi.org/10.1007/s11095-016-1933-1

Arendt J, Van Someren EJ, Appleton R, Skene DJ, Akerstedt T (2008) Clinical update: melatonin and sleep disorders. Clin Med 8(4):381–383

Armaghani SJ, Crucian GP, Heilman KM (2014) The influence of emotional faces on the spatial allocation of attention. Brain Cogn 91:108–112. https://doi.org/10.1016/j.bandc.2014.09.006

Barger LK, Flynn-Evans EE, Kubey A, Walsh L, Ronda JM, Wang W, Wright KP Jr, Czeisler CA (2014) Prevalence of sleep deficiency and use of hypnotic drugs in astronauts before, during, and after spaceflight: an observational study. Lancet Neurol 13(9):904–912. https://doi.org/10.1016/S1474-4422(14)70122-X

Carpentier W, Charles JB, Shelhamer M, Hackler AS, Johnson TL, Domingo CMM, GBI S, Wotring VE (2018) Biomedical findings from NASA's project mercury: a case series. npj Microgravity 4(6). https://doi.org/10.1038/s41526-018-0040-5

Chancellor JC, Scott GB, Sutton JP (2014) Space radiation: the number one risk to astronaut health beyond low earth orbit. Life (Basel) 4(3):491–510. https://doi.org/10.3390/life4030491

Cintrón NM, Putcha L, Chen Y-M, Vanderploeg JM (eds) (1987) In-flight salivary pharmacokinetics of scopolamine and dextroamphetamine. In: Results of the life sciences DSOs conducted aboard the space shuttle 1981–1986. Space Biomedical Research Institute, Johnson Space Center, Houston

Crucian B, Johnston S, Mehta S, Stowe R, Uchakin P, Quiriarte H, Pierson D, Laudenslager ML, Sams C (2016) A case of persistent skin rash and rhinitis with immune system dysregulation onboard the international Space Station. J Allergy Clin Immunol Pract 4(4):759–762., e758. https://doi.org/10.1016/j.jaip.2015.12.021

Cucinotta FA, Kim MHY, Chapell LJ (2010) Space radiation cancer risk projections and uncertainties – 2010. vol TP-2011-216155 NASA TP-2011-216155. NASA Center for AeroSpace Information, Hanover

Davis JR, Vanderploeg JM, Santy PA, Jennings RT, Stewart DF (1988) Space motion sickness during 24 flights of the space shuttle. Aviat Space Environ Med 59(12):1185–1189

Dijk DJ, Neri DF, Wyatt JK, Ronda JM, Riel E, Ritz-De Cecco A, Hughes RJ, Elliott AR, Prisk GK, West JB, Czeisler CA (2001) Sleep, performance, circadian rhythms, and light-dark cycles during two space shuttle

flights. Am J Phys Regul Integr Comp Phys 281(5): R1647–R1664
Du B, Daniels V, Vaksman Z, Boyd J, Crady C, Putcha L (2011) Evaluation of physical and chemical changes in pharmaceuticals flown on space missions. AAPS J 13(2):299–308
Estrada A, Kelley AM, Webb CM, Athy JR, Crowley JS (2012) Modafinil as a replacement for dextroamphetamine for sustaining alertness in military helicopter pilots. Aviat Space Environ Med 83(6):556–564
FDA (2017) CDER researchers explore the promise and potential of 3D printed pharmaceuticals. Food and Drug Administration https://www.fda.gov/Drugs/NewsEvents/ucm588136.htm. Accessed 4 Apr 2018
Flynn-Evans EE, Barger LK, Kubey AA, Sullivan JP, Czeisler CA (2016) Circadian misalignment affects sleep and medication use before and during spaceflight. npj Microgravity 2. https://doi.org/10.1038/npjmgrav.2015.19
Fond G, Micoulaud-Franchi JA, Brunel L, Macgregor A, Miot S, Lopez R, Richieri R, Abbar M, Lancon C, Repantis D (2015) Innovative mechanisms of action for pharmaceutical cognitive enhancement: a systematic review. Psychiatry Res 229(1–2):12–20. https://doi.org/10.1016/j.psychres.2015.07.006
Fucci RL, Gardner J, Hanifin JP, Jasser S, Byrne B, Gerner E, Rollag M, Brainard GC (2005) Toward optimizing lighting as a countermeasure to sleep and circadian disruption in space flight. Acta Astronaut 56(9–12):1017–1024
Goel N, Dinges DF (2012) Predicting risk in space: genetic markers for differential vulnerability to sleep restriction. Acta Astronaut 77:207–213. https://doi.org/10.1016/j.actaastro.2012.04.002
Hershkovitz D, Asna N, Shupak A, Kaminski G, Bar R, Tal D (2009) Ondansetron for the prevention of seasickness in susceptible sailors: an evaluation at sea. Aviat Space Environ Med 80(7):643–646
Homick JL (1979) Space motion sickness. Acta Astronaut 6(10):1259–1272
Hu S, Kim MH, McClellan GE, Cucinotta FA (2009) Modeling the acute health effects of astronauts from exposure to large solar particle events. Health Phys 96(4):465–476. https://doi.org/10.1097/01.HP.0000339020.92837.61
Jennings RT (1998) Managing space motion sickness. J Vestib Res 8(1):67–70. https://doi.org/10.1016/S0957-4271(97)00042-6
Johnston S (2010) Operational ground testing protocol to optimize astronaut sleep medication efficacy and individual effects. https://taskbook.nasaprs.com/publication/index.cfm?action=public_query_taskbook_content&TASKID=8225. Accessed 4/3/2018
Kaufman DW, Kelly JP, Rosenberg L, Anderson TE, Mitchell AA (2002) Recent patterns of medication use in the ambulatory adult population of the United States: the Slone survey. JAMA 287(3):337–344
Keasling JD (2010) Manufacturing molecules through metabolic engineering. Science 330(6009):1355–1358. https://doi.org/10.1126/science.1193990
Kovachevich I, Kondratenko S, Starodubtsev AK, Repenkova LG (2009) Pharmacokinetics of acetaminophen administered in tablets and capsules under long-term space flight conditions. Pharm Chem J 43(3):130–133
Law J, Van Baalen M, Foy M, Mason SS, Mendez C, Wear ML, Meyers VE, Alexander D (2014) Relationship between carbon dioxide levels and reported headaches on the international space station. J Occup Environ Med 56(5):477–483. https://doi.org/10.1097/jom.0000000000000158
Lawley JS, Petersen LG, Howden EJ, Sarma S, Cornwell WK, Zhang R, Whitworth LA, Williams MA, Levine BD (2017) Effect of gravity and microgravity on intracranial pressure. J Physiol 595(6):2115. https://doi.org/10.1113/jp273557
LeBlanc A, Matsumoto T, Jones J, Shapiro J, Lang T, Shackelford L, Smith SM, Evans H, Spector E, Ploutz-Snyder R, Sibonga J, Keyak J, Nakamura T, Kohri K, Ohshima H (2013) Bisphosphonates as a supplement to exercise to protect bone during long-duration spaceflight. Osteoporos Int 24(7):2105–2114. https://doi.org/10.1007/s00198-012-2243-z
Lee AG, Mader TH, Gibson CR, Brunstetter TJ, Tarver WJ (2018) Space flight-associated neuro-ocular syndrome (SANS). Eye (Lond). https://doi.org/10.1038/s41433-018-0070-y
Lockley SW, Brainard GC, Czeisler CA (2003) High sensitivity of the human circadian melatonin rhythm to resetting by short wavelength light. J Clin Endocrinol Metab 88(9):4502–4505. https://doi.org/10.1210/jc.2003-030570
Mader TH, Gibson CR, Pass AF, Kramer LA, Lee AG, Fogarty J, Tarver WJ, Dervay JP, Hamilton DR, Sargsyan A, Phillips JL, Tran D, Lipsky W, Choi J, Stern C, Kuyumjian R, Polk JD (2011) Optic disc edema, globe flattening, choroidal folds, and hyperopic shifts observed in astronauts after long-duration space flight. Ophthalmology 118(10):2058–2069. https://doi.org/10.1016/j.ophtha.2011.06.021
Minard CG, de Carvalho MF, Iyengar MS (2011) Optimizing medical resources for spaceflight using the integrated medical model. Aviat Space Environ Med 82(9):890–894
Oman CM (1998) Sensory conflict theory and space sickness: our changing perspective. J Vestib Res 8(1):51–56
Putcha L (2009) Data mining – pharmacotherapeutics of space motion sickness
Putcha L, Berens KL, Marshburn TH, Ortega HJ, Billica RD (1999) Pharmaceutical use by U.S. astronauts on space shuttle missions. Aviat Space Environ Med 70(7):705–708
Putcha L, Taylor P, Daniels V, Pool S (2016) Clinical pharmacology and therapeutics. In: Nicogossian AE (ed) Space physiology and medicine: from evidence to practice, 4th edn. Springer, New York p 323
Shende C, Smith W, Brouillette C, Farquharson S (2014) Drug stability analysis by Raman spectroscopy.

Pharmaceutics 6(4):651–662. https://doi.org/10.3390/pharmaceutics6040651
Smith SM, Zwart SR (2008) Nutritional biochemistry of spaceflight. Adv Clin Chem 46:87–130
Smith SM, Heer MA, Shackelford LC, Sibonga JD, Ploutz-Snyder L, Zwart SR (2012) Benefits for bone from resistance exercise and nutrition in long-duration spaceflight: evidence from biochemistry and densitometry. J Bone Miner Res 27(9):1896–1906. https://doi.org/10.1002/jbmr.1647
Stingl JC, Welker S, Hartmann G, Damann V, Gerzer R (2015) Where failure is not an option -personalized medicine in astronauts. PLoS One 10(10):e0140764. https://doi.org/10.1371/journal.pone.0140764
Stuster JW (1996) Bold endeavors: behavioral lessons from polar and space exploration. Naval Institute Press, Annapolis
Thornton WE (2011) A rationale for space motion sickness. Aviat Space Environ Med 82(4):467–468
Watkins S, Barr Y, Kerstman E (2011) The space medicine exploration medical conditions list. In: 18th IAA humans in space symposium. IAA, Houston
WHO (2017) World health organization model list of essential medicines. http://www.who.int/medicines/publications/essentialmedicines/en/. Accessed 4 Apr 2018, 20th list
Wilson JW (2000) Overview of radiation environments and human exposures. Health Phys 79(5):470–494
Wotring VE (2010) Evidence Book: Pharmacology Risk Report National Aeronautics and Space Administration, HRP-47060, Lyndon B. Johnson Space Center Houston, Texas
Wotring VE (2012) Space pharmacology. Springer briefs in space development. Springer, New York
Wotring V (2015a) Dose tracker application for monitoring medication usage, symptoms, and adverse effects during missions. NASA Johnson Space Center. https://www.nasa.gov/mission_pages/station/research/experiments/1933.html. Accessed 4 Apr 2018
Wotring VE (2015b) Medication use by U.S. crewmembers on the international Space Station. FASEB J 29(11):4417–4423. https://doi.org/10.1096/fj.14-264838
Wotring VE (2016) Chemical potency and degradation products of medications stored over 550 earth days at the international space station. AAPS J 18(1):210–216. https://doi.org/10.1208/s12248-015-9834-5
Zeitlin C, Hassler DM, Cucinotta FA, Ehresmann B, Wimmer-Schweingruber RF, Brinza DE, Kang S, Weigle G, Bottcher S, Bohm E, Burmeister S, Guo J, Kohler J, Martin C, Posner A, Rafkin S, Reitz G (2013) Measurements of energetic particle radiation in transit to Mars on the Mars science laboratory. Science 340(6136):1080–1084. https://doi.org/10.1126/science.1235989
Zhang LF, Hargens AR (2018) Spaceflight-induced intracranial hypertension and visual impairment: pathophysiology and countermeasures. Physiol Rev 98(1):59–87. https://doi.org/10.1152/physrev.00017.2016
Zoldak JT (2016) Medical consumables tracking. NASA Johnson Space Center https://www.nasa.gov/mission_pages/station/research/experiments/1259.html. Accessed 4 Apr 2018
Zwart SR, Launius RD, Coen GK, Morgan JL, Charles JB, Smith SM (2014) Body mass changes during long-duration spaceflight. Aviat Space Environ Med 85(9):897–904. https://doi.org/10.3357/asem.3979.2014

Nanotechnology in Medicine

27

M. Nikolova, R. Slavchov, and G. Nikolova

Contents

Abstract

Nanotechnology deals with materials and interactions at molecular and atomic level, sized between 0.1 and 100 nm. The small size of the nanoparticles and the changes of their physical and chemical properties compared to that of their macromolecular analogues offer many advantages for contemporary medicine, including improved drug delivery, opportunities for a better and noninvasive diagnostics, and targeted treatment with reduced adverse and systemic effects. Nonetheless, the same differences in the physical and chemical properties of nanoparticles could lead to serious and unpredictable side effects for the human body and for the global ecosystem, including accumulation, recirculation, and inflammatory, mutagenic, and oncogenic potential. The current review is focused on the current uses, benefits, disadvantages, and risks of nanotechnology in medicine.

. . .our environment, and I mean our man-made world of machines, artificial constructs, computers,

M. Nikolova (✉)
University Hospital Alexandrovska, Clinic of Nephrology, Medical University, Sofia, Bulgaria
e-mail: milena_i_dani@abv.bg

R. Slavchov
Cambridge University, Cambridge, UK

G. Nikolova
University Hospital Acad I. Penchev, Medical University, Sofia, Bulgaria

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_45

> electronic systems, interlinking homeostatic components - all this is in fact beginning more and more to possess what the earnest psychologists fear the primitive sees in his environment: animation. In a very real sense our environment is becoming alive, or at least quasi-alive, and in ways specifically and fundamentally analogous to ourselves. . .
>
> Philip K. Dick. The Android and the Human. SF Commentary #31, December 1972 (delivered as a speech at the University of British Columbia, Vancouver, Canada, February, 1972). https://genius.com/Philip-k-dick-the-android-and-the-human-annotated

Introduction

Nanotechnology is a relatively novel interdisciplinary field of modern science, on the borders between physics, chemistry, biology, and engineering sciences, which deals with materials and interactions at molecular or even atomic level – sized 0.1–100 nm (1 nm = 1×10^{-9} m). Due to their tiny size, these materials are known to have different physical and chemical properties than the larger-scale ones (magnetism, conductance, optical properties, chemical interactions, etc.) (Nikalje 2015). At these tiny-scale sizes, the surface of the material changes significantly, along with its optical, chemical, and magnetic properties, molecular recognition differs much from that on macro-level, the quantum effects may dominate over that of the classical physics, and new and unpredictable effects might appear (Idrees 2015). Their applications in the field of medicine (for diagnostic, therapeutic, and prophylactic purposes) have brought revolutionary successes in many fields, such as cancer and diabetes treatment, anti-infectious agents, vaccine delivery; cell cultures, biosensors, diagnostic tests, tissue regeneration, and nano-robots for repairing and/or replacing cell structures and genetic engineering; etc. (Emerich and Thanos 2003; Gupta and Dinda 2018; Surendiran et al. 2009).

Many natural nanoparticles occur around us, even in our food, such as oil bodies, casein micelles, cellular organelles, and even viruses (McClements and Xiao 2017). They are formed by molecular conjugates. When entering the human body, these natural nanoparticles undergo degradation, or (like viruses) they can affect the living cell and the cell will respond. On the other hand, the man-made nanoparticles represent a new invader for our body, and the cells still have no specific adaptive response against them. The interactions between the cell structures and the synthetic nanoparticles are not well understood, and many unanswered questions remain. Therefore, many ethical and safety issues in the use of nanotechnology in medicine are to be solved.

It has been considered that the nano-era began back in 1959 with the first ideas of R. Feynman on the possibilities to change the properties of the known matter by manipulating the individual atoms, i.e., the "top-to-bottom, or top-down, approach" (Vijayakumar et al. 2013). Nonetheless, these ideas emerged in science fiction long before the plenary lecture of R. Feynman, "There's plenty of room at the bottom: an invitation to enter a new field of physics," delivered at the annual meeting of the American Physical Society in 1959 (Feynman 1959). In the first half of the twentieth century, several science fiction novels pictured the use of subatomic devices for investigations, and in 1954 the great visionnaire Philip K. Dick in his novel "Autofac" described tiny self-assembling and self-reproducing robots (Milbum 2010).

The word "nanotechnology" was introduced by the Japanese scientist Norio Taniguchi in 1974 on a Conference on Production Engineering to describe semiconductor processes – "processing of, separation, consolidation, and deformation of materials by one atom or one molecule" (N Taniguchi Norio (1974). "On the basic concept of 'nano-technology.'" Proceedings of the International Conference on Production Engineering, Tokyo, 1974, Part II. Japan Society of Precision Engineering).

In 1980, E. Drexler first encountered Feynman's provocative 1959 talk "There's plenty of room at the bottom: an invitation to enter a new field of physics." And in 1986 he independently used the term "nanotechnology" in his book "Engines of creation" (Drexler 1986). In this book E. Drexler suggested the possibility that a nano-sized "assembler" could be able to build a self-copy and copies of other structures of a different complexity and size, including larger structures – or "bottom-up approach" (Vijayakumar et al. 2013). Therefore, Drexler's idea of nanotechnology is frequently

called "molecular manufacturing." Nonetheless, this idea has already been published in science the fiction novel "Autofac" by P.K. Dick 30 years ago (Dick 1955).

A major breakthrough in the development of nanotechnology was the invention of the first scanning tunneling microscope in 1981. This microscope allows imaging surfaces at the atomic level, and the inventors (G. Binning and H. Rohrer) were awarded Nobel Prize in Physics in 1986 for their discovery. Other major discoveries in the field of nanotechnology are the discovery of fullerenes by J. Kroto, R. Smalley, and R. Curl in 1985 and the discovery of carbon nanotubes in 1991 by S. Iijima. In 1999 the first book on nanomedicine ("Nanomedicine" by R. Freitas) was published.

During the first decade of the twenty-first century, 3D nanosystems, networks, and active nanosystems have been developed, and a great concern of the safety of nano-materials and nano-robots has arisen.

Other serious ideas that first arouse in science fiction literature, but nowadays are becoming more and more real, intimately related to the concept of nanotechnology, are the "transhumanism" (graduate merge of humans and machines producing new and more sophisticated creatures that would move humanity to the "next level") and "singularity" (a certain point at which a given mathematical object is not defined or its behavior becomes unpredictable – i.e., not differentiable or infinite; from anthropological and social point of view, this is the probability that at certain point of time, the development of technology will lead to the emergence of unpredictable artificial intelligence with possibly harmful for humanity effects). These ominous predictions made by the science fiction writers more than half a century ago are becoming more and more real today, as the borders between the human and the machine become more and more obscure.

Types of Nano-materials with Proven or Potential Benefits in the Field of Medicine

Different classification systems have been applied to nano-materials according to their size, shape, phase composition, etc. (Nikalje 2015). All nanoparticles have specific physical characteristics (shape, size, surface area, permeability, magnetism) that influence deeply their properties and make them very different from other larger-scale materials (Ventola 2012a). The size of all nanoparticles, by definition, is between 0.1 and 100 nm. This small size allows them to pass through biological membranes and improve drug delivery, improves their solubility, and increases their bioavailability and circulation time within the living body. It also increases the total surface of the administered material (i.e., increases the reactivity, solubility, and bioavailability, increases conjugation properties with electrostatic surfaces or other larger molecules of nanoparticles. Still, the increased permeability and bioavailability of nanoscale materials carries the risk of penetration and accumulation within certain tissues and cells and oncogenesis. Nanoparticles can have different shapes (tubes, spheres, hemispheres, cones, wires, discs, cylinders, etc.) and be solid, porous, or hollow. Therefore, nanoparticles can transport substances to and within the living cell (e.g., for targeted drug delivery, imaging, regeneration, and repair).

Nanoparticles possessing magnetic properties can be used as targeted carriers of biologically active substances as directed by a magnetic field, as MRI contrast media, or for targeted tumor lysis mediated by their magnetocaloric effect (Guo et al. 2018).

Several major types of nanoparticles suitable for medical use with specific properties have been described, including metallic nanoparticles, dendrimers, liposomes, nano-somes, nano-pores, carbon nanotubes, fullerenes, nanocrystals/quantum dots, nano-shells, nano-wires, nano-bubbles, paramagnetic nanoparticles, and nano-bombs; these structures may be combined to form nano-robots and even structures mimicking living cells (i.e., respirocytes and microbivores) (Emerich and Thanos 2003; Guo et al. 2018; Nikalje 2015; Sarfaraz et al. 2018; Surendiran et al. 2009; Vijayakumar et al. 2013).

The major types of nano-materials used in contemporary medicine are presented in Table 1.

Table 1 Major types of nano-materials used in contemporary medicine (Bhattacharya et al. 2012; Emerich and Thanos 2003; Guo et al. 2018; Gupta and Dinda 2018; Idrees 2015; Sarfaraz et al. 2018; Surendiran et al. 2009; Ventola 2012a, b; Vijayakumar et al. 2013; Wani and Kothari 2018)

Material	Properties	Fields of medical application
Liposomes	Biodegradable, biocompatible, and nonimmunogenic microvesicles	Drug carriers for both hydrophilic and hydrophobic agents (including for gene delivery)
Micelles and solid lipid nanoparticles	Nanoparticles of a "larger" size, containing natural biomolecules that form carrier structures allowing targeted transportation of pharmaceuticals	Targeted treatment Possible genetic transfer
Nano-shells	Tiny spheres coated with different materials. They can transport different agents to the living cell (treatment/prophylaxis and imaging) or absorb infrared radiation and, after being deposited within the tumor cells, give away the heat and burn the tumor without affecting the adjacent normal living cells Gold nano-shells being most extensively studied	After conjugation to antibodies, oligonucleotides, targeting ligands, polymers, or other therapeutic agents, they can transport medications and vaccines After conjugation contrast agents or isotopes, they can assist imaging tumors or other lesions
Quantum dots (nanocrystals)	Small light-emitting nanocrystals (e.g., Cd/Zn selenides) with semiconductive properties (2–10 nm corresponding to 10–5 atoms). Obey quantum laws of quantum confinement. Glow when stimulated by UV light (the wavelength depending on the size of the crystal). More stable than traditional dyes – i.e., they can be used to replace biopsy investigations In combination with MRI, they can produce images of tumors	Imaging of specific DNA fragments Imaging of tumor cells and whole tumors Imaging of tumor vascularization "Staining" of cells and tissue structure, as a parallel to biopsy Ultrasound contrast medium
Dendrimers	Branched polymers, biocompatible and biodegradable, used as nano-sized containers and transporters. Consist of a central "core" (container), branches (interior dendritic, or tree-like, structure), and surface with functional groups on it	Targeted drug delivery (i.e., in anticancer treatment) and delivery of contrast agents in MRI Gene delivery Biosensors
Fullerenes	Carbon-based carrier nanostructures in the form of hollow spheres, tubes, ellipsoids, etc. Stable to heat, superconductive properties. Inorganic fullerenes – similar to carbon-containing – but lack C-atoms; contain MoS_2, WS_2, TiS_2, NbS_2; stable to weight	Drug delivery systems
Nano-wires	Wires 40–50 nm in diameter, as long as desired Superconducting (YBCO), semiconducting (Si nano-wires, GaN, InP), metallic (Ni, Pt, Au), insulating $(Si)_2$, TiO_2) Molecular nano-wires – repeating molecular units (organic – DNA) or inorganic (containing Mo, S, and I)	Biosensors Functional MRI and PET scanning
Nanotubes	Carbon-based nanotubes that can: 1. Transport substances within the living cell via endocytosis 2. Carry radioisotopes to target sites in order to create an image of or destroy cancer cells 3. Absorb and emit near-infrared light and	Drug delivery systems Imaging studies Cancer treatment Imaging Nano-robotic devices Reparation and regeneration

(continued)

Table 1 (continued)

Material	Properties	Fields of medical application
	destroy tumor cells 4. Have near-infrared photoluminescence properties and visualize cancer cells 5. Be stimulated by light and move and create nano-robotic for cell and DNA reparation and regeneration	
Nano-electro-mechanical sensors (NEMS, cantilevers)	Nano-mechanical devices that detect the forces, motion, mechanical properties, and masses that emerge in biomolecular level A surface coated with bioreceptor that recognizes the target analyte and bends in response. The transducer transforms the mechanical response into recognizable electrical signal	Biosensors Biorecognition
Magnetic nanoparticles	Metal (Au, Ag, Co, Ni, Mn, Zn), metal oxide (γ-Fe_2O_3 and Fe_3O_4, $CoFe_2O_4$, $Mn0.6Zn0.4Fe_2O_4$, $Mn_3Zn7Fe_2O_4$, $Mn0.6Zn0.4Fe_2O_4$), and metal alloy (FeCo, FePt); ferrite nanoparticles have a stronger magnetism and a higher relaxation rate and are used in MRI. Nonvirulent and nonimmunogenic. Large specific surface area for carrying a large amount of DNA fragments, drugs, and modified compounds (after modification can be used as vector). Most modified magnetic nanoparticles are biocompatible. And some magnetic nanoparticles are superparamagnetic. Magnetic nanoparticles can produce thermal effect under the action of alternating magnetic field (tumor thermotherapy). Magnetic nanoparticles can be used for magnetic separation (under the action of the magnetic field, they can separate biomolecules). Most extensively studied – Fe_2O_3 and Fe_3O_4 nanoparticles	Vector for targeted treatment Medical imaging (MRI) Thermodestruction of tumors and non-tumor structures Separation of biomolecules
Nano-bombs	Nanoparticles that can destroy certain cells and structures via induction of hyperthermia or the targeted delivery of antineoplastic agents Also used for imaging – deposition within certain tissues and generation of a detectable signal after stimulation	Cancer treatment Imaging
Nano-pores	Wafer-like structures with high density of pores that allow entry of lower-molecular-weight substances (oxygen, glucose, insulin) to pass through but stop larger molecules, such as immunoglobulins. Can stop certain molecules or sequences. Can protect the graft after transplantation while assuring proper nutrition Can differentiate DNA strands based on differences in base pair sequences. Can differentiate purines from pyrimidines Improve longitudinal resolution for base pair	Post-transplantation graft protection DNA sequencing

(continued)

Table 1 (continued)

Material	Properties	Fields of medical application
	identification Can be used for genome sequencing	
Nano-bubbles	Nanoscale bubbles that are stable at room temperature and within the human body form microbubbles that can transport pharmaceutical agents and genes or perform sonothrombolysis when combined with ultrasound	Drug and gene delivery Sonothrombolysis
Respirocytes	Hypothetical nano-devices acting as erythrocytes with higher capacity to deliver oxygen and carbon dioxide, surface sensors and regulating system that alter the uptake of carbon dioxide and the release of oxygen as required	Artificial red blood cells
Microbivores	Hypothetical nano-devices acting as leukocytes, expected to have higher efficacy. Theoretically can clean the blood of bacteria in septicemia	Artificial white blood cells
Nano-robots	Nano-sized devices for dynamic follow-up and molecular manipulation of tissues, cells, and cellular structures	Diagnostics and treatment Artificial organs (?)

Interaction of Nanoparticles with Biological Molecules

Nanoparticles correspond and match with natural molecules and functional systems within the living body, and therefore they can actively interact with biological systems (Idrees 2015). As the majority of animal cells are within the range of 10–20,000 nm, nanoparticles and nano-sized devices can enter the living cell and subcellular structures (organelles) and interact with intracellular molecules, such as proteins and ribonucleic acids (DNA and RNA). Nanoscale devices can both detect and possibly cure cellular defects by correcting subcellular defects (e.g., via interaction with intracellular biopolymers such as proteins and ribonucleic acids, with receptor and enzymes) and can be used for active monitoring of cellular processes, (sub)cellular and tissue reparation and regeneration. One should not forget that at these tiny-scale sizes, the characteristics of matter change significantly with domination of quantum effects and significant changes in chemical reactivity that may lead to the occurrence of new and unpredictable effects and properties (Idrees 2015). Nanoparticles have been used as "carriers" of biologically active molecules (antibodies, vaccines, anticancer agents, etc.), where the core of the new nano-biomaterial is formed by the nanoparticles and is covered by inert material (i.e., silica) and then by a linker molecule, and finally the biologically active structures are attached. Nano-materials can also be used to cover and to bind biomolecules, and after transportation to the living cells by the biomolecules (i.e., in this case the carrier is the biomolecule), the targeted nanoparticles may detect and affect the cell. In other words, nanoparticles may be used for the detection and destruction of damaged cells (including neoplastic) after being targeted to these cells by a carrier biomolecule. Moreover, magnetic nanoparticles can be used as carriers, MRI contrast agents, and for the induction of magnetic hyperthermia (Guo et al. 2018).

Areas of Application in Medicine

As in all fields of human life, nanotechnologies have found promising new applications in medical practice – for diagnostic, therapeutic, and prophylactic purposes. Table 2 represents the major

Table 2 Nanoparticle size and applications in medicine (Idrees 2015; Nikalje 2015)

Type of nanoparticle	Size (nm)	Use
Carbon nanotubes	Diameter 0.5–3 Length 20–100	Detection of DNA changes and/or protein biomarkers
Dendrimers	<10	Controller drug release Imaging
Nanocrystals/quantum dots	2–9.5	Improve drug solubility Detect cancer cells Detect mutant DNA fragments Tumor and lymph node visualization
Nano-shells	Various	Tumor imaging Tumor ablation
Nano-wires	Various	Sensing of proteins and chemicals (biomarker and mutant DNA detection) Detection of gene expression products
Nanoparticles	10–100	Contrast agents (MRI and ultrasound) Targeted drug delivery Permeation enhancers Detection of apoptosis and angiogenesis
Liposomes, micelles, emulsions	Various	Drug delivery Liposomes are nontoxic!

uses of nanoparticles in medicine, and Table 3 presents the benefits of nano-sized materials for the contemporary medical practice.

Diagnosis

Nano-materials offer the unique opportunity of noninvasive, fast, and inexpensive in vitro and in vivo diagnostics, even in real time, for nontoxic imaging studies and for the simultaneous diagnostic and therapeutic approaches (the so-called theranostics – targeted diagnostic approach combined with immediate therapeutic intervention, for instance – visualization of tumor cells using magnetic nanoparticles with subsequent magnetism-induced thermal ablation or visualization of the tumor using antibody-bound gold nano-shells with subsequent irradiation with infrared light and thermal ablation). NEMS can be used for precise and noninvasive in vivo sensitizing. Moreover, the tiny-sized nanoparticles permit precise and fast DNA sequencing, in vivo staining and visualization, and targeted imaging of damaged cells and tissues while decreasing the need of biopsy. Therefore, nanotechnologies give us the change to perform both super-targeted screening in high-risk groups and large mass-screening programs for at population level.

The nano-wires also permit detection of pathological changes in remote and inaccessible areas, such as the central nervous system, and detection of alterations in the small blood vessels and nerve fibers that are hard to evaluate using the existing techniques. Moreover, they are capable of detecting tumor antigens located on or within remote tumor cells.

Magnetic and luminescent nanoparticles offer noninvasive in vivo imaging without the risk of renal toxicity with high sensitivity and specificity.

Probably the most impressive breakthrough based on nanotechnologies in contemporary diagnostics is that in molecular diagnostics, nanotechnologies allow detection/sequencing and reparation of DNA changes, even single-nucleotide alterations. Moreover, tiny nano-based quasi-robot systems are capable of correcting the defects. These astonishing new discoveries even allow the construction of new DNA molecules for biomedical and biocomputer researches, such as nano-sensors, nano-electronic and bioelectronic devices, and DNA computation (Zahid et al. 2013).

Table 3 Fields of application and benefits of nano-materials in medicine (Bhattacharya et al. 2012; Emerich and Thanos 2003; Guo et al. 2018; Gupta and Dinda 2018; Idrees 2015; Sarfaraz et al. 2018; Surendiran et al. 2009; Tasciotti et al. 2016; Ventola 2012a, b; Vijayakumar et al. 2013; Wani and Kothari 2018; Yadav et al. 2018; Yambe 2009; Zahid et al. 2013)

Field of application	Benefits
Diagnostics	
Imaging	Magnetic and luminescent nanoparticles for contrast enhancement Contrast-medium carriers Nano-wires allowing noninvasive detection of problems within the central nervous system and other dangerous-to-investigate sites Less invasive and nontoxic imaging. Real-time and early detection Inexpensive mass screening Theranostics
Molecular diagnostics	Nano-probes Less invasive diagnostics, giving information on molecular level with the opportunity to correct the defect Inexpensive mass screening
Treatment	Targeted anticancer treatment with antineoplastic drugs, enzymes/enzyme inhibitors, radiation, thermal ablation, gene therapy, DNA reparation/DNA transfer Lower dose, lower systemic availability, faster response Less toxic or nontoxic for the adjacent tissues/cells and the whole body Anti-infectious treatment (tuberculosis, leishmaniasis, antifungal agents) Ag-containing nanoparticles serving as antimicrobial "scaffold" Mechanisms to overcome therapeutic resistance in infections and in neoplastic diseases Neuroregeneration and reparation (Alzheimer's and Parkinson's diseases). Epilepsy – preventing seizures
Vaccines	Delivery and stabilization of vaccines (against HBV, staphylococci, etc.)
Regeneration and reparation	Wound healing – biocompatible wound healing material, Ag-containing dressings and creams/solutions, chitosan, Cu-containing nanoparticles Glutathione and collagen nanoparticles for cosmetic purposes
Nano-robotics	A nano-sized device for dynamic follow-up and molecular manipulation of tissues, cells, and cellular structures, as a parallel to viruses and bacteria living within our body Must transmit data and receive commands and should be able to repair or correct living structures or processes The simultaneous use of nano-photonic and nanotube-based technologies allows DNA manipulation, and larger and more complex nano-robots/artificial organs are expected to emerge
Transplantology and prosthetics	Targeted and controlled delivery of immunosuppressive medications, preservation of graft function, achievement and preservation of tolerance Imaging Implantable drug delivery systems Nano-engineered prosthetics Development of artificial organs

Treatment

Nanotechnologies are applied in many fields of treatment, including anti-infectious medications (against tuberculosis, fungi, and leishmania), anticancer medications, wound healing and cosmetics, artificial fabrics and tissues, vaccines, dentistry, neuroprotection and neuroregeneration, artificial cells and organs, and nanorobotics (Emerich and Thanos 2003; Gupta and Dinda 2018; Nikalje 2015; Sarfaraz et al. 2018; Surendiran et al. 2009; Tasciotti et al. 2016; Ventola 2012a, b; Vijayakumar et al. 2013; Wani and Kothari 2018; Yadav et al. 2018; Yambe 2009; Zahid et al. 2013).

Nanoparticles can be used for targeted drug delivery offering higher specificity, higher bioavailability, lower dose, and lower incidence of

both systemic and local side effects (i.e., increasing cost-effectiveness). Different types of nanoparticles are used for targeted drug delivery, including liposomes, micelles, dendrimers/nanopolymers, nano-pores, nano-shells, NEMS, etc. Still, their toxicity and safety for the human body and for the environment remain unclear, and further research is required.

Nanoparticles are widely used as carriers of anticancer drugs, such as paclitaxel and doxorubicin. Newly designed nanoparticles are called "minicells" (Nikalje 2015). They are composed of the biomembranes of mutant bacteria and carry different types of antineoplastic medications. After being engulfed by the tumor cells, these nanoparticles are capable of destroying the tumor. Nano-porous materials can also carry different medications and offer the benefit of prolonged release. Tolerogenic nanoparticles have also been developed – after loaded with certain antigens (e.g., myelin) that can be introduced to the human body and induce tolerance to this antigen, thus decreasing the autoimmune response against the same substance.

Another field is the antibiotic delivery and antibiotic resistance (Zn oxide nanoparticles can decrease antibiotic resistance and increase antibiotic effect of ciprofloxacin). Nanoparticles have been used to dissolve thrombi (both in combination with ultrasound and as carriers of anticoagulant substances). Nanoparticles have been used as carriers of nucleic acids (including gene transfer), peptides, etc.

For vaccine delivery, nanoparticles have several important advantages (especially liposomes!): they are inert, their small size permits them to cross biological barriers, and they can be targeted to specific organs. This is particularly important, especially in the development of vaccines against viral pathogens. Moreover, one should not forget that viruses themselves represent natural nanoparticles that become "alive" after entering the host cell and exploiting the host life, supporting and reproducing intracellular systems.

In neurodegenerative disorders, nanoparticles can be used for targeted drug delivery to the damaged cells or areas that otherwise remain hidden behind the blood-brain barrier. Different types of nanoparticles have been investigated, including liposomes, dendrimers, nanogels, emulsions, polymeric and solid lipid nanoparticles, and nano-suspensions. In Parkinson's disease, nanoparticles are investigated as drug carriers to specific brain areas (Nikalje 2015) aimed at improving the patient's condition and quality of life, without being able to improve the prognosis.

The same strategies are under investigation for central nervous system tumors, HIV-induced involvement of the central nervous system, and Alzheimer's disease. In the latter, investigations are also aimed at removal of amyloid precursors and depositions.

In ophthalmology, nanoparticles can be used for diagnosis (measurement of intraocular pressure), treatment (ocular infections and retinal degeneration), and prevention (prevention of scarring and development of secondary glaucoma, of eye dehydration, etc.).

In surgery, nano-techniques can be used for visualization of tumors, for vascular repair, for tissue repair and removal, wound dressings and textiles with antimicrobial properties, etc.

Nanoparticles can affect the immune cells and the immune and allergic response and can therefore be used as drug carriers in autoimmune and allergic diseases. The question concerning the possibility for triggering an autoimmune disease (for instance, by activating complementary cascade or by altering the properties of self-antigens, such as proteins and DNA) remains unanswered.

In transplantation, nanoparticles can be used for drug delivery, for induction and maintenance of donor tolerance, and for diagnostic purposes (including detection of rejection).

A new field of nanomedicine is the development of nano-based prosthetic materials and artificial organs (biocomputer-based systems, incorporating the properties of the living organism and of the computer sensoring and responding systems). The investigations in this field have started with the development of artificial red and white blood cells that are not applicable in the clinical practice due to their unpredictable accumulation and destruction in the spleen. The nanosensors and the opportunities for the development

of nano-based-guided synthesis of specific DNA molecules are a promising new area of investigation.

Gene therapy is among the newest areas of investigation, using nano-based techniques. Multiple investigations, aimed at treatment of genetic diseases (including diabetes, alpha-1-antitrypsin deficiency, cystic fibrosis, and other diseases), are held all over the world. Nanotechnologies are used for the synthesis, vectoring (transfer), and targeting of DNA. These techniques are expected to provide a cure for multiple genetic diseases. Still, they carry the risk of improper and/or wrongly positioned insertion of the DNA fragment and unpredictable further consequences, including oncogenesis.

Pharmacokinetical and Pharmacodynamical Characteristics of Nanoparticles

Pharmacokinetics (PK) is what the living body does to a foreign agent (drug, substance, toxin, etc.) when it enters the body, and pharmacodynamics (PD) is what the agent does to the body (Bhattacharya et al. 2012). The PK and PD profiles of nano-materials are difficult to predict, because of their tiny-scale size and the huge changes in their chemical, mechanical, electromagnetic, and optical properties that happen in the nanoscale size and during their interactions with the living structures. Moreover, the accumulation of nanoparticles within the living body is hard to predict which makes the wide use of nano-materials dangerous. Nano-materials have increased absorption and bioavailability compared to their micro- and macro-analogues (because of their smaller size and free penetration through biological barriers), but their elimination is difficult to predict and therefore is thoroughly investigated. Three known factors that determine the probability of accumulation of nanoparticles are their size, charge, and hydrophilic or hydrophobic nature (Bhattacharya et al. 2012). The size of nanoparticles determines the risk of accumulation: <3 mm accumulate in a nonspecific manner within the living tissues, 3–8 mm are eliminated by the kidneys (especially dendrimers and other nano-polymers), 30–80 mm accumulate in the lungs and well-vascularized tissues, and those >80 nm accumulate within the liver and the spleen (Bhattacharya et al. 2012). The charge of nanoparticles determines their binding to plasma and tissue proteins, distribution, and their renal clearance. It also determines their immunogenicity and propensity to change normal proteins and biomembranes. And last but not least, hydrophobic particles tend to accumulate in parenchymal organs. Another important factor that determines the PK profile of nanoparticles is their carrier and cargo (i.e., the molecules they are bound to). Another confounding factor is the fact that certain cargos in nanoparticles can be released after the action of a specific triggering factor (targeted drug release) or are bound to an agent that decreases release (the so-called slow-release formulas), undergo significant recirculation within the human body (e.g., carbon nanotubes), or are inert and can undergo significant recirculation within the food chain (excretion by the human body, followed by liberation in the environment, uptake by other organisms, and eventually reentering the human body many, many times) (Bhattacharya et al. 2012).

One should not forget that at the nanoscale, small changes of structure and conjugation of different substances lead to large and unpredictable alterations in the behavior of the nanoparticle and therefore in the PK and PD profiles of the nano-material.

To make the long story short, the PK and PD profiles of many nanoparticle-based formulations are not well understood, and because of the nanoscale interactions and changes in the physical and chemical properties of material at these tiny sizes, no extrapolation can be made from the data concerning their micro- and macro-analogues. The carriers and the cargos of nano-materials are crucial for the PK and PD of the nano-material, and even small changes in nanoparticles, carriers, and cargos can lead to significant alterations and formation of new materials with new properties (Wani and Kothari 2018). Moreover, many nano-materials are virtually inert and undergo significant (maybe sometimes even endless)

recirculation within the human body and within the food chain and often tend to accumulate and cause undesirable effects, such as inflammation, induction of immune response, and maybe enhancement or triggering of oncogenesis (Ventola 2012c; Zhang et al. 2017). Nanoparticles are also known to be able to release or trigger the reactive oxygen species that lead to oxidative stress and inflammation and to activate the complement cascade and trigger immune response (Gomez Lopez 2013).

Pharmacogenetics (PG) is a relatively new interdisciplinary field on the borders between pharmacology, genetics, and personalized medicine that evaluates the influence of genetic markers upon the therapeutic response in every individual. It evaluates the role of genetic factors (genetic variations, mutations and polymorphisms, or drug-regulated gene expression and/or epigenetic factors) for the different therapeutic response (i.e., PK profile). Nanotechnologies, particularly nano-pores, can play important role in determining and predicting pharmacogenetic variations in drug metabolism and response by detecting single nucleotide polymorphisms and other genetic variations, as they offer fast and noninvasive DNA sequencing (Bhattacharya et al. 2012).

The toxicological profile of nano-materials remains unclear. Because of the huge changes in the physical and chemical properties of materials at the nanoscale and their multicomponent structure, they have unpredictable toxicity for the living body and for the environment. The main mechanisms of toxicity of nano-materials are associated with the generation of radical oxygen species with subsequent damage of the living structures and inflammatory changes (Lanone and Boczkowski 2006). The interaction with nuclear DNA with the development of subsequent oxidative changes may lead to oncogenesis. On the other hand, the interaction of nano-materials with mitochondrial structures and DNA might further impair cellular metabolism and energy stores.

The interaction with red blood cells may cause hemolysis, and the interaction with platelets and vascular endothelium may precipitate thrombosis.

The accumulation and interaction with liver cells may lead to hepatotoxicity, especially in hepatic accumulation of the nanoparticles. The accumulation within renal structures may lead to oxidative stress and inflammation with subsequent chronic kidney damage and renal failure. As the liver and the kidneys are the two major clearance points of all foreign materials that enter the human body, increasing concern is arising of the hepatic and renal side effects of nano-materials, including oxidative stress, inflammation, and oncogenesis, especially having in mind the increasing use of these materials, their stability, and circulation in the global ecosystem. This issue is very important not only for the human health but for the global safety of all living organisms on our planet (Bhattacharya et al. 2012; McClements and Xiao 2017).

Another serious problem is the lack of reliable methods to assess the toxicity and oncogenicity of nanoparticles, because of the unpredictable properties they may have compared to their macro-analogues (Bhattacharya et al. 2012). Currently we are facing a compelling need for new tools to assess the safety of nanoparticles.

Carbon-based nanotubes are a well-known nano-material, used for diagnostic and therapeutic purposes. Nonetheless, these nanoparticles are quite stable and tend to accumulate in the living body causing oxidative stress and chronic inflammation and changes resembling asbestosis (Bhattacharya et al. 2012). Quantum dots are also composed of toxic elements that can accumulate and harm the human body (Bhattacharya et al. 2012).

Cytotoxic drugs for cancer treatment and in organ transplantation are successfully delivered using modern nanosystems, but after excretion in the environment, they remain stable and may be incorporated by other organisms and undergo prolonged circulation within the ecosystem causing mutagenic and oncogenic effects. New drug delivery systems with decreased toxicity are being developed in response to this issue.

Nanocrystals (quantum dots) and metal- and silica-based nanoparticles can also initiate oxidative, inflammatory, and cytotoxic alterations in the living body, and their safety is still to be evaluated.

Moreover, all metal-containing nanoparticles may cumulate in the human body and cause chronic intoxications (Bhattacharya et al. 2012). Nano-polymers (dendrimers) can cause changes in cell interaction and activation but are known to have low immunogenicity and pro-inflammatory effects (especially the anion-carrying ones).

A serious problem represents the effect of targeted (especially cytotoxic drug carrying) nanoparticles on adjacent and remote tissues and cells (the so-called "off-target" effects) (Bhattacharya et al. 2012). This issue remains to be solved.

Safety and Ethical Issues: To "Nano" or Not to "Nano"?

With the increasing use of nano-materials in all fields of science, including medicine, the question of their safety becomes more and more pressing. They have multiple benefits, but, like a double-edged sword, they carry multiple and sometimes unpredictable and serious risk not only for the patient and the producing workers but for the ecosystem in general. Despite their wide application and the indisputable benefits for the modern food and energy industry, environmental protection, electronics, and medical care, the safety of nanoparticles remains unclear (Table 4), because of several serious problems (Anderson et al. 2016; Bhattacharya et al. 2012; Lanone and Boczkowski 2006; McClements and Xiao 2017; Prodanov 2017; Ventola 2012c; Zhang et al. 2017):

- Insufficient experience with nanoparticles
- Unpredictable physical and chemical properties in vivo because of the significant changes in the characteristics of the materials in the nanoscale with domination of quantum effects and appearance of new and unexpected characteristics
- Significant stability
- Insufficient data on their cumulation in the human body and circulation in the environment after excretion
- Unknown dose-response relationship and dose limitations in vivo
- Limited data on their physical, chemical, and toxicological properties and changes within the ecosystem
- Insufficient in vitro and in vivo experimental data
- Insufficient data on occupational exposure and hazards
- Lack of adequate markers for the evaluation of toxicity, dose, overdose, etc. – it remains unclear which markers are best, i.e., particle size, number, concentration, mass, etc.

According to their safety profile, nano-materials are divided in five major categories: A (no significant health risk), B (slight hazard and toxicity), C (moderate hazard), D (serious hazard), and E (severe hazard) (Prodanov 2017). Carbon-based nanotubes are known to belong to group E, because of their insolubility, cumulation, and toxic and DNA-damaging effects.

Another safety concern, besides the health and environmental risks, is the *preservation of our own self*. It remains unknown whether all modern nanoparticle-containing biosensors, especially when reaching the central nervous system or being able to change and control human DNA or brain activity, can be used for illegal or immoral purposes, such as mind control, eugenics, malevolent changes in human genome, etc. In other words, the nanoparticles are bringing once again the question whether we have reached the point at which full control upon the human body and own self (i.e., DNA) has come? Another serious question, especially concerning nano-prosthetics, artificial organs, and DNA-reparation techniques, is who covers the expenses; does every person deserve these cutting-edge new high-tech services? And what happens if the human body cannot tolerate these new technologies and the devices cannot be extracted without additional health risks? In the presence of biocompatible nano-based devices, who is human and who is not; where is the boundary between the human and the technology? Are the newly build DNA molecules and biocomputer systems alive, or are they merely a technological product or a machine?

Table 4 Advantages and disadvantages of nano-materials (Anderson et al. 2016; Bhattacharya et al. 2012; Gomez Lopez 2013; Prodanov 2017; Ventola 2012c; Wani and Kothari 2018; Zhang et al. 2017)

Advantages	Disadvantages
Fast and cheap in vitro and in vivo diagnostics	Adverse effects profile and toxicology are not well understood and unpredictable
High selectivity and specificity in diagnostics (DNA sequencing, specific DNA, RNA, protein and other targets, possibilities for specific and more stable dying of certain structures vs. conventional biopsy evaluation, real-time diagnostics, nontoxic and specific contrast media, etc.)	Unpredictable interactions with the living cells and structures due to changes in the physical and chemical properties of the materials at the nanoscale and possible domination of quantum effects, unknown magnetic and paramagnetic properties, and electroconductivity
Noninvasive diagnostics	Unpredictable PK and PD profiles. Carriers and cargos significantly change PK and PD
Selective and nontoxic imaging techniques	Generation of radical oxygen species
Targeted (highly specific) drug and radiation (isotope) delivery with minimal toxic effects – both systemic and for the adjacent tissue	Complement activation with subsequent inflammatory and/or autoimmune response, unpredictable oncogenic risk
Improved bioavailability of the therapeutic agent – i.e., lower doses and reduced adverse effects	Endothelial toxicity (particularly on cerebrovascular endothelium)
Targeted thermal ablation	Unknown risks or accumulation in parenchymal organs
Possibilities for PG evaluation (DNA sequencing)	Unknown risks of recirculation within the living body
Opportunities for theranostics – targeted diagnostic approach combined with immediate therapeutic intervention	Unknown effects on the environment and risks of circulation of unknown duration within the ecosystem

These questions have found no answers, and we should be extremely cautious with each and every new invention and technology, carefully weighting the advantages and risks, because we are entering an era when the boundaries between the macromorphological and the quantum worlds and between the human and the technology are becoming more and more obscure and we are facing the risk of losing and even destroying our human self. This "patent age of new invention" (Lord GG Byron (1824) Don Juan, I, 132) may literally bring us to the stars, but unfortunately it may lead us to our end.

References and Further Reading

Anderson DS, Sydor MJ, Fletcher P, Holian A (2016) Nanotechnology: the risks and benefits for medical diagnosis and treatment. J Nanomed Nanotechnol 7:4–5

Bhattacharya S, Alkharfy KM, Janardhanan R, Mukhopadhyay D (2012) Nanomedicine: pharmacological perspectives. Nanotechnol Rev 1:235–253

Dick PK. Autofac. In: Galaxy Magazine, November 1955

Dick PK. The android and the human. In: SF Commentary #31, December 1972

Drexler KE (1986) Engines of creation: the coming era of nanotechnology. Anchor Books, Doubleday

Emerich D, Thanos CG (2003) Nanotechnology in medicine. Expert Opin Biol Ther 3(4):655–663

Feynman R (1959) Engineering and science. California Institute of Technology

Gomez Lopez A (2013) Chapter 44: Nanotechnology and autoimmunity. In: Anaya JM, Shoenfeld Y, Rojas-Villarraga A et al (eds) Autoimmunity: from bench to bedside. El Rosario University Press, Bogota, pp 727–738

Guo T, Lin M, Huang J et al (2018) The recent advances of magnetic nanoparticles in medicine. J Nanomater 2018:1–8, Article ID 7805147. https://www.hindawi.com/journals/jnm/2018/7805147/

Gupta YK, Dinda AK (2018) Nano-pharmaceuticals in medicine. Proc Indian Natl Sci Acad 84(1):255–266. http://www.insajournal.in/insaojs/index.php/proceedings/article/view/470

Idrees M (2015) Role of nanotechnology in medical sciences: a review. Internat J Innov Drug Discov 5 (1):14–24. http://www.ijidd.com/File_Folder/14-24(ijidd).pdf

Lanone S, Boczkowski J (2006) Biomedical applications and potential health risks of nanomaterials: molecular mechanisms. Curr Mol Med 6(6):651–663

Lord GG Byron (1824) Don Juan, I, 132

McClements DJ, Xiao H (2017) Is nano safe in foods? Establishing the factors impacting the gastrointestinal

fate and toxicity of organic and inorganic food-grade nanoparticles. Sci Food 1:6–18
Milbum C (2010) Nanotechnology in the age of posthuman engineering. Science as science fiction. In: Nanovision. Engineering the future. Duke University Press, North Carolina, pp 19–59
Nikalje AP (2015) Nanotechnology and its application in medicine. Med Chem 5(2):81–89
Prodanov D (2017) Management of health risks related to the use of engineered nanomaterials. An analogy to biosafety. Biomed Rev 28:100–104
Sarfaraz S, Bano T, Fatima W (2018) Nanotechnology and its therapeutic application – a review. MOJ Bioequiv Availab 5(1):24–27
Surendiran A, Sandhiya S, Pradhan SC, Adithan C (2009) Novel applications of nanotechnology in medicine. Indian J Med Res 130:689–701
Taniguchi N (1974) On the basic concept of 'nano-technology'. In: Proceedings of the international conference on production engineering, Tokyo, 1974, Part II. Japan Society of Precision Engineering
Tasciotti E, Cabrera FJ, Evangelopoulos M et al (2016) The emerging role of nanotechnology in cell and organ transplantation. Transplantation 100:1629–1638
Ventola CL (2012a) The nanomedicine revolution. Part 1: emerging concepts. P&T 37(9):512–525
Ventola CL (2012b) The nanomedicine revolution. Part 2: current and future clinical applications. P&T 37(10): 582–591
Ventola CL (2012c) The nanomedicine revolution. Part 3: regulatory and safety challenges. P&T 37(11):631–637
Vijayakumar R, Agarwal R, Jagannathan S et al (2013) Application of nanotechnology in medicine. In: Navani NK, Sinha S (eds) Nanotechnology series. Volume 7: Diagnostics and therapeutics. Studium Press LLC, Houston, pp 89–100
Wani KA, Kothari R (2018) Nano technology in medicine and future implications: a mini review. J Nanomed Nanotechnol 9:497–499
Yadav HKS, Alsalloum GA, Al Halabi NA (2018) Chapter 14. Nanobionics and nanoengineered prosthetics. In: Grumezescu A (ed) Nanostructures for the engineering of cells, tissues and organs. Elsevier, Oxford, pp 513–587
Yambe T (2009) Artificial Organs with Nanotechnology and Development of the New Diagnosis Tool. 167–176. https://www.semanticscholar.org/paper/Artificial-Organs-with-Nanotechnology-and-of-the-Yambe/22555b5c39995c4df7c4b5a8c07752a6287f759b
Zahid M, Kim B, Hussain R et al (2013) DNA nanotechnology: a future perspective. Nanoscale Res Lett 8(1): 119–131
Zhang WK, Gu HW, Li XJ et al (2017) The dark side of "the force" – lipid nanoparticles enhance the oncogenesis of diethylnitrosamine and result in liver cancer in mice. Nanomedicine 13(2):701–711. https://www.ncbi.nlm.nih.gov/pubmed/27729235

Gender Differences in Drug Therapy

28

Anthony G. Fenech and Vanessa Petroni Magri

Contents

A. G. Fenech (✉)
Department of Clinical Pharmacology and Therapeutics,
University of Malta, Msida, Malta
e-mail: anthony.fenech@um.edu.mt

V. Petroni Magri
Department of Clinical Pharmacology and Therapeutics,
Faculty of Medicine and Surgery, University of Malta,
Msida, MSD, Malta
e-mail: vanessa.petroni@um.edu.mt

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_69

Abstract

Therapeutics has steadily evolved over the last two decades, from a broad system designed to target multiple patients with one management approach to a personalized model that focuses on the characteristics of each individual patient and tailors treatment accordingly. In order to achieve the maximum benefit and minimum risk, personalized medicine needs to address multiple individual parameters including pathology, physiology, genetics, comorbidity, sex, and gender. While considerable personalized pharmacotherapeutic advances have been made in all fields, sex and gender have somewhat lagged behind. This has possibly been the result of a long-standing premise that females of childbearing age should be generally omitted from clinical trials. This is now changing. For several years, drug regulatory agencies have been advocating the inclusion of both sexes in preclinical and clinical trial design, ensuring adequate statistical power to discriminate sex-gender differences. This has led to a gradual but steady increase in published data, which highlights sexually dimorphic differences in the pharmacokinetics, pharmacodynamics, and pharmacogenetics of several drugs. Male and female drug metabolism patterns, gene expression profiles, drug interaction pathways, and adverse drug reaction reporting have all contributed to a *corpus* of knowledge that needs to be constantly expanded and solidified. Furthermore, this information needs to be understood within the context of gender-bias in adherence to therapy and adverse effect reporting. Adherence to medication is a complex matter based on multiple factors including socioeconomic, health system, and patient-specific factors.

Further sex-gender research is required to consolidate and expand current data, within bench-to-bedside translational application frameworks. This needs to be incorporated into all commercial drug development stages. The inclusion of sex-gender differences in independent research needs to be incentivized through peer-review guidelines, grant application priorities and institutional review boards. This drives prospective study design towards research models capable of generating information that contributes towards the inclusion of sex-gender differences in personalized medicine treatment algorithms.

Abbreviations

ACE	Angiotensin converting enzyme
ADME	Absorption, distribution, metabolism and excretion
ADR	Adverse drug reaction
ALT	Alanine transaminase
AST	Aspartate transaminase
AUC	Area under the curve
C_{max}	Peak plasma concentration
CPY450	Cytochrome P450 enzyme system
DMET	Drug metabolizing enzymes and transporters
EMA	European Medicines Agency
FDA	Food and Drug Administration
GWAS	Genome-wide association study
HAGMA	High anion gap metabolic acidosis
HDL	High-density lipoprotein
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
LDL	Low-density lipoprotein
NSAID	Non-steroidal anti-inflammatory drug
SGD	Sex-gender differences
UGT	Uridine diphosphate-glucuronosyltransferase

Introduction

The meaning of the term "gender" has diversified over the years, to carry an independent meaning to "sex," and this has been reflected both in the scientific, as well as in the popular press. "Gender" is today recognized to refer more to aspects regarding social roles, cultural differences, behavioral aspects and nonbinary identities, while "sex"

remains the term which biologically identifies males from females in terms of an individual's genetics, reproductive system, and secondary sexual characteristics.

Both these aspects are relevant to drug therapy. While sex-related physiology imposes different ways through which drugs are handled by the body, gender imparts cultural and behavioral differences which affect the way drugs are used by the patient, as well as the patient's perception of therapeutic and adverse drug outcomes. For example, males and females may have different gender-related perceptions and reactions to the same pain stimulus, and this may influence the perceived efficacy of analgesia in both sexes. The same holds for differences in male and female self-prescribing attitudes.

It is often difficult to discriminate whether observed male/female differences in pharmacological effects are due to sex, gender, or a combination of both. While sex physiology may be considered to be an objective contributor to pharmacological activity, gender-associated differences (e.g., adherence to therapy, dietary habits and perceived outcomes) may influence the outcomes of studies that are not powered enough to factor in these covariates. The term sex-gender differences (SGDs) is often used to refer to such difficult-to-discriminate variability.

The recognition and embodiment of sex-gender differences into pharmacological management is an important component of personalized medicine, a term which stems from concepts first introduced by Hippocrates (460–c. 370 BC) (Sykiotis et al. 2005). This has today evolved into a patient-centered science that envelops consideration of lifestyle factors, comorbidities, cultural influences, biomarkers, and genomics, among others, within the context of an evidence-based management system.

In order to address this, sex- and gender-related therapeutic aspects need to be introduced early into the drug development process. The cyclic nature of hormonal regulation in females has often resulted in their exclusion from preclinical and clinical studies, with the premise of maintaining homogeneity in the test population. This approach has been challenged. Preclinical animal studies, as well as all clinical phases, need to include both males and female covariates. Additionally, studies need to be statistically strongly powered enough in order to be able to detect potential sex-gender differences.

Indeed, as early as July 1993, the FDA published the "Guideline for the study and evaluation of gender differences in the clinical evaluation of drugs" (Food and Drug 1993). The development of this guideline was driven by concerns that the drug development process in place did not allow room for adequate identification of sex-gender differences in therapeutics, due to the general exclusion of female candidates in premarketing drug trial phases. The guideline called for a consensus that "women should be allowed to determine for themselves the appropriateness of participating in early clinical trials" and "sponsors should collect gender-related data during research and development and should analyse the data for gender effects in addition to other variables such as age and race." In order to successfully accomplish this, the guideline required sponsors to "include a fair representation of both genders as participants in clinical trials so that clinically significant gender-related differences in response can be detected." The guideline further emphasized the importance of addressing the effects of the menstrual cycle and exogenous hormonal therapy including oral contraceptives on drug outcomes. Furthermore, the converse of this, i.e., the effect of the drug on the efficacy of oral contraceptives, was recommended.

Pharmacological Response

The outcome profile of drug therapy is the culmination of a multifactorial process that involves interrelated events related to pathophysiology of disease, pharmacokinetics, pharmacodynamics, biochemistry, genetics, behavioral and environmental issues, as well as commonly recognized characteristics such as age and body weight. Moreover, gender-biased prescribing patterns, as well as gender-dependent adherence to therapy contribute to the overall picture. For example, a recent scoping review found that women with dementia living in the community were more

likely to receive potentially inappropriate medication than men, whereas among residents of nursing homes, men tended to receive potentially inappropriate medication at a greater frequency than women (Trenaman et al. 2019). The dissection of mechanisms that drive inter-individual differences in drug treatment outcomes is a complex endeavor that necessitates an integrative approach to multiple contributory factors. The contributions of sex and gender to drug response operate within this multivariate framework and may therefore be difficult to dissect. Furthermore, temporal factors such as age-related differential developmental processes in males and females and pre-and post-menopausal changes in females present a time-axis along which sex-gender differences may fluctuate.

Pharmacokinetics

Pharmacokinetics is the study of the influence of the body on drugs, and it follows the fate of a drug from administration, through distribution, metabolism, and finally elimination (ADME). It rationalizes drug dosing regimens based on various quantifiable parameters of physiologically dependent drug properties, such as bioavailability, volume of distribution, half-life, and clearance. These parameters may be affected by intrinsic factors such as body weight, disease, renal and hepatic function, as well as extrinsic factors such as smoking, alcohol, and diet. Drug pharmacokinetics may also be subject to sex-dependent differences, either due to direct hormonal influence such as hormone-dependent gene expression or due to indirect influences such as changes in body water composition at different phases of the menstrual cycle (Nicolas et al. 2009). Table 1 provides some examples of SGDs that may influence drug pharmacokinetic profiles.

Absorption: The most common route of drug absorption, that is via the gastrointestinal tract, is subject to multiple physiological parameters, including gastric acid secretion as well as gastric and bowel transit times. Women have been reported to have lower rates of gastric acid secretion and consequently a somewhat higher gastric pH than men. This is coupled with slower gastric emptying, and a longer intestinal transit time especially during progesterone-high phases. These parameters may exert an influence on pH-dependent drug dissolution, as well as on the overall time available for intestinal absorption of the drug from its respective formulation. Male/female differences in P-glycoprotein efflux transporter activity, especially within intestinal tissue, also contribute to dimorphic profiles in drug absorption (Bebawy and Chetty 2009).

Distribution: The distribution of a drug throughout the various body compartments as well as within subcellular organelles (Li et al. 2018) is a complex event that is determined by multiple physiological variables including disease (Gonzalez et al. 2011). Besides well-recognized issues such as tissue perfusion rate, plasma protein binding, barrier membrane systems, and tissue/plasma equilibria kinetics, other parameters may be influenced by physiological differences between the sexes, and by cyclical variations in these differences. Perhaps the major cyclical variation contributing to changes in drug distribution is the estrogen-mediated salt and water retention, which contributes to a larger water compartment, larger volume of distribution, and potentially reduced tissue absorption of hydrophilic drugs. A larger water compartment may imply a lower plasma protein concentration, thus influencing the free/bound kinetics of highly protein bound drugs. These changes are further compounded by differences in the distribution of adipose tissue deposition between males and females, and by the increased testosterone-induced muscle mass which occurs in males, thus causing shifts in the compartmentalization ratios of administered drugs.

Metabolism: Drug metabolism is a multistep process which normally converts the drug into less active or inactive components, prior to excretion. This stepwise conversion is conventionally divided into three phases with Phase I being a chemical modification phase, often driven by members of the CYP450 enzyme system, Phase II being a conjugation step, and Phase III being additional final modifications in preparation for the metabolite excretory process. While a number of Phase I drug metabolism pathways has been

Table 1 Examples of differences between male and female pharmacokinetic parameters. These differences may drive pharmacokinetic-based sex-gender dimorphisms in therapeutic and adverse drug responses

Examples of affected drugs	Male/female dimorphism	Reference
Absorption		
Rifampicin	Longer gastric emptying time, longer transit time and higher gastrointestinal pH in females Higher gastrin secretion in women than men	(Soldin and Mattison 2009; Coskun et al. 1995; Feldman et al. 1983)
Distribution		
Diazepam Imipramine Verapamil	Increased volume of distribution for lipophilic drugs in females due to higher proportion of body fat. Increased volume of distribution for hydrophilic drugs in males due to higher plasma volume Increased serum binding globulins in females Lower levels of plasma alpha-1 acid glycoprotein in females Lower Pgp (MDR1, ABCB1) efflux transporter intestinal expression in females than in males	(Jochmann et al. 2005; Soldin et al. 2011; Kishino et al. 2002; Bebawy and Chetty 2009)
Metabolism – Phase I		
Rilpivirine Ritonavir Codeine Clarithromycin Erythromycin	Lower activity of some CYP450 enzymes in females (e.g., CYP1A2, CYP2E1, CYP2D6) and higher activity of others (e.g., CYP3A4)	(Oertelt-Prigione and Regitz-Zagrosek 2009; Jochmann et al. 2005; Austin et al. 1980)
Metabolism – Phase II		
Voriconazole Azathioprine	Higher UDP glucuronosyl transferase activity in males Higher methyltransferase activity in males	(Morissette et al. 2001; Szumlanski et al. 1992)
Elimination		
Propranolol Verapamil	Lower renal clearance in women	(Franconi et al. 2011; Krecic-Shepard et al. 2000)

shown to exhibit sexually dimorphic features, there is little such evidence in the literature for Phase II and III metabolism.

Phase I CYP450 enzymes are the products of an extensively polymorphic family of genes[1] some of which have over 100 variants reported in the literature, and which may exhibit different allelic frequencies across different ethnic groups. Differences in the activity of specific CYP450 enzymes may be the result of genetic variability but may also be due to variations in genetic expression caused by factors such as age and sex. The variables influencing CYP450 activity, which may include sex-hormone-regulated gene expression, are not straightforward. An extensive study of CY450 activity determinants (Yang et al. 2010) reported females to show higher hepatic CYP3A4T activity than males; a difference that was maintained across age. However, increasing age was found to also be a contributor to higher CYP1A2 activity in females but not in males, and a similar trend was found for CYP2B6 activity in males but not in females. This suggests that sex-differences in activities of these Phase I enzymes are not necessarily consistent across all age groups. Indeed, the authors reported females to have lower overall CYP450 activities than males in the younger age groups, with the situation being reversed in later years. This observation has also been reported from other studies involving CYP1A2, CYP2D6, and CYP2E1. For example, the angiotensin receptor blocker telmisartan, shows a higher C_{max} in females than in males, possibly due to lower female CYP2C8 and CYP2C9 expression. Interestingly, some studies have reported female CYP3A4 gene expression to be higher

[1]Pharmacogene Variation Consortium. PharmVar database. https://www.pharmvar.org/genes

than in males; however, it is coupled with a female-dominated higher clearance of CYP3A4 substrates, suggesting that the expected functional outcome would be the net result of balancing both factors (Zhu et al. 2003; Schwartz 2003, 2007; Krecic-Shepard et al. 2000; Cabaleiro et al. 2013). Other studies have, however, reported approximately twofold higher protein expression of CYP3A4 in female liver tissue and suggested that this may be potentially responsible for up to 50% of sexually dimorphic pharmacokinetic variability in the metabolism of CYP3A4 substrates such as antipyrine, alfentanil, erythromycin, midazolam, and verapamil (Zanger and Schwab 2013).

Further evidence of the general tendency for females to exhibit higher CYP450 gene expression than males comes from two large hepatic expression profiling studies. Zhang and co-workers identified 40 ADME-related genes with higher female expression than male, and 3 CYP3A5, CYP27B1, and UGT2B15 showing male bias (Zhang et al. 2011). A year later, Yang and co-workers reported 77 drug metabolizing enzyme and transporter (DMET) genes with sex-dependent differential hepatic RNA expression from a panel of 374 studied DMET genes (Yang et al. 2012). The top 10 genes consisted of four Phase I enzyme coding genes (*CYP7A1*, *CYP3A7*, *CYP3A4*, *ADH1*) all of which showed greater expression in females, four Phase II enzyme coding genes (*ACSL4*, *GSTA1*, *GSTA2*, *UGT2B17*), of which all showed positively female-biased expression except UGT2B17 which showed a male-biased expression and two transporter genes (*SLC3A1and SLC13A1*) of which the first was expressed at higher levels in females, and the second showed higher expression in males. The authors proceeded to carry out a co-expression gene network analysis of these 77 sexually dimorphically expressed genes, in order to identify groups of genes, which share common drivers of expression. *CYP2A6*, a known estrogen-inducible Phase I gene, shared expression similarities with *ALDH5A1*, *CYP2B6*, *CYP2B7P1*, *SLC10A1*, *GSTA1*, *GSTA2*, and *GSTA5*, while *CYP3A4* shared expression similarities with *FMO3*, *GSTA1*, *GSTA2*, *GSTA5*, *ALDH5A1,* and *SLC10A1*, possibly suggesting common sex-dependent direct expression drivers or an interdependency and genetic cross-talk network being driven by sexually dimorphic factors.

Phase II metabolic conjugation processes are little understood with respect to sex-gender differences, and the published literature is very meagre in this respect. There is little evidence to support hormonal influences on expression of Phase II enzymes or a sexually dimorphic distribution of functional Phase II gene variants. The reported data arise from measurements of drug concentrations and/or their Phase II metabolites in patients, and therefore may also include disease-related confounding factors. For example, it generally appears that Phase II drug glucuronidation processes (e.g., with β_1-adrenoceptor blockers) occur faster in men than women, but the mechanism is unclear.

Elimination: Recognized sex-differences in excretion rates are mainly tied to the renal elimination pathway mechanisms. Males have long been recognized to have higher serum creatinine clearance than females, and this difference is maintained across different age groups (Berg 2006). This may result in a risk for increased toxicity of renally excreted drugs with a low therapeutic window, such as digoxin, in females compared to males administered the same dosage regimen (Rathore et al. 2002)**.** Biliary excretion sexual dimorphism also represents a poorly understood mechanism. Implications are generally directed towards the P-glycoprotein efflux transporter, since its expression from the *ABCB1* gene is twofold lower in human females than in males. This may potentially explain, for example, the more sustained intracellular concentration of fexofenadine in females than in males (Kim et al. 2001).

Pharmacodynamics

In order to exert their function, drugs interact with complex molecular physiology in order to induce cellular changes which propagate to an organ level and to eventual clinical manifestations. Such manifestations include therapeutic outcomes as well as unwanted adverse effects. These

outcomes are also subject to psychological factors, which are most clearly embodied in the well-established placebo and nocebo effects, some of which have a degree of sex-gender dependence (Enck and Klosterhalfen 2019). The interaction of multiple physiological and psychological factors creates a multivariate environment with multiple potential sex-gender-dependent contributions which may be difficult to individually dissect. This is also evident in drugs of abuse as well as recreational drugs such as alcohol and tobacco. For example, cocaine and amphetamine induce more pleasurable responses in women than in men. Additionally, women escalate doses and dosing frequency faster than men, eventually becoming addicted earlier. Similarly, women experience greater smoking withdrawal symptoms than men, while they experience less alcohol withdrawals (Becker et al. 2017).

Therapeutic Outcomes: Normalization of drug doses for equivalent efficacy in different individuals is commonly carried out using age, body weight, or body surface area parameters, in a sex-independent manner, and these are considered to be adequate for clinical applications. Other widely recognized considered variables include disease and comorbidities, and concomitantly administered drugs. Sex-gender differences in pharmacodynamic variables have been documented both at a molecular as well as at a clinical level. These may offset the otherwise accepted clinical dosage normalization procedures. For example, sex-differences in tissue receptor densities, intracellular calcium, and sympathetic neurotransmission are recognized events, though there seems to be little evidence to support an overall hormonal influence on these molecular outcomes (Ziegler et al. 2017).

Moreover, clinical endpoints in drug efficacy may include an element of patient subjectivity, especially in areas which include neuropharmacology where patient-independent therapeutic outcome assessments are not always possible. For example, the reported efficacy of anti-depressant or analgesic therapy partially depends on patients' perceptions, and gender-dependent differences in these perceptions may influence such management outcomes. A systematic analysis of 18 studies of placebo (expectation of therapeutic outcomes from unknowingly pharmacologically inert products) and nocebo (expectation of adverse outcomes from unknowingly pharmacologically inert products) effects in males and females, identified placebo effects to be more predominant in males while nocebo effects to be more predominant in females (Vambheim and Flaten 2017). This is in line with the established increased adverse effect profile generally reported in females. The authors postulate that this could be due to combined affective responses and psychophysiological mechanisms, with stress being a frontline factor. Since higher stress levels are normally reported in females than males, this may predispose to an expectation of adverse effects to the administration of a product which is effectively pharmacologically inert. The nocebo response may be further augmented by previous negative drug experiences, which contribute to further psychological assertion of negative therapeutic expectations. Table 2 summarizes some examples.

Cancer therapeutics represent a particularly relevant pharmacological area where SGDs may be an important contributor. For example, there is evidence to suggest that immunotherapy outcomes may be influenced by sex-driven differences in immune response. These differences may be the result of X-linked genes providing an immunoreactivity advantage in females, coupled to complex estrogen-driven influences on various cellular components of immune pathways (Gabriele et al. 2016; Markle and Fish 2014). The sex-dependent effects of tyrosine kinase inhibitors such as imatinib, sorafenib, axitinib, and others have been succinctly reviewed (Schmetzer and Florcken 2012). These drugs, though belonging to the same class, appear to exert different sex-dependent profiles of efficacy and toxicity, though there is no general agreement in the literature. The heterogeneity of cancers and disparities between the clinical background of different patients, such as age, smoking habits, and tumor histological and molecular subtype, often makes it difficult to establish SGD contributions within a treatment setting. Sex may therefore be an important stratification variable to be

Table 2 Sexually dimorphic pharmacodynamic responses in commonly prescribed drugs. (*TC* total cholesterol, *LDL* low density lipoprotein, *AD* Alzheimer's disease, β_2AR β_2 adrenoreceptor)

Examples of affected drugs	Sexually dimorphic therapeutic effect	Reference
Simvastatin atorvastatin	Greater decrease in TC and LDL levels in women	(Smiderle et al. 2014)
Galanthamine	73% greater chance of positive response by males with AD than by females	(MacGowan et al. 1998)
Amlodipine	Increased hypotensive effect in females	(Kloner et al. 1996; Spratt 1997)
Aspirin	Increased resistance to anti-platelet activity in females	
Fenofibrate	Improved lipid profile in females compared to males	(d'Emden et al. 2014)
Rosuvastatin	Improved atheroma volume regression in females at start of treatment, which however became similar to that in males after 24 months	(Puri et al. 2014)
Citalopram	Better antidepressant response in females than males	(Young et al. 2009)
Imipramine	Better antidepressant response in males than females	(Kornstein et al. 2000)
β_2-adrenoceptor antagonists	Reduced antagonist response in females due to increased β_2AR expression and signaling	(Kneale et al. 2000; Mills et al. 1996)
Endothelin receptor antagonists	Improved therapeutic response to pulmonary arterial hypertension in women	(Gabler et al. 2012)

included in all cancer clinical trials, as knowledge on its contribution to prognosis may serve to improve therapeutics in the clinic.

Adverse Effects: Females have often been reported to suffer from a greater overall incidence of adverse drug reactions than males, and pharmacovigilance data has generally supported this observation. Reports of overall increased female ADR incidence have been as high as twofold over males (Nakagawa and Kajiwara 2015), and this difference seems to predominate in the adult age-groups (Damien et al. 2016). A position document of the European Society of Cardiology states that women experience more frequent adverse drug reactions from diuretics (hyponatremia, hypokalemia, and severe arrhythmias), ACE inhibitors (cough), anticoagulants (hemorrhagic complications), statins (myopathy), and antiarrhythmic drugs (longer QT interval). These effects, especially QT prolongation, may be estrogen-facilitated but may be also contributed by the polypharmacy management, which is generally more observed in females than in males (Rosano et al. 2015).

Sexual dimorphism in ADRs may be due to multiple reasons; (a) sex-dependent pharmacokinetic differences, often leading to higher plasma concentrations of some drugs in females and the occurrence of Type A adverse reactions, (b) sex-dependent pharmacodynamic differences occurring at a cellular or tissue level, which are often secondary to male/female physiological and endocrinological differences, and (c) gender-dependent differences in self-reporting or perception of adverse effects (e.g., nocebo effects discussed earlier). For example, females exhibit a longer baseline QT interval then men, and therefore drugs that cause QT interval prolongation adverse effects are likely to negatively influence women more than men (Stramba-Badiale et al. 1997; Lehmann et al. 1996; Ruan et al. 2004). Additionally, there are more reported instances of resistance to the anti-platelet activity of aspirin in females, than in males, and this is supported by early evidence showing that androgens support the aspirin anti-platelet effect while estrogens oppose it by increasing platelet-activity (Spranger et al. 1989). Moreover, the male/female risk of adverse reactions is also potentially influenced by ethnicity, although solid evidence for this is scarce, possibly due to most studies not being adequately powered to allow for ethnic population stratification.

A particularly interesting SGD-related target is the renin-angiotensin system. Angiotensin II plasma concentrations have been reported to increase under the influence of estrogens, while both ACE activity and *AGTR1* gene (angiotensin II, type 1; AT1 receptor) expression decrease (Fischer et al. 2002). In males, androgen activity increases transcription of the angiotensin I precursor and renin substrate, angiotensinogen. This may contribute to sexually dimorphic responses to the antihypertensive effects of drugs acting on the renin-angiotensin system, when administered in similar dosage regimens (Table 3).

Drug Interactions: Drug combinations present a unique situation where drug-interaction-based adverse effects may follow different sex-biased patterns due to divergent sexual dimorphism of the interacting drugs. This complicates the analysis of ADR reporting data, because of potentially unrecognized confounding SGD drug interaction events. Unfortunately, there have been difficulties with identifying the role/s of sex-gender in drug interaction studies that included sex as a covariate, possibly because such studies need to be very highly statistically powered in order to extract meaningful data from the multivariates of polypharmacy treatment. Most drug interaction studies in humans are based on retrospective data, due to the ethical implications of designing prospective studies in this area. The differences in polypharmacy between males and females makes retrospective studies even more difficult, and other variables such as age further complicate the matter. For example, in a retrospective study carried out on elderly male and female patients, women younger than 80 years old were identified to use more drugs than men in the same age group. In this study, 32.6% of men and 49.2% of women had at least one drug-drug interaction, while 8.1% of men and 10.6% of women described four or more potential drug-drug interactions (Venturini et al. 2011). Moreover, drug-interaction studies carried out in small animals are difficult to extrapolate to humans, especially in view of the multiple pharmacokinetic differences between the species.

Some studies have reported co-administration of statins and fibrates to result in a significantly higher incidence of cardiovascular related adverse events in females than males; an observation that was not evident with either drug administered alone (Ginsberg et al. 2010; d'Emden et al. 2014). High anion gap metabolic acidosis

Table 3 Examples of reported sexually dimorphic adverse drug reactions. (*ADR* adverse drug reaction, *AST* aspartate transaminase, *ALT* alanine transaminase)

Examples of affected drugs	Sexually dimorphic ADR	Reference
Amlodipine	Increased incidence of peripheral edema in women	(Kloner et al. 1996; Spratt 1997)
Insulin	Females have a higher risk experiencing severe hypoglycemia	(McGill et al. 2013)
Fingolimod	Females have a higher risk of developing an infectious episode Males have a higher risk of AST and ALT alterations	(Manni et al. 2017)
Cyclophosphamide Ifosfamide	Increased rates of hematologic toxicity and stomatitis in females	(Wrobel et al. 2011)
5-fluorouracil	Females have a higher incidence of oral mucositis, but lower rates of gut toxicity than males Nausea and vomiting is a greater problem in females during therapy due to the lower activity of anti-emetic drugs	(Schmetzer and Florcken 2012)
Simvastatin Atorvastatin	Higher frequency of myalgia in women Increase in CPK and/or abnormal liver function more frequent in in men	(Smiderle et al. 2014)

(HAGMA), a rare drug interaction between paracetamol and flucloxacillin also showed a higher incidence in females than males, possibly due to sex-dependent differences in cellular biochemical physiology (Jessurun et al. 2016). Being male was listed as a contributory risk factor for drug-drug interaction events in intensive coronary care patients (Jankovic et al. 2018).

Rodent studies showed simvastatin metabolism to be mainly CYP2C11-driven in males and CYP3A-driven in females. This sexually dimorphic biochemistry, contributes to different male/female rates of simvastatin metabolism due to the different ability of itraconazole to inhibit CYP3A and CYP2C11 (Ishigami et al. 2001). Itraconazole is known to inhibit hepatic simvastatin metabolism in humans, leading to simvastatin toxicities such as myopathy.

Adherence to Therapy

The World Health Organization defines adherence as "the extent to which a person's behaviour taking medication, following a diet and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider." Adherence is frequently incorrectly used interchangeably with compliance in the medical scenario. However, as opposed to compliance, adherence involves the participation of patients when deciding on the recommended therapeutic management (WHO 2003).

Adherence is a multidimensional concept. Rather than being the sole responsibility of the patient, adherence to medication is the result of the relationship existing among social-economic and health-system factors, as well as patient-, condition-, and health-related factors (WHO 2003). A sociodemographic variable which has been evidenced through numerous studies to affect adherence is gender.

Therapeutic adherence is of empirical importance in disease prevention and control. Erratic adherence to medication can lead to exacerbations, which would increase probability for hospital admission. This may adversely impact the economy since it results in greater hospital fees, coupled with a reduced output from the population's working force. A study which depicted this scenario clearly showed that adherence to therapy in patients suffering from chronic vascular disease leads to a decrease in the annual health care expenditure. This was mainly a result of less emergency department visits as well as less inpatient admissions (Roebuck et al. 2011).

Significant differences in therapeutic adherence have been reported between males and females. In-depth study of such considerable disparities has led to the conclusion that ulterior clinical outcomes can be achieved by improved, personalized therapeutic management and medicine selection (Manteuffel et al. 2014).

Disease-Dependent Therapy Adherence

Disease-dependent medication adherence pivots upon a multitude of aspects which affect males and females differently in specific conditions. Some important factors include disease and symptom severity, treatment availability, disability level, as well as the rate of disease progression. Disease progression, as well as associated complications, relies on the psychosocial and biological differences observed in men and women. Adherence is also affected by past or current alcohol and drug abuse problems, as well as comorbidities, notably depression (WHO 2003; Kautzky-Willer and Harreiter 2017). Such issues are often confounded by concomitant sex-gender differences in lifestyle, work commitments, and the financial income that may be necessary to provide access to medications.

Comorbidity Effect on Therapy Adherence

All mental health conditions can have a detrimental effect on a patient's adherence to their prescribed therapeutic regimen, a trait which is studied intently in depression. Depression has been described as a long-lasting or recurrent

condition which affects the patient's normal daily functioning. Individuals suffering from depression experience sadness, guilt feelings, low levels of concentration and interest, loss of appetite, sleep disturbances, tiredness, and little self-worth. Females tend to suffer more from this condition than males. A consequence of very severe depression may be suicide (WHO 2017).

Depression, especially when left untreated, may be debilitating, which will in turn alter the patient's ability to appreciate the importance of adhering to their treatment and management plan. Hence, there may be negative effects on both depression, as a stand-alone condition, and other comorbidities the patient may be suffering from such as serious potentially terminal diseases including cancer.

Depression as a Comorbidity in Oncology Patients

Cancer patients suffering from depression suffer from an impaired quality of life as well as an elevated mortality risk (Dauchy et al. 2013; Pinquart and Duberstein 2010). Depression in cancer can be managed via pharmacotherapy (e.g., tricyclic antidepressants (TCAs) and selective serotonin reuptake inhibitors (SSRIs)), as well as with nonpharmacological approaches such as psychosocial interventions including problem-solving therapy, relaxation strategies, cognitive behavioral therapy, acceptance and commitment therapy, psychoeducation, and exercise therapy (Smith 2015).

Linden and coworkers report that women, as opposed to men, experience a higher tendency to suffer from depression for all cancer types. Explanations for such a discrepancy included female's aptitude to react emotionally to certain life circumstances, as opposed to men finding it difficult to express certain emotions. Additionally, certain cancer types which men are diagnosed with are associated with a better prognosis and life-expectancy outcome (Linden et al. 2012). Hence, adherence to the depression-associated treatment is of empirical importance. Cancer outcomes can vary with the presence of depression as a component in the clinical picture, since this can impact the consistency to maintaining the whole therapeutic regimen (Dauchy et al. 2013).

Screening for depression as a prevailing comorbidity is not enough when it comes to patient management. Planned routine patient counselling for both males and females would be beneficial in assessing whether the patients are adhering to their therapy management for their diagnosed depression. This will be advantageous since optimal control of depression can affect the adherence and thus the outcomes of other coexisting conditions.

Medication Side-Effects and Adherence

Medication side-effects can be sex-specific, hence presenting differently in males and females. Certain drug classes, e.g., respiratory, genitourinary, antiparasitic, antineoplastic, and sex-hormone medications have been associated with side-effects which are more commonly reported in women rather than males (Montastruc et al. 2002). A study conducted by Bilimek and coworkers found that side-effects associated with intense lipid-lowering therapy regimens were particularly experienced by females. This resulted in medication nonadherence and thus poor lipid control in the female group. This assessment was reported even though both men and women were offered comparable quality of medical care (Billimek et al. 2015).

The effective management of several oncologies tends to incorporate multiple modalities, where patients can be exposed to various approved therapies. These include surgery, radiation, as well as the administration of chemotherapeutic drugs. Therapeutic regimens have improved over the years, both in their efficacy to improve quality of life as well as in their associated reduced toxicity. However, toxicity experienced by cancer patients due to such therapies is still considered to be a major deterrent in their quality of life. The majority of these said therapy-associated side-effects can nowadays be effectively managed by certain drugs. Nevertheless, patients suffering an impeded life as a result

of the disease itself, as well as the associated management, leads to the nonadherence of both cancer therapies as well as the prophylactic medication. This practice tends to be noted in both female and male sexes.

Other Factors Affecting Therapy Adherence

Response to therapy can vary in men and women and is thus another important concept to consider when selecting medication. An unsuccessful response to therapy may eventually result in nonadherence resulting in disease progression or exacerbation. A tangible example lies with cardiovascular drugs which are characterized by different pharmacological profiles in women and men. Physiological differences are evident in women and men, and this is especially true in cardiovascular disease, which is associated with gender-specific clinical manifestations (Stolarz and Rusch 2015).

Although women tend to be more keen than men in preventing and treating disease, Menteuffel and coworkers reported that females are less adherent than males with respect to diabetes and cardiovascular medications. Plausible reasons center around the fact that women tend to have more complex therapeutic regimens, making adherence a challenging task. As previously mentioned, unpleasant side-effects may also pose as a contributor to this discrepancy. Furthermore, women may feel that it is challenging to address their own health issues, since they feel responsible for their family's care. Hence, most of their effort and energy is invested elsewhere, rather than for their self-care (Manteuffel et al. 2014).

A contribution to therapy nonadherence which is commonly overlooked is that of medication cost. Cancer medications are known to be highly expensive, normally exceeding the average price-point of nononcologic medications. This creates a difficult situation for the patients and their families, since they would not be financially competent to make such hefty payments for cancer, especially considering the immediate time-frames normally required to increase therapeutic success. This contributes to therapeutic nonadherence and possibly cancer progression. Medication cost issues can affect males and females differently according to income, the country they reside and work in, whether they are part of a medical insurance scheme, or whether their country of residence runs a national health service that provides free or subsidized drugs.

Improving Patient Adherence

In the modern world of medicine, nonadherence to therapy is considered to be a significant public health problem, particularly in patients who are suffering from a chronic condition. Counteracting this growing epidemic shall involve a collective effort from all governmental and private health clinics and hospitals who would be willing to offer interventions focusing on behavior, education, integrated care, and self-management (Costa et al. 2015).

Thus, when deciding on the therapeutic management plan, health care professionals must carry out detailed discussions with the patient in question. Health care providers should acknowledge and encompass every variable which might affect patient adherence to therapy. This will pave the pathway to increasing the promise of adherence among the patient population suffering from a range of various diseases.

Gender and Pharmacogenetics

Pharmacogenetics has rapidly established itself as a major specialty within both the academic and applied fields of pharmacology. As of the current date, the leading PharmGKB knowledgebase[2] provides pharmacogenetic data on over 650 drugs and provides links to more than 100 pharmacogenetic-based dosing guidelines published by the Clinical Pharmacogenetics Implementation Consortium.[3]

Pharmacogenetic variability is the result of functional variability in genes which contribute to the pharmacological outcomes of the drug.

[2]PharmGKB (http://www.pharmgkb.org)

[3]Clinical Pharmacogenetics Implementation Consortium (https://cpicpgx.org)

Such variability could influence any pharmacokinetic (absorption, distribution, metabolism, and excretion) or pharmacodynamic parameter of the drug and may have a bearing on the therapeutic outcome as well as the adverse effect risk. Such pharmacogenes may be responsible for the expression of variant proteins (often due to polymorphic variation in coding regions or exon/intron junctions) or may be responsible for the over or underexpression of wild-type gene products (often due to polymorphic variation in regulatory regions such as promoters).

Although the existence of differences in allelic frequencies between males and females is well recognized (Lucotte et al. 2016), the effect of sex on pharmacogenetic outcomes is poorly studied even to date. Within the drug development arena, preclinical functional molecular pharmacology studies rarely consider the gender of cell cultures, pre-clinical animal studies are mainly conducted in males, and there is little sex-dependent conclusive pharmacogenetic data emerging from premarketing clinical trials. Most data available has emerged from postmarketing clinical academic research.

The study of sex-influenced pharmacogenetic outcomes is confounded by the fact that very few known pharmacogenes are located on the sex chromosomes. Therefore, male/female pharmacogenetic differences are mainly autosomal in origin and may also be epigenetically influenced. More specifically, these may relate to (a) differences in autosomal allelic frequencies between males and females (Zuo et al. 2015); sexually dimorphic allelic frequencies have been frequently reported but are poorly understood (b) sex-influenced gene expression, which may or may not be hormonally controlled, (c) epigenetic differences between males and females, such as differences in DNA methylation and histone modifications (Ratnu et al. 2017), and (d) male/female differences in phenotypes of behavior, especially in the case of drugs which act on the central nervous system. For example, males and females may perceive pain differently and the effect of the same pharmacogenetic variant which influences the efficacy of an analgesic drug may therefore be reported differently by male and female patients.

There have been difficulties in establishing with statistical certainty the sex-dependent penetrance of pharmacogenetic variants. For example, estrogen receptor (*ESR1*) and apolipoprotein (*APOA1*) gene polymorphisms have been reported to exert a sex-dependent effect on statin responses to High Density Lipoprotein (HDL), Low Density Lipoprotein (LDL), total cholesterol, and triacylglycerol profiles (Kajinami et al. 2005; Smiderle et al. 2016). However sex-dependent findings were not evident in apolipoprotein E (*APOE*) gene variants, which also influence statin-mediated reductions in LDL–cholesterol (Mega et al. 2009). A later conducted Genome Wide Association Study (GWAS) found strong associations with the ATP-binding cassette subfamily G member 2 (*ABCG2*), lipoprotein a (*LPA*), and *APOE* genes and statin therapeutic outcomes but no genome wide statistical significance was attained with the sex covariate (Chasman et al. 2012).

Perhaps one of the better understood mechanisms of sex-dependent pharmacogenetics is the methylation-driven epigenetic influence on gene expression. Some DNA methylases and demethylases are sex chromosome-linked, and transcriptionally regulated by estrogens (Gaignebet and Kararigas 2017; Ratnu et al. 2017). This contributes to a sex-dependent expression profile. This sex-dependent epigenetic control may be further influenced by behavioral differences, such as sleep patterns and dietary habits, which may themselves be gender-related, and also contribute to the epigenetic control of specific genes. Table 4 lists examples of pharmacogene variants which have been reported to express a sexually dimorphic drug response phenotype.

Interestingly, there is little evidence to suggest functional sexual diversity in the genotype frequencies among Phase I metabolism CYP450 genes. As discussed earlier, sexual dimorphism in Phase I metabolism appears to be mainly due to sex-dependent expression rather than sex-dependent allelic diversity. Although male/female differences in frequencies of functional alleles of CYP450 genes may exist, their potential contribution to SGDs appears to be overshadowed by stronger sex-dependent physiological variables

Table 4 Examples of pharmacogenetic variants and reported sexually dimorphic phenotypes. The variant nomenclature has been kept the same to that used by the authors. Specific alleles have been provided where stated by the authors (*ACE* angiotensin converting enzyme, *ACEi* angiotensin converting enzyme inhibitor)

Drug class	Gene	Variant	Sexual dimorphism	Reference
General antidepressants	*PPGAL*	rs948854 [promoter] T-allele	Decreased response of antidepressant treatment in females	(Unschuld et al. 2010)
Simvastatin	*SLCO1B1*	rs4363657 C-allele	Higher odds ratios of males with respect to myopathy associated with the CC genotype	(Group et al. 2008)
Atorvastatin	*MDR1 (ABCB1)*	G2677 T/A	Females carrying at least one 2677 T or 2677A allele showed significantly higher LDL levels than wild type. This was not seen in men.	(Kajinami et al. 2004)
Simvastatin	*INSIG1*	rs73161338 A-allele	Associated with lower statin-induced *INSIG1* expression and greater reductions in triglyceride levels in males.	(Theusch et al. 2017)
ACE inhibitors[a]	*ACE*	rs4344 allele	Increased risk of ACEi-induced cough in females	(Grilo et al. 2011)
ACE inhibitors[a]	*ACE*	rs4646994 (indel Alu repetitive element)	Increased risk of ACEi-induced cough in females	(Grilo et al. 2011)
Captopril	*ACE2*	rs2106809 T-allele	Reduced captopril antihypertensive response and higher risk of hypertension in women than men.	(Fan et al. 2007)
ACE inhibitors[a]	*XPNPEP2*	C2399A	"A" allele associated with an increased risk of ACEi-associated angioedema in men but not in women	(Woodard-Grice et al. 2010)

[a]Authors did not specify which ACE-inhibitor drugs were being administered

that influence their expression. For example, the low-expressing *CYP3A4∗22* (rs35599367, C > T) allele shows no sex-dependent phenotypic contribution though it markedly reduces expression of *CYP3A4* (Wang et al. 2011). Fujita and coworkers studied the effects of various *CYP2B6* and *UGT1A9* alleles on measured blood propofol concentrations 4 h post initial target-controlled infusion, together with sex-dependent effects. Measured propofol concentrations in females were closer to predicted values than in males; however, *CYP2B6* and *UGT1A9* genotypes were not statistically different between the sexes (Fujita et al. 2018). Similarly, multivariate analysis which included multiple *CYP2B6*, *CYP3A5*, and *ABCB1* genotypes, age, sex, body weight, albumin, alanine transaminase (ALT), urea, and serum creatinine, identified the three alleles *CYP2B6∗6*, *CYP2B6∗11*, *ABCB1* rs3842 and sex as major predictors of plasma concentrations of the antiretroviral efavirenz following single-dose administration. However, there was no evidence of any of the individual allele effects being sex-dependent (Mukonzo et al. 2009).

Regulatory Aspects

Food and Drug Administration (FDA)

The US Food and Drug Administration (FDA) clearly recognizes the importance of assessing sex-gender differences in drug therapy, and it acknowledges that "understanding mechanisms of sex differences in medical product development is crucial for regulatory decisions and optimal treatment outcomes."[4] A chronology of the development of sex-gender recommendations is provided on the FDA website under "Regulations, Guidance, and Reports related to Women's Health"[5] and Table 5 lists selected milestones.

The FDA assigns importance to three central questions for assessing sex-gender influences

[4]https://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm131182.htm

[5]https://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm472932.htm

Table 5 Selected milestones in the FDA-driven recognition and recommendations related to sex-gender differences in drug therapy

Year	Recommendation	Reference
1987	Both sexes of animals should be included in preclinical drug safety studies for products targeted for use by both sexes	Nonclinical pharmacology/toxicology section of an application[a]
1988	Recommends data analysis of clinical pharmacology studies safety, effectiveness and by sex, race and age. – guidance	1988 FDA guidance for the format and content of the clinical and statistical section of an application[b]
1993	Women should be allowed to determine for themselves the appropriateness of participating in early clinical trials Sponsors should collect gender-related data during research and development and should analyze the data for gender effects in addition to other variables such as age and race Sponsors should include a fair representation of both genders as participants in clinical trials so that clinically significant gender-related differences in response can be detected	Guideline for the study and evaluation of gender differences in the clinical evaluation of drugs[c]
1997	FDA and NIH along with representatives of the drug manufacturing industry should review and develop guidance on inclusion of women and minorities in clinical trials	FDAMA sect. 115: Clinical investigations (b) women and minorities regulation[d]
1998	Investigational new drug (IND) data regarding participation in clinical trials should be presented in annual reports by sex, age, and race	Investigational new Drug (IND) applications - annual reports (21 CFR 312.33)[e]
1999	Dose adjustments should be based on multiple factors such as gender, smoking habit, creatinine clearance, and body weight	FDA guidance for industry population pharmacokinetics[f]
2014	Recommendations for improving the completeness and quality of analyses of data on demographic subgroups in the summaries of product safety and effectiveness, and in labeling Recommendations for improving the public availability of such data to patients, health care providers, and researchers	FDA action plan to enhance the collection and availability of Demographic subgroup data[g]
2014	Outlining of FDA's expectations regarding sex-specific patient enrolment, data analysis, and reporting of study information for medical device applications	Evaluation of sex-specific data in medical device clinical studies: Guidance for industry and FDA staff[h]
2016	Initiative to raise awareness about the importance of participation of diverse groups of women in clinical research, and to share best practices about clinical research design, recruitment, and subpopulation analyses	Diverse women in clinical trial initiative OWH in collaboration with the NIH Office on research in Women's health[i]

[a] https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM079234.pdf

[b] https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/WomensHealthResearch/UCM131200.pdf

[c] https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/WomensHealthResearch/UCM131204.pdf

[d] https://www.fda.gov/downloads/RegulatoryInformation/Legislation/SignificantAmendmentstotheFDCAct/FDAMA/FullTextofFDAMAlaw/UCM089145.pdf

[e] https://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm472932.htm

[f] https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM072137.pdf

[g] https://www.fda.gov/downloads/RegulatoryInformation/Legislation/SignificantAmendmentstotheFDCAct/FDASIA/UCM410474.pdf

[h] https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM283707.pdf

[i] https://www.fda.gov/ForConsumers/ByAudience/ForWomen/ucm118508.htm

during drug development. The first asks how to best determine if and when a dosage regimen should be adjusted for males or females. The second asks at what stage or stages during the drug development process should information on such an adjustment be elaborated, and the third asks what magnitude of a sex-gender difference should be sufficient to warrant inclusion in a drug label with recommendations for dosing adjustments. In order to address these questions, the FDA recommends that male/female differences be addressed at multiple stages, including[6]:

- *In vitro* studies in human or nonhuman tissue
- *In vivo* studies in nonhuman species
- Early phase clinical trials and late-phase confirmatory studies in humans
- Postmarketing studies

In order to attain these targets, the FDA advocates for sufficient numbers of both sexes to be included in order for studies to be sufficiently statistically powered to detect potential differences. It also suggests that a suitable panel of pharmacokinetic endpoints could potentially be identified, which can then be used to signal the presence or absence of a sex-gender-related effect for a particular drug. If present, this would encourage more detailed sex-gender studies on the drug in question in order to investigate the male/female impact on safety and efficacy.

The FDA has also identified drugs where enough clinical evidence exists to recommend male/female dose adjustments and has included these recommendations in the drug label. For example, zolpidem has been recommended to be used at a dose of 1.75 mg in women and 3.5 mg in adult men due to a 45% higher female C_{max} and area under the curve (AUC) values at equivalent male doses (label update: 6 February 2013). Information has also been included in the labels of other drugs where sex-gender differences have been studied but not enough evidence was found to warrant sex-based dose adjustments. For example, golimumab showed 13% higher clearance in females than in males after weight adjustment; however, the clinical outcomes were not different (label update: 27 December 2013). Ezogabine showed AUC values that were 20–30% higher and C_{max} values that were 50–100% higher in females than males, across both young and old-age categories, with however no sex-gender differences in weight-adjusted clearance of this drug (label update: 6 September 2013) (Table 6).

European Medicines Agency (EMA)

Within the European regulatory framework, the recognition of sex-gender differences in pharmacological treatment has closely followed US trends.

The European Medicines Agency promotes the consideration of SGDs in the majority of its guidelines and reflection documents. Official documents, such as the EMA guidelines for drug development for the management of alcohol dependence,[7] primary osteoporosis,[8] hypertension,[9] Alzheimer's disease,[10] and urinary incontinence,[11] to name a few, all include specific sex-stratification recommendations, in order to identify potentially sexually dimorphic drug responses early during development. Sex is also an integral data field and a primary stratification variable of the EMA EudraVigilance[12] suspected ADR reporting system.

[6] https://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm134469.htm

[7] https://www.ema.europa.eu/documents/scientific-guideline/guideline-development-medicinal-products-treatment-alcohol-dependence_en.pdf

[8] https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-primary-osteoporosis_en.pdf

[9] https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-primary-osteoporosis_en.pdf

[10] https://www.ema.europa.eu/documents/scientific-guideline/guideline-medicinal-products-treatment-alzheimers-disease-other-dementias-revision-1_en.pdf

[11] https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-urinary-incontinence_en.pdf

[12] http://www.adrreports.eu/en/index.html

Table 6 Examples of drugs for which the FDA includes sex-dependent pharmacological variation information in the drug label. Tradenames are given in brackets (C_{max} peak plasma concentration, *AUC* area under the curve). (Adapted from Fadiran and Zhang 2015)

Affected drugs	Drug approval	Cited approval label	Label statement
Bevacizumab (Avastin)	26 Feb 2004	16 Dec 2013	Males have a higher body-weight corrected clearance than females
Darifencin hydrobromide (Enablex)	22 Dec 2004	15 Mar 2012	C_{max} and AUC at steady-state are higher in females than in males
Enfuvirtide (Fuzeon)	13 Mar 2003	31 Oct 2013	Lower clearance in females than males
Perflexane phospholipidmicrospheres (Imagent)	31 May 2002	31 May 2012	Females eliminate perflexane through the expired air more slowly than males
Zolpidem (intermezzo)	23 Nov 2011	6 Feb 2013	Clearance following sublingual administration is slower in females, leading to higher Cmax and AUC values
Mirabegron (Myrbetriq)	28 Jun 2012	28 Jun 2012	Body weight-corrected systemic exposure based on C_{max} and AUC is 20–30% higher in females
Trospium chloride (Sanctura)	28 May 2004	23 Jul 2012	Conflicting evidence. Single daily dosing reported a C_{max} and AUC based lower exposure in elderly females than elderly males. Divided doses, however, reported higher exposure in elderly females
Lisdexamfetamine dimesylate (Vyvanse)	23 Feb 2007	6 Dec 2013	In adults aged 55-64y, C_{max} and AUC were 15% and 13% higher respectively, in females compared to males

International Council for Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH)

The ICH,[13] a major scientific, technical, and regulatory organization, whose founding members include the European Commission, the FDA, and the MHLW/PMDA (Pharmaceuticals and Medical Devices Safety Information division of the Ministry of Health, Labor and Welfare, Japan), strongly promotes the importance of recognizing SGDs in drug development, through several of its documents.

In "Sex-related considerations in the conduct of clinical trials,"[14] the ICH emphasizes that "patients entering clinical trials should be reasonably representative of the population that will be later treated by the drug, as subpopulations may respond differently to a given drug treatment." This is supported by several later ICH guidelines, especially those related to drug efficacy (ICH M4E part of the M4 Common Technical Document),[15] dose response (ICH E4), "General considerations for clinical trials" (ICH E8),[16] and the "Structure and content of clinical study reports" (ICH E3),[17] which specifically recommend the inclusion of sex as an important demographic variable, and which should be used for stratification analysis where sufficient statistical power exists.

Conclusion

There is increasing evidence of sex-gender influences in therapeutic outcomes, and regulatory agencieshave repeatedly recommended that this

[13] https://www.ich.org

[14] https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Consideration_documents/ICH_Women_Revised_2009.pdf

[15] https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4E_R2_Efficacy/M4E_R2__Step_4.pdf

[16] https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E8/Step4/E8_Guideline.pdf

[17] https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf

should be addressed during the drug-development process. However, this area has unfortunately remained more of an academic activity rather than an applied endeavor with prescribing relevance. Few, if any, therapeutic guidelines actually recommend sex-dependent prescribing modifications.

The issues that potentially contribute to this are many (Table 7). Perhaps the primary issue relates to the multiplicity of variables that actually contribute to SGDs in therapeutics. This leads to a difficulty in extracting meaningful data that can be successfully applied to patients within different age groups, ethnicities, comorbidities, and polypharmacy regimens. Moreover, pharmacological SGDs which are hormonally dependent may have outcomes which present differently (a) during prepubertal age, (b) postpubertal, with dependency on cyclical hormonal states, and (c) during the menopausal and postmenopausal age. These multiple variables require clinical studies to be designed with a high level of stratification and statistical power in order to identify differential outcomes. This adds a significant cost to the clinical trial phases which normally already make up a considerable financial component of the drug development process. Notwithstanding, knowledge of SGDs in pharmacological therapy has applications in personalized prescribing. However, in order for such data to be useful, a multivariate model, which includes all respective known contributing factors, may need to be established, at least for drugs with well-known sexually dimorphic responses. For example, a multivariate model was used to explain plasma clozapine changes in male and female smokers and nonsmokers, using variables which also included age and body weight. Female sex was found to contribute to 17% increased plasma levels in this model. The authors generated nomograms which included multiple variables, which could be used to individualize dosage to achieve a closer target plasma clozapine concentration (Rostami-Hodjegan et al. 2004).

Table 7 Some issues that hinder evidence-based identification of sex-gender differences in pharmacological responses. (*SGD* sex-gender difference)

	Problems associated with evidence-based identification of pharmacological SGDs
1.	Multiplicity of variables that potentially contribute to SGDs in therapeutics, making an inclusive multivariate analysis difficult
2.	Number of candidates required for sufficient statistically powered patient stratification in clinical trials
3.	Lack of understanding of the molecular or physiological mechanisms leading to specific SGDs
4.	Risk of involving women of childbearing age in clinical trials
5.	Lack of any mandatory requirement for the pharmaceutical industry to incorporate the study of SGDs in the drug development process
6.	Pharmacoeconomical implications of the marketed product
7.	Potential of SGDs to be individual drug related rather than a drug class effect
8.	Difficulty of extrapolation of pharmacological monotherapy SGD data to multi-drug prescribing scenarios

There still also exist the risks of involving women of childbearing age within a clinical trial. The willingness of such women to participate may be weakened by their perception of unknown risks which may potentially exist for longer than the treatment period. This is compounded by the knowledge that for several years, women have been omitted from the clinical trial scenario, except for drugs which are specifically targeted to females, such as hormonal therapy.

Proposals for the Future

Clinical therapeutics has expanded in several directions and, among other fronts, has significantly progressed in the area of personalized medicine. This is perhaps most clearly evident in the pharmacogenomics and biomarker development areas, where predictive testing is used for individualized patient-specific focused prescribing.

Sexual dimorphism in therapeutics provides a primary opportunity for sex and gender to become an integral component of personalized therapy approaches. This necessitates integration at multiple levels. Kim and coworkers suggested that scientific journal editors should require authors to clearly label single-sex studies and justify the

rationale for pursuing them (Kim et al. 2010). With respect to new drugs under development, (a) early preclinical molecular pharmacology studies need to be carried out in cell line or primary cell models from both male and female donors, (b) preclinical studies need to be carried out in both male and female animal models, (c) the FDA recommendations[18] need to be followed with respect to the clinical phases of drug development, and both genders need to be incorporated into the clinical trial structure. This holds both for clinical outcome studies as well as for drug development pharmacogenomics. Indeed, the expert document entitled "Sex and Gender Analysis in Medical and Pharmacological Research" of the United Nations Division for the Advancement of Women states that "Efforts to avoid bias are as important when reviewing the results of animal models as when reviewing the results of clinical trials."[19] Furthermore, funding agencies and institutional review boards should give preference to drug studies that include sex and/or gender as a component of the study design – a point that has already recently been emphasized by McGregor and coworkers (McGregor et al. 2016).

With respect to drugs already in the post-marketing phase, researchers need to incorporate suitably powered sex and gender variables into clinical research, in order to generate data which may have applicability in the clinical field. Sexual dimorphism data which originate from pharmacovigilance reports constitute *posthoc* reporting which may lack the robustness of pre-designed studies. Therefore, pharmacovigilance reports may be potentially used as pointers to indicate which currently marketed drugs need to be studied for sexually dimorphic outcomes.

The financial implications of such research are also a consideration. Drug development is a costly and high-risk process, which needs to be commercially profitable to the pharmaceutical industry. Additional developmental costs are not welcome, especially when the outcome data are not *mandated* but simply *recommended* by regulators and may thus be considered unnecessary by the industry. Mitigation of this may therefore need to be promoted on multiple fronts: (a) there needs to be stronger pressure by regulatory agencies on the industry, to incorporate the identification of potential SGDs in drug development, (b) early animal studies and Phase I clinical trials could be used to identify the drug candidates that are most likely to merit SGD studies in Phase II/III trials, (c) financial incentives (e.g., tax incentives) could be offered to the pharmaceutical industry to specifically promote SGD studies without increasing the final marketed cost of the drug.

SGD drug therapy data needs to be validated and incorporated into management guidelines. The establishment of a knowledgebase, populated with evidence-based data and curated by a panel from the scientific and drug regulatory areas, would help to transfer such knowledge to management guideline developers and prescribers. Moreover, the incorporation of such data into regulatory agency-approved drug labels, further contributes to a sex-gender component in the personalized medicine portfolio. There are various informative resources which are intended to guide the incorporation of sex and gender into clinical research design and analysis (McGregor et al. 2016). Of particular interest is the position taken by the US National Institutes of Health, which since 2016 have required all grant applicants to address the role of sex as a biological variable in vertebrate animal and human studies.

These changes need to be implemented in a strategic manner, prioritizing areas where SGDs are most likely to contribute to therapeutic outcomes. Therefore, there is need for internationally accepted guidelines and agreed priority areas. The availability of trained teams capable of assisting investigators to integrate suitably powered and relevant sex and gender components, into research project development is recommended. Research, regulatory and funding agencies perhaps constitute the best tripartite

[18] https://www.fda.gov/science-research/womens-health-research/gender-studies-product-development-scientific-issues-and-approaches

[19] https://www.un.org/womenwatch/daw/egm/gst_2010/Franconi-EP.12-EGM-ST.pdf

collaborating structure, which can potentially deliver the necessary drivers to guide the inclusion of sex and gender into therapeutic research and practice.

Challenges

Perhaps the most important and least costly challenge to improve sex-gender-related therapeutic outcomes in personalized medicine, is the promotion of awareness among prescribers. Prescribers need to be aware of sex-gender differences that may be expected in the therapeutic and adverse effect outcomes of the drugs they prescribe. They also need to be provided with the information to mitigate such SGDs when necessary, such as dosing modifications or cycle-dependent prescribing.

Increased prescriber awareness also leads to an improved adverse drug reaction reporting activity, through which suspected sexually dimorphic reactions may be flagged in the prescriber's report. This enhances centralized pharmacovigilance data collection and subsequent flagging of drugs which merit further SGD research.

The problem of designing suitably powered studies from which clinically applicable data can be derived remains a challenge. Age, ethnicity, comorbidity, pharmacogenetic variability, hormonal cyclical variability, and gender-dependent attitudes in self-reporting of adverse drug reactions are but a few of the variables that need to be considered. Data obtained from pharmacovigilance reporting may be biased due to higher tendencies of females than males to engage in post-general marketing surveillance activities (Harmark et al. 2013) and report ADRs to their physician. Furthermore, SGDs may or may not be attributable to drug class effects. For example, a study establishing SGDs with respect to a specific statin does not necessarily imply the same effect occurring with any other current or future-developed statin, making data extrapolation across similar compounds unfeasible.

Finally, pharmacoeconomic aspects need to be also considered. Therapeutic benefit-risk ratios need to be viewed in the light of cost-risk ratios. Any increased financial burden on patients or national health services must be justified by evidence-based benefits originating from well-structured and powered studies.

Pharmacological SGDs deserve to be an integral component of the personalized medicine framework. In order to exploit the maximum therapeutic benefit from the knowledge of such differences, the three key actions remain (a) prescriber awareness, (b) evidence-based SGD identification, and (c) the integration of relevant SGD-based prescribing into disease management guidelines.

References and Further Reading

Austin KL, Mather LE, Philpot CR, McDonald PJ (1980) Intersubject and dose-related variability after intravenous administration of erythromycin. Br J Clin Pharmacol 10(3):273–279. https://doi.org/10.1111/j.1365-2125.1980.tb01755.x

Bebawy M, Chetty M (2009) Gender differences in p-glycoprotein expression and function: effects on drug disposition and outcome. Curr Drug Metab 10(4):322–328

Becker JB, McClellan ML, Reed BG (2017) Sex differences, gender and addiction. J Neurosci Res 95(1–2):136–147. https://doi.org/10.1002/jnr.23963

Berg UB (2006) Differences in decline in GFR with age between males and females. Reference data on clearances of inulin and PAH in potential kidney donors. Nephrol Dial Transplant Off Publ Eur Dial Transplant Assoc Eur Renal Assoc 21(9):2577–2582. https://doi.org/10.1093/ndt/gfl227

Billimek J, Malik S, Sorkin DH, Schmalbach P, Ngo-Metzger Q, Greenfield S, Kaplan SH (2015) Understanding disparities in lipid management among patients with type 2 diabetes: gender differences in medication nonadherence after treatment intensification. Women's Health Issues Off Publ Jacobs Inst Women's Health 25(1):6–12. https://doi.org/10.1016/j.whi.2014.09.004

Cabaleiro T, Roman M, Ochoa D, Talegon M, Prieto-Perez R, Wojnicz A, Lopez-Rodriguez R, Novalbos J, Abad-Santos F (2013) Evaluation of the relationship between sex, polymorphisms in CYP2C8 and CYP2C9, and pharmacokinetics of angiotensin receptor blockers. Drug Metab Dispos 41(1):224–229. https://doi.org/10.1124/dmd.112.046292

Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM (2012) Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the justification for the use of statins in prevention: an intervention trial evaluating Rosuvastatin (JUPITER)

trial. Circ Cardiovasc Genet 5(2):257–264. https://doi.org/10.1161/CIRCGENETICS.111.961144

Coskun T, Sevinc A, Tevetoglu I, Alican I, Kurtel H, Yegen BC (1995) Delayed gastric emptying in conscious male rats following chronic estrogen and progesterone treatment. Research in experimental medicine Zeitschrift fur die gesamte experimentelle Medizin einschliesslich experimenteller Chirurgie 195(1):49–54

Costa E, Giardini A, Savin M, Menditto E, Lehane E, Laosa O, Pecorelli S, Monaco A, Marengoni A (2015) Interventional tools to improve medication adherence: review of literature. Patient Prefer Adherence 9:1303–1314. https://doi.org/10.2147/PPA.S87551

Damien S, Patural H, Trombert-Paviot B, Beyens MN (2016) Adverse drug reactions in children: 10 years of pharmacovigilance. Archives de pediatrie: organe officiel de la Societe francaise de pediatrie 23(5):468–476. https://doi.org/10.1016/j.arcped.2016.01.015

Dauchy S, Dolbeault S, Reich M (2013) Depression in cancer patients. EJC Suppl EJC Off J EORTC European Organization for Research and Treatment of Cancer [et al] 11(2):205–215. https://doi.org/10.1016/j.ejcsup.2013.07.006

d'Emden MC, Jenkins AJ, Li L, Zannino D, Mann KP, Best JD, Stuckey BG, Park K, Saltevo J, Keech AC, Investigators FS (2014) Favourable effects of fenofibrate on lipids and cardiovascular disease in women with type 2 diabetes: results from the Fenofibrate intervention and event lowering in diabetes (FIELD) study. Diabetologia 57(11):2296–2303. https://doi.org/10.1007/s00125-014-3344-3

Enck P, Klosterhalfen S (2019) Does sex/gender play a role in placebo and nocebo effects? Conflicting evidence from clinical trials and experimental studies. Front Neurosci 13:160. https://doi.org/10.3389/fnins.2019.00160

Fadiran EO, Zhang L (2015) Effects of sex differences in the pharmacokinetics of drugs and their impact on the safety of medicines in women. In: Harrison-Woolrych M (ed) Medicines for women. Springer International Publishing, Cham, pp 41–68. https://doi.org/10.1007/978-3-319-12406-3_2

Fan X, Wang Y, Sun K, Zhang W, Yang X, Wang S, Zhen Y, Wang J, Li W, Han Y, Liu T, Wang X, Chen J, Wu H, Hui R, Study Group for Pharmacogenomic Based Antihypertensive Drugs Selection E, Side Effects iRAC (2007) Polymorphisms of ACE2 gene are associated with essential hypertension and antihypertensive effects of captopril in women. Clin Pharmacol Ther 82(2):187–196. https://doi.org/10.1038/sj.clpt.6100214

Feldman M, Richardson CT, Walsh JH (1983) Sex-related differences in gastrin release and parietal cell sensitivity to gastrin in healthy human beings. J Clin Invest 71(3):715–720. https://doi.org/10.1172/jci110818

Fischer M, Baessler A, Schunkert H (2002) Renin angiotensin system and gender differences in the cardiovascular system. Cardiovasc Res 53(3):672–677

Food US, Drug A (1993) Guideline for the study and evaluation of gender differences in the clinical evaluation of drugs; notice. Fed Regist 58(139):39406–39416

Franconi F, Carru C, Malorni W, Vella S, Mercuro G (2011) The effect of sex/gender on cardiovascular pharmacology. Curr Pharm Des 17(11):1095–1107

Fujita A, Hayamizu K, Yoshihara T, Zaitsu M, Shiraishi F, Arima H, Matsuo K, Shiokawa K, Setoguchi H, Sasaguri T (2018) Effect of sex and polymorphisms of CYP2B6 and UGT1A9 on the difference between the target-controlled infusion predicted and measured plasma propofol concentration. JA Clin Rep 4(1):59. https://doi.org/10.1186/s40981-018-0196-8

Gabler NB, French B, Strom BL, Liu Z, Palevsky HI, Taichman DB, Kawut SM, Halpern SD (2012) Race and sex differences in response to endothelin receptor antagonists for pulmonary arterial hypertension. Chest 141(1):20–26. https://doi.org/10.1378/chest.11-0404

Gabriele L, Buoncervello M, Ascione B, Bellenghi M, Matarrese P, Care A (2016) The gender perspective in cancer research and therapy: novel insights and on-going hypotheses. Annali dell'Istituto superiore di sanita 52(2):213–222. https://doi.org/10.4415/ANN_16_02_13

Gaignebet L, Kararigas G (2017) En route to precision medicine through the integration of biological sex into pharmacogenomics. Clin Sci 131(4):329–342. https://doi.org/10.1042/CS20160379

Ginsberg HN, Elam MB, Lovato LC, Crouse JR, 3rd, Leiter LA, Linz P, Friedewald WT, Buse JB, Gerstein HC, Probstfield J, Grimm RH, Ismail-Beigi F, Bigger JT, Goff DC Jr, Cushman WC, Simons-Morton DG, Byington RP (2010) Effects of combination lipid therapy in type 2 diabetes mellitus. N Engl J Med 362(17):1563–1574. https://doi.org/10.1056/NEJMoa1001282

Gonzalez D, Conrado DJ, Theuretzbacher U, Derendorf H (2011) The effect of critical illness on drug distribution. Curr Pharm Biotechnol 12(12):2030–2036

Grilo A, Saez-Rosas MP, Santos-Morano J, Sanchez E, Moreno-Rey C, Real LM, Ramirez-Lorca R, Saez ME (2011) Identification of genetic factors associated with susceptibility to angiotensin-converting enzyme inhibitors-induced cough. Pharmacogenet Genomics 21(1):10–17. https://doi.org/10.1097/FPC.0b013e328341041c

Group SC, Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, Gut I, Lathrop M, Collins R (2008) SLCO1B1 variants and statin-induced myopathy–a genomewide study. N Engl J Med 359(8):789–799. https://doi.org/10.1056/NEJMoa0801936

Harmark L, Lie-Kwie M, Berm L, de Gier H, van Grootheest K (2013) Patients' motives for participating in active post-marketing surveillance. Pharmacoepidemiol Drug Saf 22(1):70–76. https://doi.org/10.1002/pds.3327

Ishigami M, Kawabata K, Takasaki W, Ikeda T, Komai T, Ito K, Sugiyama Y (2001) Drug interaction between simvastatin and itraconazole in male and female rats. Drug Metab Dispos 29(7):1068–1072

Jankovic SM, Pejcic AV, Milosavljevic MN, Opancina VD, Pesic NV, Nedeljkovic TT, Babic GM (2018) Risk factors for potential drug-drug interactions in

intensive care unit patients. J Crit Care 43:1–6. https://doi.org/10.1016/j.jcrc.2017.08.021

Jessurun N, van Marum R, Hermens W, van Puijenbroek E (2016) Advanced age and female sex as risk factors for high anion gap metabolic acidosis after a Drug interaction between paracetamol and Flucloxacillin: a case series. J Am Geriatr Soc 64(10):e90–e93. https://doi.org/10.1111/jgs.14332

Jochmann N, Stangl K, Garbe E, Baumann G, Stangl V (2005) Female-specific aspects in the pharmacotherapy of chronic cardiovascular diseases. Eur Heart J 26(16):1585–1595. https://doi.org/10.1093/eurheartj/ehi397

Kajinami K, Brousseau ME, Ordovas JM, Schaefer EJ (2004) Polymorphisms in the multidrug resistance-1 (MDR1) gene influence the response to atorvastatin treatment in a gender-specific manner. Am J Cardiol 93(8):1046–1050. https://doi.org/10.1016/j.amjcard.2004.01.014

Kajinami K, Brousseau ME, Lamon-Fava S, Ordovas JM, Schaefer EJ (2005) Gender-specific effects of estrogen receptor alpha gene haplotype on high-density lipoprotein cholesterol response to atorvastatin: interaction with apolipoprotein AI gene polymorphism. Atherosclerosis 178(2):331–338. https://doi.org/10.1016/j.atherosclerosis.2004.08.034

Karagiannis TC (2014) The timeless influence of Hippocratic ideals on diet, salicytates and personalized medicine. Hell J Nucl Med 17(1):2–6. https://doi.org/10.1967/s0024499100110

Kautzky-Willer A, Harreiter J (2017) Sex and gender differences in therapy of type 2 diabetes. Diabetes Res Clin Pract 131:230–241. https://doi.org/10.1016/j.diabres.2017.07.012

Kim RB, Leake BF, Choo EF, Dresser GK, Kubba SV, Schwarz UI, Taylor A, Xie HG, McKinsey J, Zhou S, Lan LB, Schuetz JD, Schuetz EG, Wilkinson GR (2001) Identification of functionally variant MDR1 alleles among European Americans and African Americans. Clin Pharmacol Ther 70(2):189–199. https://doi.org/10.1067/mcp.2001.117412

Kim AM, Tingen CM, Woodruff TK (2010) Sex bias in trials and treatment must end. Nature 465(7299):688–689. https://doi.org/10.1038/465688a

Kishino S, Nomura A, Itoh S, Nakagawa T, Takekuma Y, Sugawara M, Furukawa H, Todo S, Miyazaki K (2002) Age- and gender-related differences in carbohydrate concentrations of alpha1-acid glycoprotein variants and the effects of glycoforms on their drug-binding capacities. Eur J Clin Pharmacol 58(9):621–628. https://doi.org/10.1007/s00228-002-0530-x

Kloner RA, Sowers JR, DiBona GF, Gaffney M, Wein M (1996) Sex- and age-related antihypertensive effects of amlodipine. The amlodipine cardiovascular community trial study group. Am J Cardiol 77(9):713–722

Kneale BJ, Chowienczyk PJ, Brett SE, Coltart DJ, Ritter JM (2000) Gender differences in sensitivity to adrenergic agonists of forearm resistance vasculature. J Am Coll Cardiol 36(4):1233–1238

Kornstein SG, Schatzberg AF, Thase ME, Yonkers KA, McCullough JP, Keitner GI, Gelenberg AJ, Davis SM, Harrison WM, Keller MB (2000) Gender differences in treatment response to sertraline versus imipramine in chronic depression. Am J Psychiatry 157(9):1445–1452. https://doi.org/10.1176/appi.ajp.157.9.1445

Krecic-Shepard ME, Barnas CR, Slimko J, Schwartz JB (2000) Faster clearance of sustained release verapamil in men versus women: continuing observations on sex-specific differences after oral administration of verapamil. Clin Pharmacol Ther 68(3):286–292. https://doi.org/10.1067/mcp.2000.109356

Lehmann MH, Hardy S, Archibald D, quart B, MacNeil DJ (1996) Sex difference in risk of torsade de pointes with d,l-sotalol. Circulation 94(10):2535–2541

Li Q, Zhou T, Wu F, Li N, Wang R, Zhao Q, Ma YM, Zhang JQ, Ma BL (2018) Subcellular drug distribution: mechanisms and roles in drug efficacy, toxicity, resistance, and targeted delivery. Drug Metab Rev 50(4):430–447. https://doi.org/10.1080/03602532.2018.1512614

Linden W, Vodermaier A, Mackenzie R, Greig D (2012) Anxiety and depression after cancer diagnosis: prevalence rates by cancer type, gender, and age. J Affect Disord 141(2–3):343–351. https://doi.org/10.1016/j.jad.2012.03.025

Lucotte EA, Laurent R, Heyer E, Segurel L, Toupance B (2016) Detection of allelic frequency differences between the sexes in humans: a signature of sexually antagonistic selection. Genome Biol Evol 8(5):1489–1500. https://doi.org/10.1093/gbe/evw090

MacGowan SH, Wilcock GK, Scott M (1998) Effect of gender and apolipoprotein E genotype on response to anticholinesterase therapy in Alzheimer's disease. Int J Geriatr Psychiatry 13(9):625–630

Manni A, Direnzo V, Iaffaldano A, Di Lecce V, Tortorella C, Zoccolella S, Iaffaldano P, Trojano M, Paolicelli D (2017) Gender differences in safety issues during Fingolimod therapy: evidence from a real-life relapsing multiple sclerosis cohort. Brain Behav 7(10):e00804. https://doi.org/10.1002/brb3.804

Manteuffel M, Williams S, Chen W, Verbrugge RR, Pittman DG, Steinkellner A (2014) Influence of patient sex and gender on medication use, adherence, and prescribing alignment with guidelines. J Women's Health 23(2):112–119. https://doi.org/10.1089/jwh.2012.3972

Markle JG, Fish EN (2014) SeXX matters in immunity. Trends Immunol 35(3):97–104. https://doi.org/10.1016/j.it.2013.10.006

McGill JB, Vlajnic A, Knutsen PG, Recklein C, Rimler M, Fisher SJ (2013) Effect of gender on treatment outcomes in type 2 diabetes mellitus. Diabetes Res Clin Pract 102(3):167–174. https://doi.org/10.1016/j.diabres.2013.10.001

McGregor AJ, Hasnain M, Sandberg K, Morrison MF, Berlin M, Trott J (2016) How to study the impact of sex and gender in medical research: a review of resources. Biol Sex Differ 7(Suppl 1):46. https://doi.org/10.1186/s13293-016-0099-1

Mega JL, Morrow DA, Brown A, Cannon CP, Sabatine MS (2009) Identification of genetic variants associated with response to statin therapy. Arterioscler Thromb Vasc Biol 29(9):1310–1315. https://doi.org/10.1161/ATVBAHA.109.188474

Mills PJ, Ziegler MG, Nelesen RA, Kennedy BP (1996) The effects of the menstrual cycle, race, and gender on adrenergic receptors and agonists. Clin Pharmacol Ther 60(1):99–104. https://doi.org/10.1016/S0009-9236(96)90172-1

Montastruc JL, Lapeyre-Mestre M, Bagheri H, Fooladi A (2002) Gender differences in adverse drug reactions: analysis of spontaneous reports to a regional pharmacovigilance Centre in France. Fundam Clin Pharmacol 16(5):343–346

Morissette P, Albert C, Busque S, St-Louis G, Vinet B (2001) In vivo higher glucuronidation of mycophenolic acid in male than in female recipients of a cadaveric kidney allograft and under immunosuppressive therapy with mycophenolate mofetil. Ther Drug Monit 23(5):520–525

Mukonzo JK, Roshammar D, Waako P, Andersson M, Fukasawa T, Milani L, Svensson JO, Ogwal-Okeng J, Gustafsson LL, Aklillu E (2009) A novel polymorphism in ABCB1 gene, CYP2B6∗6 and sex predict single-dose efavirenz population pharmacokinetics in Ugandans. Br J Clin Pharmacol 68(5):690–699. https://doi.org/10.1111/j.1365-2125.2009.03516.x

Nakagawa K, Kajiwara A (2015) Female sex as a risk factor for adverse drug reactions. Nihon Rinsho 73(4):581–585

Nicolas JM, Espie P, Molimard M (2009) Gender and interindividual variability in pharmacokinetics. Drug Metab Rev 41(3):408–421. https://doi.org/10.1080/10837450902891485

Oertelt-Prigione S, Regitz-Zagrosek V (2009) Gender aspects in cardiovascular pharmacology. J Cardiovasc Transl Res 2(3):258–266. https://doi.org/10.1007/s12265-009-9114-9

Pinquart M, Duberstein PR (2010) Depression and cancer mortality: a meta-analysis. Psychol Med 40(11):1797–1810. https://doi.org/10.1017/S0033291709992285

Puri R, Nissen SE, Shao M, Ballantyne CM, Barter PJ, Chapman MJ, Erbel R, Libby P, Raichlen JS, Uno K, Kataoka Y, Nicholls SJ (2014) Sex-related differences of coronary atherosclerosis regression following maximally intensive statin therapy: insights from SATURN. J Am Coll Cardiol Img 7(10):1013–1022. https://doi.org/10.1016/j.jcmg.2014.04.019

Rathore SS, Wang Y, Krumholz HM (2002) Sex-based differences in the effect of digoxin for the treatment of heart failure. N Engl J Med 347(18):1403–1411. https://doi.org/10.1056/NEJMoa021266

Ratnu VS, Emami MR, Bredy TW (2017) Genetic and epigenetic factors underlying sex differences in the regulation of gene expression in the brain. J Neurosci Res 95(1–2):301–310. https://doi.org/10.1002/jnr.23886

Roebuck MC, Liberman JN, Gemmill-Toyama M, Brennan TA (2011) Medication adherence leads to lower health care use and costs despite increased drug spending. Health Aff 30(1):91–99. https://doi.org/10.1377/hlthaff.2009.1087

Rosano GM, Lewis B, Agewall S, Wassmann S, Vitale C, Schmidt H, Drexel H, Patak A, Torp-Pedersen C, Kjeldsen KP, Tamargo J (2015) Gender differences in the effect of cardiovascular drugs: a position document of the working group on pharmacology and Drug therapy of the ESC. Eur Heart J 36(40):2677–2680. https://doi.org/10.1093/eurheartj/ehv161

Rostami-Hodjegan A, Amin AM, Spencer EP, Lennard MS, Tucker GT, Flanagan RJ (2004) Influence of dose, cigarette smoking, age, sex, and metabolic activity on plasma clozapine concentrations: a predictive model and nomograms to aid clozapine dose adjustment and to assess compliance in individual patients. J Clin Psychopharmacol 24(1):70–78. https://doi.org/10.1097/01.jcp.0000106221.36344.4d

Ruan YF, Liu N, Zhou Q, Li Y, Wang L (2004) Experimental study on the mechanism of sex difference in the risk of torsade de pointes. Chin Med J 117(4):538–541

Schmetzer O, Florcken A (2012) Sex differences in the drug therapy for oncologic diseases. Handb Exp Pharmacol 214:411–442. https://doi.org/10.1007/978-3-642-30726-3_19

Schwartz JB (2003) The influence of sex on pharmacokinetics. Clin Pharmacokinet 42(2):107–121. https://doi.org/10.2165/00003088-200342020-00001

Schwartz JB (2007) The current state of knowledge on age, sex, and their interactions on clinical pharmacology. Clin Pharmacol Ther 82(1):87–96. https://doi.org/10.1038/sj.clpt.6100226

Smiderle L, Lima LO, Hutz MH, Van der Sand CR, Van der Sand LC, Ferreira ME, Pires RC, Almeida S, Fiegenbaum M (2014) Evaluation of sexual dimorphism in the efficacy and safety of simvastatin/atorvastatin therapy in a southern Brazilian cohort. Arq Bras Cardiol 103(1):33–40

Smiderle L, Fiegenbaum M, Hutz MH, Van Der Sand CR, Van Der Sand LC, Ferreira ME, Pires RC, Almeida S (2016) ESR1 polymorphisms and statin therapy: a sex-specific approach. Pharmacogenomics J 16(6):507–513. https://doi.org/10.1038/tpj.2015.60

Smith HR (2015) Depression in cancer patients: pathogenesis, implications and treatment (review). Oncol Lett 9(4):1509–1514. https://doi.org/10.3892/ol.2015.2944

Soldin OP, Mattison DR (2009) Sex differences in pharmacokinetics and pharmacodynamics. Clin Pharmacokinet 48(3):143–157. https://doi.org/10.2165/00003088-200948030-00001

Soldin OP, Chung SH, Mattison DR (2011) Sex differences in drug disposition. J Biomed Biotechnol 2011:187103. https://doi.org/10.1155/2011/187103

Spranger M, Aspey BS, Harrison MJ (1989) Sex difference in antithrombotic effect of aspirin. Stroke 20(1):34–37

Spratt KA (1997) Sex- and age-related antihypertensive effects of amlodipine. Am J Cardiol 79(6):843–844

Stolarz AJ, Rusch NJ (2015) Gender differences in cardiovascular drugs. Cardiovasc Drugs Ther 29(4):403–410. https://doi.org/10.1007/s10557-015-6611-8

Stramba-Badiale M, Locati EH, Martinelli A, Courville J, Schwartz PJ (1997) Gender and the relationship between ventricular repolarization and cardiac cycle length during 24-h Holter recordings. Eur Heart J 18(6):1000–1006

Sykiotis GP, Kalliolias GD, Papavassiliou AG (2005) Pharmacogenetic principles in the Hippocratic writings. J Clin Pharmacol 45(11):1218–1220. https://doi.org/10.1177/0091270005281091

Szumlanski CL, Honchel R, Scott MC, Weinshilboum RM (1992) Human liver thiopurine methyltransferase pharmacogenetics: biochemical properties, liver-erythrocyte correlation and presence of isozymes. Pharmacogenetics 2(4):148–159

Theusch E, Kim K, Stevens K, Smith JD, Chen YI, Rotter JI, Nickerson DA, Medina MW (2017) Statin-induced expression change of INSIG1 in lymphoblastoid cell lines correlates with plasma triglyceride statin response in a sex-specific manner. Pharmacogenomics J 17(3):222–229. https://doi.org/10.1038/tpj.2016.12

Trenaman SC, Rideout M, Andrew MK (2019) Sex and gender differences in polypharmacy in persons with dementia: a scoping review. SAGE Open Med 7:2050312119845715. https://doi.org/10.1177/2050312119845715

Unschuld PG, Ising M, Roeske D, Erhardt A, Specht M, Kloiber S, Uhr M, Muller-Myhsok B, Holsboer F, Binder EB (2010) Gender-specific association of galanin polymorphisms with HPA-axis dysregulation, symptom severity, and antidepressant treatment response. Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol 35(7):1583–1592. https://doi.org/10.1038/npp.2010.30

Vambheim SM, Flaten MA (2017) A systematic review of sex differences in the placebo and the nocebo effect. J Pain Res 10:1831–1839. https://doi.org/10.2147/JPR.S134745

Venturini CD, Engroff P, Ely LS, Zago LF, Schroeter G, Gomes I, De Carli GA, Morrone FB (2011) Gender differences, polypharmacy, and potential pharmacological interactions in the elderly. Clinics 66(11):1867–1872. https://doi.org/10.1590/s1807-59322011001100004

Wang D, Guo Y, Wrighton SA, Cooke GE, Sadee W (2011) Intronic polymorphism in CYP3A4 affects hepatic expression and response to statin drugs. Pharmacogenomics J 11(4):274–286. https://doi.org/10.1038/tpj.2010.28

WHO (2003) Adherence to long-term therapies: evidence for action

WHO (2017) Depression and other common mental disorders: global health estimates

Woodard-Grice AV, Lucisano AC, Byrd JB, Stone ER, Simmons WH, Brown NJ (2010) Sex-dependent and race-dependent association of XPNPEP2 C-2399A polymorphism with angiotensin-converting enzyme inhibitor-associated angioedema. Pharmacogenet Genomics 20(9):532–536. https://doi.org/10.1097/FPC.0b013e32833d3acb

Wrobel G, Mauguen A, Rosolen A, Reiter A, Williams D, Horibe K, Brugieres L, Le Deley MC, European Inter-Group for Childhood N-HL (2011) Safety assessment of intensive induction therapy in childhood anaplastic large cell lymphoma: report of the ALCL99 randomised trial. Pediatr Blood Cancer 56(7):1071–1077. https://doi.org/10.1002/pbc.22940

Yang X, Zhang B, Molony C, Chudin E, Hao K, Zhu J, Gaedigk A, Suver C, Zhong H, Leeder JS, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich RG, Slatter JG, Schadt EE, Kasarskis A, Lum PY (2010) Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. Genome Res 20(8):1020–1036. https://doi.org/10.1101/gr.103341.109

Yang L, Li Y, Hong H, Chang CW, Guo LW, Lyn-Cook B, Shi L, Ning B (2012) Sex differences in the expression of Drug-metabolizing and transporter genes in human liver. J Drug Metab Toxicol 3(3):1000119. https://doi.org/10.4172/2157-7609.1000119

Young EA, Kornstein SG, Marcus SM, Harvey AT, Warden D, Wisniewski SR, Balasubramani GK, Fava M, Trivedi MH, John Rush A (2009) Sex differences in response to citalopram: a STAR∗D report. J Psychiatr Res 43(5):503–511. https://doi.org/10.1016/j.jpsychires.2008.07.002

Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol Ther 138(1):103–141. https://doi.org/10.1016/j.pharmthera.2012.12.007

Zhang Y, Klein K, Sugathan A, Nassery N, Dombkowski A, Zanger UM, Waxman DJ (2011) Transcriptional profiling of human liver identifies sex-biased genes associated with polygenic dyslipidemia and coronary artery disease. PLoS One 6(8):e23506. https://doi.org/10.1371/journal.pone.0023506

Zhu B, Liu ZQ, Chen GL, Chen XP, Ou-Yang DS, Wang LS, Huang SL, Tan ZR, Zhou HH (2003) The distribution and gender difference of CYP3A activity in Chinese subjects. Br J Clin Pharmacol 55(3):264–269

Ziegler SM, Beisel C, Sutter K, Griesbeck M, Hildebrandt H, Hagen SH, Dittmer U, Altfeld M (2017) Human pDCs display sex-specific differences in type I interferon subtypes and interferon alpha/beta receptor expression. Eur J Immunol 47(2):251–256. https://doi.org/10.1002/eji.201646725

Zuo L, Wang T, Lin X, Wang J, Tan Y, Wang X, Yu X, Luo X (2015) Sex difference of autosomal alleles in populations of European and African descent. Genes Genomics 37(12):1007–1016. https://doi.org/10.1007/s13258-015-0332-z

Methodologies of PD Assessment: Scales

29

Roman Görtelmeyer

Contents

Abstract

Scales and behavioral tests are the primary source of data in human research for measuring, modeling, and testing perception, thoughts, feelings, complains, and performance. In drug development scales and tests are used in clinical trials as tools to support diagnosis; evaluate states and traits, as well as lifestyle and performance of patients; or collect ratings by relatives and health carers before, during, and after treatment. Clinical scales can be classified according to several criteria; one classification is according to the intended responder and the method of data generation (FDA 2016; Walton et al. Value Health 18:741–752, 2015). Well-developed and ready-for-purpose scales and tests can be used to describe, characterize, or quantify (localize) objects and processes within a framework of a pharmacological clinical study or preferably of a drug development program of interest. In the present article the conceptualization, development, selection, analysis, and application of items and scales in various phases of clinical drug development are presented in an overview and discussed.

Purpose and Rational

Scales and behavioral tests are the primary source of data in human research for measuring, modeling, and testing perception, thoughts, feelings, complains, and performance. In drug development scales and tests (for convenience further subsumed as clinical scale) are used in clinical trials as tools to support diagnosis; evaluate states and traits, as well as lifestyle and performance of patients; and collect ratings by relatives and health carers before, during, and after treatment or for the duration of a clinical study. Clinical scales have been classified according

R. Görtelmeyer (✉)
Scientific Consulting, Frankfurt am Main, Germany

Biological and Clinical Psychology, University of Mannheim, Mannheim, Germany
e-mail: roman.goertelmeyer@t-online.de; goertelm@mail.uni-mannheim.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_27

various criteria; one of these classifications is the qualification according to the intended responder and the method of data generation (FDA 2016; https://www.fda.gov/Drugs/%20DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm; Walton et al. 2015). As such "outcome assessments include survival, clinical outcome assessments (COAs), and biomarkers. A COA is any evaluation that can be influenced by human choices, judgment, or motivation. There are four types of COAs: patient-reported outcome (PRO), clinician-reported outcome (ClinRO), observer-reported outcome (ObsRO), and performance outcome (PerfO)" (Powers et al. 2017). Beyond these qualifications scales may be of great help in exploring the mechanisms of action on human behavior and attitudes. The assessment program of pharmacological intervention comprises evaluation of concept, methods, disease model, definition of (target) population, concept of effect, domains of potential drug response, and benefit and risk of pharmacological intervention and could include estimates of probable success in the market. In most phases of drug development, scales can be efficiently used for differential benefit of the program. Measurement and scaling are fundamental processes in the empirical sciences and of special importance in drug research and development. In natural science measurement may be merely regarded as the use of an existing scale to quantify the object, and scaling can be regarded as "... the assignment of objects to numbers according to a rule" (Stevens 1951). Measurement is only possible if some scale is defined. The use of "scale" in literature is not unique. The term is used in various contexts with different meanings. A more generic definition can be found in dictionaries, e.g., relative size or extent (Oxford English Dictionary).

In the present paper, scale is used synonymously for a standardized and validated questionnaire or behavioral test which delivers a score that has been empirically proven to measure or indicate the underlying characteristic or process of interest. The result of measurement by clinical scales usually is a score on an abstract dimension, e.g., probability, intensity, frequency, as well as change or difference, or the result of scoring and scaling will be transposed into some classification. The way many psychometricians are defining *scale* includes the concept, addressed by some task, a verbal expression, or a question, combined with a concept and design of response option, which may be some kind of differential verbalization or any other ordering of response format that is presented to the respondent. Further, a method for proper quantification of the response and some evidence on reliability and validity is expected. Scale defined in this sense is the standardized quantification of a response of someone in a well-defined test situation (e.g., Lienert and Raatz 1998). The definition of exact standards of measurement is related to units that refer to specific conditions and quantitative attributes. In the natural sciences, the metric international system of measurements (système international d'unités, SI) is used, in which scientifically spoken some quantities are designated as the fundamental units. The first fundamental units, referring to specific empirical conditions and quantitative attributes, were:

Meter (m) SI unit of length
Second (s) SI unit of time
Kilogram (kg) SI unit of mass
Kelvin (K) SI unit of temperature

In 1971, the last of today's accepted seven basic units was Mol (mol), the SI unit of amount of substance from which all other needed units can be derived.

With increasing use and relevance of electronic communication, the *Unified Code for Units of Measure* (UCUM) has gained fundamental importance. UCUM is "a code system intended to include *all* units of measures being contemporarily used in international science, engineering, and business... to facilitate unambiguous electronic communication of quantities together with their units. The focus is on electronic communication, as opposed to communication between humans." (Schadow and McDonald 2014, http://unitsofmeasure.org/ucum.html). The failure of the NASA Mars Climate Orbiter, which was destroyed on a mission to the Mars in September 1999 instead of entering the planet's orbit, was due to miscommunications about the value of forces (ftp://ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf.).

Social, educational, and clinical sciences lack standard measures and units. Continuous efforts are made by various working groups to develop frameworks and models to link physiological, pharmacological, and biological units and systems with behavior to understand basic dimensions of functioning. A prominent example is the Research Domain Criteria (RDoC) which is "centered around dimensional psychological constructs (or concepts) that are relevant to human behavior and mental disorders, as measured using multiple methodologies and as studied within the essential contexts of developmental trajectories and environmental influences." (National Institute of Mental Health, NIH, https://www.nimh.nih.gov/research-priorities/rdoc/index.shtml). This approach, like similar others, could be addressed in conceptual frameworks for drug development programs, not limited to neuropsychopharmacology. Ideally, in drug research and development, the measurement concept should be developed right from the beginning of the search for a new substance or method when first consideration on medical indications is in play or in case of planned modification of a known drug (FDA Clinical Outcome Assessment Qualification Program 2017). The early decision on clinical measures may help to link pharmacological evidence to the assessment of patient's benefit after approval by authorized bodies.

The measurement of hypothesized effect as well as the confirmation of known effects of a substance or compound needs a theory-based quantification strategy to make the hypothesis testable and enable new hypotheses. There are two major strategic mistakes that should be avoided from the perspective of methodology: usage of measurement tools of unknown validity or reliability in explorative or even pivotal clinical trials and usage of well-known measurement instruments in the wrong experimental setting or population.

Clinical measurement and scaling methods are used to describe, qualify, and quantify (localize) objects, constructs, or behavior during the clinical studies and document the treatment outcomes according to the study protocol. Some scales may either be used as diagnostic tools for the classification or severity grading of patients, and as criteria for selection of patients, or as outcome variables. In justified cases, scales may be used for both diagnostic purposes and outcome variables. In case of diagnostic tools, it is important that the scale has been shown to be of sufficient content and construct validity to cover essential aspects of the target disease, condition, symptom, or syndrome. Of special interest in binary decision are the sensitivity and specificity of the diagnostic instrument to support classification (e.g., "patient shows signs of disease X": yes/no), with "sensitivity" defined as the detection rate of correct positive and "specificity" defined as the detection rate of correct negative identified persons, e.g., not having symptoms of disease X. These criteria are prerequisites for the correct selection of valid samples from the target population as defined in the clinical drug development plan.

The observer-rated scales on the other side are helpful to collect additional data on treatment effects outside the visits with the doctor or investigator. They should always be introduced with an explicit recommendations on who (expert, physician, medicinal personal, etc.) and at which time during the study they have to be used. It is mandatory to define the accountable respondent in the study protocol to make sure that the appointed rater(s) will be trained for the correct use of the scale throughout the study.

Diagnosis for sure is a complex cognitive process which frequently, especially in very experienced experts, is emerging from implicit processing of information. Various studies have shown that medical experts are better in differential diagnostic, tend to process less information, and decide faster than nonexperts (e.g., Kundel et al. 2007). Acknowledged experts in the field of interest should not only be involved in the selection and development of clinical outcome measures but also take part in the rater training sessions. In addition to the training mostly focused on the proper use of study documentation, rater training may increase the awareness for the necessity to make explicit reasoning and documentation with regard to the diagnostic process and conclusion.

A further class of clinical scales, the self-reports often are of primary interest, apart from tolerance ratings, for the assessment of the

efficacy of the treatment. Self-reports are the only method to measure and qualify symptoms, thoughts, beliefs, and opinions. Performance tests adopted to clinical conditions are important to complete the clinical outcome assessment. The use of performance tests measuring learning and memory functions, attention, and problem-solving, for example, is in no way restricted to neuropsychopharmacology. These tests have shown to contribute to the understanding of CNS effects in various other indications (e.g., hepatic encephalopathy).

In one of their draft guidelines (2006) on patient-reported outcome (PRO) measures, the FDA stated that these measurement tools are of increasing importance in drug development. Self-reported questionnaires that are given directly to patients without the intervention of clinicians are often preferable to the clinician-administered interview and rating of the given answer. Questionnaires which are self-completed capture directly the patient's perceived response to treatment, without a third party's interpretation, and may be more reliable than observer-reported measures because they are not affected by interobserver variability (FDA 2006).

There are various topics which may be directly captured with the patient's response to a PRO. The choice of the item content, design, and response format will very much depend on the targets of the clinical program. A concept of PRO measures may be, for example, one of the following:

- Discrete symptoms or signs, for example, frequency and/or intensity of pain sensation, and frequency of seizures
- Overall conditions, e.g., depressiveness, fatigue, and general physical complaints
- Feelings about the health condition, e.g., worry about disease getting worse
- Feelings and opinions about the treatment, e.g., feeling the treatment is effective, and feeling relief of depressive mood since the start of treatment
- General assessments, for example, improvement in physical functioning, treatment satisfaction, overall quality of life ratings, and health-related quality of life (HRQL) ratings
- Specific assessments, e.g., decreased pain intensity and how bothersome specific symptoms are
- Assessments of change after treatment (e.g., "symptom improved very much since the start of treatment"; "symptom improved most of the time after start of treatment," etc.) in contrast to absolute assessments

Clinicians have recognized that understanding the patient's perspective on the impact of disease and treatment on functioning and well-being is important for pharmaceutical, biologic, and medical device product development and evaluation. Clinical studies are increasingly incorporating health-related quality of life (HRQL) and other concepts of patient-reported outcome instruments into clinical trial programs for new drugs with the expectation that these outcomes will help inform physicians and patients on the beneficial effects of these treatments (e.g., Wilke et al. 2004). PROs may be further useful in differentiating the patient benefits among competing products with similar clinical efficacy and translating clinical effects into outcomes more meaningful to patients, relatives, and treating physicians. For a first classification of the various clinical scales, see Table 1.

Technically speaking, a diagnostic scale may be constructed as a questionnaire or a checklist asking for a step-by-step response from one item to the next or it may be used as a guidance for the experienced physician to conduct a patient interview (e.g., Hamilton Depression Scale). Diagnostic scales should reflect relevant diagnostic criteria in ICD-10 (http://www.who.int/classifications/icd/en/), other diagnostic classification systems, and/or medical or regulatory classification systems, e.g., DSM-5 for mental disorders. If available the scales selected for the program should also correlate with criteria from validated structured or semi-structured interviews established in the relevant medical speciality. The ongoing discussions about the scientific usefulness of diagnostic systems such as DSM or ICD-10 (e.g., Kendell and Jablensky 2003) may be taken into account. Both systems may not provide distinct classification in some disciplines, for instance, in psychiatry, that cannot always be transcribed into neurobiological pathways and genetic entities (cf. NIH Research Domain Criteria).

Table 1 Taxonomy of the role of observer- and patient-rated scales in clinical development of drugs

Scope	Source of information	Content	Processing
Diagnosis; selection; exclusion	Physician/expert	Overall health status Symptoms, signs, or syndrome Functional status	Classification Identification Ranking
	Patient	Overall health status Symptoms/signs, individually or as a syndrome associated with a medical condition Functional status Activities of daily living Perception/worries about health Health-related quality of life	
	Relative/ caregiver	Functional status in daily situations Efforts in caring for the patient/carer burden	
Measurement of treatment outcome	Physician/expert	Clinical global impression of health status, severity of illness, change in condition, therapeutic effect, side effects, etc. (e.g., AMDP&CIPS 1990) Overall health status Health-related quality of life status and/or change Symptoms, signs, or syndrome Functional status	Effect, efficacy Benefit Tolerability, adverse events, estimation of Risk factors Covariate to primary effect measure
	Patient	Absolute and/or relative measure of change: Overall health status Symptoms, signs, syndrome Functional status Activities of daily living Perception/worries about health Patient satisfaction with the treatment and its results (e.g., Asadi-Lari et al. 2004) Health-related quality of life	
	Relative/ caregiver	Functional status in daily situations Carer burden	
Monitoring	Professional personal/ caregiver/relative	Documentation of study outside events	Nuisance variables, system description

As already emphasized physicians, experts, and other professional personal responsible for assessments should be individually trained in the use of the diagnostic and observer-reported scales. Rater training should be an integrated part of any clinical explorative and pivotal trial to further valid diagnosis and reliable selection of study populations and generate informative co-variates for posteriori statistical adjustments and explorations in accordance to the study protocol.

Procedure

How to select, develop, and/or modify clinical scales eligible for the planned clinical program or a clinical study? Which rater (expert rater, observer, patient) will be informative on which level and in which domain of treatment effects? How to operationalize the research concept and to make hypotheses testable and find informative answers from explorative tests and questions?

The search for suitable clinical scales can be started with scanning guidelines and consensus publication from the medical speciality of interest. There are also critical reviews on scales available (e.g., AMDP&CIPS 1990). Usually, the first step to find appropriate scales will be to collect and compile items from various published scales (tasks, questions, or expressions); compare those items with regard to content, wordings, and response format; check for doubles; and decide

on a first selection of items to be tested as draft item collection with experts in the field and ideally with people who are familiar with symptoms and signs. The following criteria may be important for the selected items and scales check:

Wording
Comprehensibility
Completeness of topics
Response format
Availability of statistical item/scale characteristics.

These criteria are as well applicable to qualify standardized published scales. In case of foreseeable multicenter-multicultural studies, the planning team should consider translation of the eligible items/scales. This may, and mostly will, make linguistic validation and cultural transformations of concepts and cognitive debriefing necessary (Wild et al. 2009). These procedures are often very time consuming and should be planned right in time.

After compilation and adoption of the items according to the constructs, dimensions, and domains pre-specified in the clinical development plan, the scale(s) will be applied in a test sample of subjects (best choice would be a subsample from the target population of the developmental drug) in order to analyze item and scale characteristics and confirm or develop a scaling and scoring procedure ready for purpose. As discussed before, scaling is the part of measurement that involves the construction of an instrument that associates qualitative constructs with quantitative metric units. Scaling evolved out of efforts mostly in psychologic and educational research to measure constructs like intelligence, motivation, authoritarianism, self-esteem, and so on. In many ways, scaling remains a mostly misunderstood aspect of social research measurement. Remarkably it attempts to do one of the most difficult of research tasks, to measure abstract concepts, which have no obvious empirical conditions or which are not directly observable.

The *item* is regarded the unit of any scale (questionnaire, behavioral test). It is usually composed of three parts: (1) the test instruction (information on what the item or the complete scale is used for and how the choice for best response should be performed); (2) the presentation of a question, task, or an expression to which the respondent is invited to respond to; and (3) the choice of best response which is expected in the third part of the item offering an arrangement of words, numbers, idioms, symbols, or any other signal (response option of a specific format) to make best discriminative choice likely. The response format may be binary, multi-categorical, ordered categorical, or a response line with one or more anchors (visual analog scale, VAS). In some behavioral tests, the response is expected in a blank field (e.g., a gap at a certain position in a sentence where the respondent has to complete the sentence). Free form tests and questionnaires may collect interesting information. The free answers have to be separately analyzed and categorized to enable further statistical processing. In closed formats, the stimulus for choice of the best response should be well designed to induce a discrimination process reflecting the degree of agreement or disagreement.

The rating of the expression or question may in relation to the question be

Direct (e.g., are you satisfied with the results of this treatment: yes/no or another format with more than two categories)
Comparative (e.g., my headache is much better than before treatment: true/not true)
Magnitude (any frequency or intensity rating, e.g., for pain)

According to Stevens (1951), measurements in science are generally on four levels (see Table 2). The taxonomy may be extended by a fifth measurement level, the absolute scale (e.g., number of objects, probabilities).

Any numeric ordering or intuitive arrangement of verbal or symbolic response options does per se not reflect the true level of the measurement. In fact it has to be empirically shown that ratings, for example, on the frequency dimension with a typical arrangement like 0 = "never," 1 = "occasionally," 2 = "frequently," and 3 = "always" fulfil the criteria of ordinal or even interval level of measurement. Especially questionnaires in behavioral and social sciences rarely reach the level of

Table 2 Levels of measurement and statistical processing

Level	Statistical description	Example	Test statistics
Nominal (or categorical)	Modus, frequency	Gender	Nonparametric tests
Ordinal	Median, percentile	Degree of agreement	
Interval	Arithmetic mean, standard deviation	Body temperature in Celsius or Fahrenheit	
Ratio	Geometric mean	Age, body weight	Parametric tests

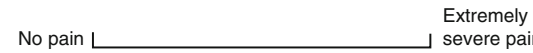
interval scale. Psychometricians are interested in scaling on a high level of measurement. Despite these considerations, the choice of the response format should always be guided by the content and scalability of the construct. Therefore, just attaching ordered natural numbers or percent numbers to design a "metric" response format is not scaling in a mathematical sense. The qualification of level of measurement of a scale has to be proven on response distribution characteristics based on empirical data.

Scales may further be divided into two broad categories: unidimensional and multidimensional. The unidimensional scaling methods were developed in the first half of the twentieth century, and some of them have been named after their inventor. Among the various scaling methods, the *psychophysical scaling* has a separate theoretical background. Psychophysics is a psychological discipline that has its roots back in the work of G. T. Fechner, E. H. Weber, and Wilhelm Wundt, founder of the first laboratory for experimental psychological at the University of Leipzig, 1879. Psychophysics deals with the relationship between physical stimuli and subjective correlates, in general the percept. Psychophysicists employ experimental stimuli that can be objectively measured, such as pure tones varying in intensity or lights varying in luminance or frequency. All the traditional senses have been studied including the enteric perception and the sense of time. Regardless of the sensory domain, three main procedures of investigation have been used: the definition of absolute threshold, discrimination threshold, and various scaling procedures using constant or systematically varied stimuli characteristics. The absolute threshold is the level of intensity or frequency at which the subject can just detect the presence of the signal. The difference threshold is defined as the magnitude of difference between two stimuli of differing intensity or frequency that the subject is able to detect. The just noticeable difference, also named difference limen (DL), is the difference in stimuli properties the subject notices with a defined proportion of the cases (mostly $p = 0.50$).

The determination of critical flicker fusion frequency and that of critical fusion frequency are examples of psychophysical measurement that have often been used and are still in use in psychopharmacology.

The *visual analog scale* (VAS) is in most cases a 100 mm horizontal line with two named poles or verbal anchors, like "not at all" and "very much" or similar wordings. The typical use of VAS is, for example, the following:

How severe was your pain today? Please place a mark on the line below to indicate how severe your pain was!



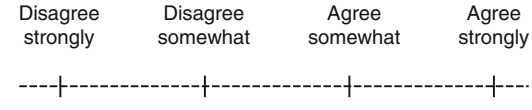
When using VAS as a measurement instrument, one tries to quantify a sensation, a trait, or any other entity on a ratio scale level assuming further that the entity's characteristics are ranging across a continuum from "none" to "very severe" or "very intensive" or within a similar concept. The assessment is highly subjective, in a practical way "imprecise," in regard to the positioning of the tic mark. VAS may be of value when looking at the intraindividual change of the entity, but they are most likely of less value for comparing groups. It has been argued that a VAS is trying to deliver interval or even ratio measures. But there are no convincing arguments for the values being more than ordinal data. Despite the simplicity and "face validity" of this measurement

method, data handling and interpretation have to be done with caution.

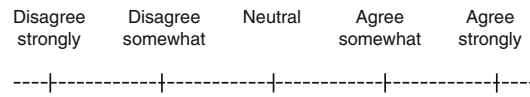
Thurstone scaling. Thurstone was one of the first scaling theorists. He invented three different methods for developing a unidimensional scale (e. g., 1927): the method of equal-appearing intervals, the method of successive intervals, and the method of paired comparisons. The three methods differed in how the scale values for items are constructed, but in all three cases, the resulting scale is rated the same way by respondents. The method of equal-appearing intervals is explained as it is the easiest method. Because this is a unidimensional scaling method, the concept one is trying to scale is reasonably thought of as one-dimensional. When starting the procedure, the description of this concept should be as clear as possible so that the persons who are going to create the statements (items) have a clear idea of what the investigator is trying to measure. Next, the developer will ask people to generate similarly worded statements about the concept. Then the participants are asked to rate each statement on an 11-point response scale with a predefined criterion, like how favorable the statement appears to them with regard to the construct. Next the ratings will be analyzed. For each statement, one needs to compute the median and the interquartile range. The median is the value above and below which 50% of the ratings fall. The first quartile (Q1) is the value below which 25% of the cases. The median is the 50th percentile. The third quartile, Q3, is the 75th percentile. The interquartile range is the difference between third and first quartile, or Q3–Q1. To facilitate the final selection of items for the scale, one might write the parameters into a table; maybe we want to sort the statements in the table of medians and interquartile range in ascending order by median and, within that, in descending order by interquartile range.

For the final scale, one should select statements that are equally distributed across the range of medians. Within each median value, try to select the statement that has the smallest interquartile range. Once the items for the final scale are selected, one should test the new scale in an independent sample of eligible patients or members of the target population.

Likert scaling is a unidimensional scaling similar to Thurstone scaling. The term "Likert scale" is used in various ways in literature. Sometimes the term actually seems to describe Likert or Likert-like items. Likert scales are the four- to nine-point scales much used in clinical trials and in many other fields of research. The scale is often used in a semantic polar format, that is, in relation to the statement or question, the response options correspond to wordings like "agree strongly" and "disagree strongly." Example:



The five-point scale is probably the most commonly used response format version, which in many cases in contrast to the four-point scale will include a midpoint "neutral."



The assumption with the VAS as well as behind numeric rating scales, including Likert scale, is that the geometrical distance between markers on a line or of tic boxes in combination with verbal expressions and/or numbers is homologous. This is one of the reasons why the graphical layout of those items should guarantee equally spaced elements of the response format.

After definition of the concept, verbalization of the statements, tests of the draft version, and confirming those items that form a reasonable scale, the final score for the respondent on the multi-item scale is the (weighted) sum of their ratings on all items (sometimes called a "summated" scale). On some scales, one will have items that are reversed in meaning from the overall direction of the scale. To cumulate item scores, one will have to inverse the response score of this item to maintain unidirectional scaling. Likert scales, like other item scales, may be problematic in comparison across groups. The expectation of a researcher would normally be that the mean response will vary across treatment groups. The problem is that in many cases the variances will

also differ. The variance has to be less at the ends of the scale, as there is no alternative response to one side of the endpoint. For example, with a five-point scale, the variance would be expected to be largest at the midpoint, 3, and smallest at the extremes. A possible solution to this problem might be to use the arc sine square root transformation of the scores. The responses are divided by 5, to yield a number between 0 and 1. The square root is taken (still between 0 and 1). The angle whose trigonometric *sine* is that number is the transformed response and can be used for further statistical analysis.

Osgood et al. (1957)'s *semantic differential* was designed to measure the connotative meaning of concepts. The respondent is asked to choose his or her position on a scale between two bipolar adjectives (e.g., "adequate–inadequate" or "valuable–worthless").

Sometimes it may be difficult to find properly defined poles of the differential! Therefore, many researchers prefer unipolar item scales (e.g., mood scales and multi-item pain scales).

Guttmann scaling. This method is also known as cumulative scaling. Like with the other examples of item response scaling, this method starts with the definition of the construct of interest, the generation of a large set of statements that are judged by some experts or members of the target group how favorable (yes/no rating) the expressions are in regard to the construct. Following this, one constructs a matrix or table that shows the responses of all the respondents on each of the items. Afterward this matrix is sorted so that respondents who agree with more statements are listed at the top and those agreeing with fewer are at the bottom of the matrix. For respondents with the same number of agreements, the statements are sorted from left to right from those that most agreed to those that fewest agreed to. In case of only a few items, one can easily examine this matrix. In larger item sets, the method of choice may be the *scalogram analysis* to determine the subsets of items from our pool that best approximate the cumulative property. After the review of these items follows the selection of the final scale elements. In many cases, there is no perfect cumulative scale and the researcher will have to test for goodness of fit. These statistics will estimate a scale score for each of the items that are used in the calculation of a respondent's score.

In the late 1950s and early 1960s, measurement theorists developed more advanced techniques for creating multidimensional scales. *Multidimensional scaling* (*MDS*) is a data reduction technology normally using a direct similarity or dissimilarity matrix. MDS fits a set of points in a space such that the distances between the points are as closely as possible to a given set of dissimilarities between a set of objects, for example, ratings. MDS does not make distribution assumptions necessary. As MDS is a spatial method, there are metric assumptions, for example, the distance from some point A to B shall be the same as from B to A. This might sound strange to the reader, but in some situations, two points A and B may not be bidirectionally equidistant. Consider, for instance, the distance between home and work, which may be due to specific situations in the morning and the evening not of identical length. If the equidistance assumption cannot be fulfilled, one should not use MDS. Anyway, the decision whether to use or construct one- or multi-dimensional scales depends very much on the clinical target (will efficacy in one dimension clinically be sufficient or will relevant drug effect be expected in more than one dimension) as defined in the clinical development plan. If the construct is one dimensional, one also will use one-dimensional scales, and if the construct is of known multidimensionality, one should consider multidimensional scales or several one-dimensional scales. Both ways will offer their special advantages and disadvantages with regard to the upcoming point of decision on the clinical trial outcome.

The application of mathematical models to response data from questionnaires, clinical, educational, and psychological tests, is discussed and described in test theory. Test theory is a body of theory that offers mathematical models to make statistical adjustments in response data in order to predict, describe, or estimate a person's trait, ability, attitude, or any other construct. There are in general two different test theories, which are of relevance in the present context and which help to

understand the steps from single item or item pool generation to item construction, definition of adequate response options to testing the first draft questionnaire, and confirmation of the final measurement instrument. The methods are also very helpful for the reevaluation of known scales and items. The two main test theories are:

1. *Classical test theory*, which assumes that for each person, we have a true score of some ability or characteristic, T, which would be obtained if there were no errors in our measurement. Because instruments used for measurement (and sometimes the users of those too) are imperfect, the score that is observed for each entity, for example, a person's ability, most times is different from the person's true abilities or attitudes. It is concluded that the difference between the true score and the observed score is the result of measurement errors. Classical test theory is dealing with the relation of the true score T, the error E, and the observed score X. Formally:

$$X = T + E.$$

Further assumption: True score T and error E are not correlated, $\rho(T,E) = 0$.

The most important concept is that of reliability. The reliability of the observed test scores X, denoted as $\rho^2(X, T)$, is defined as the ratio of true score variance $\sigma^2 T$ to the observed score variance $\sigma^2 X$. Because it can be shown the variance of the observed scores to equal the sum of the variance of true scores and the variance of error scores, it follows that

$$\rho^2(X,T) = \frac{\sigma^2 T}{\sigma^2 X} = \frac{\sigma^2 T}{\sigma^2 T + \sigma^2 E}.$$

The reliability of test scores becomes higher as the proportion of error variance in the test scores becomes lower and vice versa. The reliability is equal to the proportion of the variance in the test scores that could be explained if we knew the true scores. The square root of the reliability is the correlation between true and observed scores.

2. *Item response theory* (*IRT*), also known as latent trait theory, is a set of probabilistic models and the application of mathematical models to response data from behavioral tests and questionnaires measuring abilities, characteristics, or other variables. IRT models apply functions to quantify the probability of a discrete outcome, such as a correct response to an item, in terms of *person and item parameters* (see, e.g., Linden and Hambleton 1997; Rost 2004).

   Person parameters may, for example, represent the cognitive ability of a patient or the severity of a patient's symptom. Item parameters may include item difficulty (location), discrimination (slope), and random guessing (lower asymptote). IRT does not only apply to discrete binary data but may also deal with ordered categorical data to indicate level of agreement and other response dimensions as already discussed earlier. One of the purposes of IRT is to provide a framework for evaluating how well assessments and individual questions on assessments work. In drug development programs, IRT may be very helpful to collect and construct items and maintain item pools for clinical trials in a defined indication and develop or adopt new scales within the conceptual framework of the clinical program.

The performance of an item in a test is described by the *item characteristic curve* (ICC). The curve gives the probability that a person with a given ability level will answer the item correctly or give an answer in line with the expectations according to the construct definition. Persons with lower ability ($\theta < 0.0$) have less of a chance to answer correctly or agree on a yes/no item, while persons with high ability are very likely to answer correctly.

IRT models can be divided into two families: one-dimensional and multidimensional models. One-dimensional models require a single trait (e.g., ability) dimension θ. Multidimensional IRT models analyze response data arising from multiple traits. However, because of the greatly increased complexity with increasing number of included traits, the majority of IRT research and applications utilize a one-dimensional model. The models are further on named according to the number of

parameters estimated. The one-parameter logistic model (1PL) assumes that there is only minimal guessing by the respondent and that items have equivalent discriminations, so that items can be described by a single parameter (b_i). The 1PL uses only b_i, the 2PL uses b_i and the parameter a_i, and the 3PL uses b_i, a_i, and item parameter c_i.

A given model describes the probability of a correct response (or a yes/no response option where one is defined as correct and the other as incorrect in the frame of some syndrome or disease theory) to the item as a function of a *person parameter*, which is in the case of multidimensional item response theory, a vector of person parameters. For simplicity we will stay with the model of only one person parameter. The probability of a correct response depends on one or more item parameters for the item response function (IRF). For example, in the three-parameter logistic (3PL) model, the probability of a correct response to an item i is given by

$$p_i(\theta) = c_i + \frac{(1 - c_i)}{1 + e^{-a_i(\theta - b_i)}}$$

where θ signifies the person parameter, e is the constant 2.718, and a_i, b_i, and c_i are the item parameters.

As can be seen from Fig. 1, the item parameters simply determine the shape of the IRF. The figure depicts an example of the 3PL model of the ICC with an explanation of the parameters. The parameter b_i represents the item location (item difficulty). It is over the point on θ where the IRF has its maximum slope. The simulated example item is of medium to higher difficulty, since $b_i = 1.3$, which is to the right of the center of the distribution. The item parameter *a i* represents the (rather good) discrimination, the degree to which the item discriminates between persons in different regions on the latent continuum. This item parameter characterizes the slope of the IRF where the slope is at its maximum. The item parameter $c_i = 0.20$ indicating that persons with low ability may endorse correct response.

One of the major contributions of item response theory is the extension of the concept of reliability. Traditionally, reliability refers to the precision of measurement (i.e., the degree to which measurement is free of error). And traditionally, it is measured using a single index, such as the ratio of true and observed score variance (see above). This index is helpful in characterizing an average reliability. But IRT makes it clear that precision is not uniform across the entire range of test scores. Scores at the edges of the test score

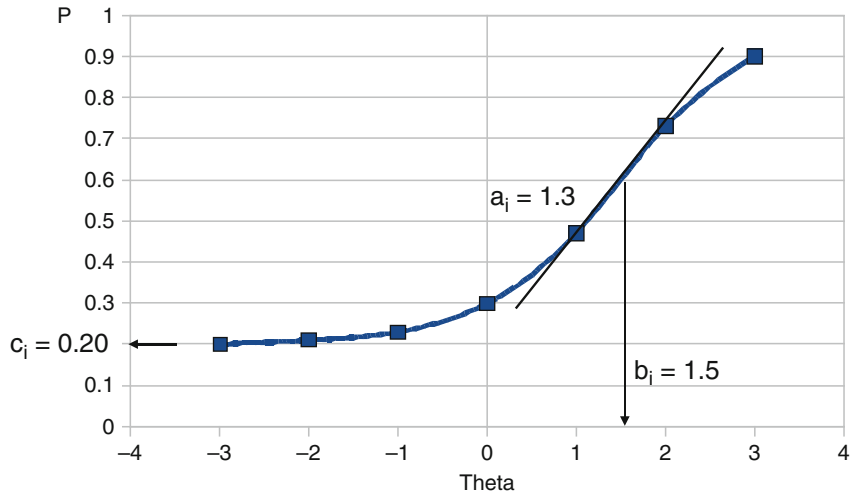


Fig. 1 IRF for a hypothetic item data with item parameters $b_i = 1.5$, $a_i = 1.3$, and $c_i = 0.2$

range generally have more error associated with them than scores closer to the middle of the range.

Item response theory elaborated the concept of item and test information to replace reliability. Information is also a function of the model parameters. According to Fisher information theory (named after the inventor and famous statistician R.A. Fisher), the item information supplied in the case of the Rasch model (Rasch 1960) for dichotomous response data is simply the probability of a correct response multiplied by the probability of an incorrect response:

$$I(\theta) = p_i(\theta)q_i(\theta).$$

The standard error (*SE*) is the reciprocal of the test information at a given trait level:

$$\mathrm{SE}(\theta) = \frac{1}{\sqrt{I(\theta)}}.$$

Intuitively, one can agree to the conclusion that more information implies less error of measurement.

After this short excursion into some basics of test theory, one may agree that measuring is not just assigning numbers to empirical objects or events (see the straightforward definition of S.S. Stevens 1951). In the classical definition, measurement is the estimation of ratios of quantities. Quantity and measurement are mutually defined: quantitative attributes are those which make measuring possible. In terms of representational theory, numbers are assigned based on similarities between the structure of number systems and the structure of qualitative systems. A property is quantitative if such structural similarities can be established. This definition is much stronger than the definition of Stevens (1951).

Evaluation

Drug development in known or new indications will start with an extensive literature search for measurement models and scales in the target area. When searching and selecting suitable scales for the clinical development plan, essential information about the scale characteristics and properties are of major importance. Information with regard to the following topics may be needed:

Completeness and representativeness with regard to the concept of interest
Relation to medical and mathematical measurement models available
Published indices or data for at least scale reliability and validity
Sufficient evidence on satisfying scale properties
Evidence for validated linguistic and/or cultural versions, as in many cases pharmacological drug development will increasingly often be performed in multicultural and multilingual studies

In addition instructions for the standardized application of the test or questionnaire, procedures of training raters for the proper use of those instruments are as important as the mostly cited scale properties reliability and validity. Further information about the way the items shall be presented, the scoring rule, for which experimental conditions and in which population the indices and coefficients are valid is needed. It is almost always necessary to consider some kind of reevaluation of the selected scale(s) for the own new project. Clinical scales used as outcome assessment tools are key elements in patient-focused drug development. Their reasonable use is recommended not only by the authorities in the USA and Europe.

The necessity of a conceptual framework before starting the clinical development program of a drug is known for decades in the field of pharmacodynamic research and development. It has been explicitly named in the guideline for patient-reported outcome measures (FDA 2009). The conceptual framework should combine three major concepts:

1. Treatment or interventional concept
2. Target population for the intended treatment
3. Measurement concept (including the endpoint model)

Figure 2 gives an overview of some important steps in the developmental process from the

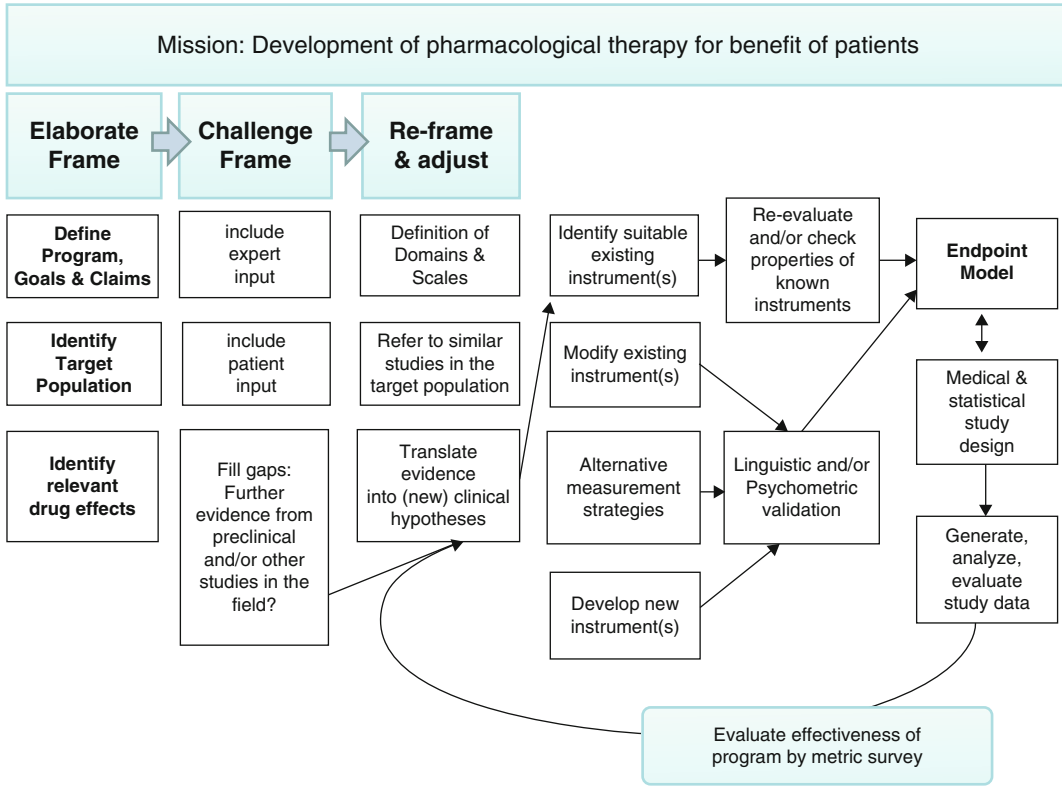


Fig. 2 Critical path description from concept development to endpoint model. The *arrows* shall indicate the flow of information during the process

measurement and scaling concept until the final definition of the endpoint model and the entrance of these findings and decisions into a clinical study protocol. These steps can be planned and performed in a way of inductive and deductive reasoning including strategic decision nodes. Suffice to say scales can be used to document and evaluate these processes in innovative drug development. Theories of judgment and decision making are interesting fields, but their discussion is beyond the scope of this chapter. The interested reader may find interesting topics in Kahneman (2011).

Critical Assessment of the Method

The selection of scales for a drug development program should always be guided by the program goals and claims that are intended to reach for. Sometimes the decision to select and use known scales seems to be influenced by their availability, their application in similar clinical programs or clinical studies, and published data. In case of diagnostic scales, this strategy may be appropriate as long as the drug development program is within a well-known traditional indication and patient population. On good reason to rely on established scales is the comparability of new results with former study outcomes. Innovative drug development, e.g., based on a new mode of action, or intended for special populations (e.g., children, or very old persons), will, in many cases, make modification and/or development of diagnostic scales, scales for inclusion criteria assessments, and above all outcome measures necessary. Some of the following criteria in case of search and selection of known scales may be helpful for the decision. For what purpose and according to which clinical and methodological framework

was the scale developed did scale development care for sufficient content validity? Does the item concept (question, task, or statement, and the response format) match the intended construct to be investigated? Is the scale a single-item or multi-item scale and does it appropriately fit the one-dimensional or a multidimensional construct? In the preceding paragraphs, some information about item and scale development methods have been discussed which might be of additional help to answer these questions during search, selection, evaluation, and decision on eligible scales for the planned clinical program.

Critical *item analysis and construction* will include semantic and mathematical methods as well as consideration on the design of item and/or scale. Further the planning team may consider more modern technical solution to present items and scales (e.g., tablets, smartphones, WEB-based surveillance platforms) which could offer new perspectives on drug assessment. These new technologies offer opportunities to "continuously" gather data (e.g., dairy data). Privacy and data protection, validation of the system, testing the validity of items and scales, and planning adequate data handling and analysis will be a challenge. The computerized presentation of clinical scales will need some attention on wording, layout, and arrangement of response options to direct the respondent's attention to the task or question and to encourage to an open unbiased response. There is a considerable number of literature which reports on research issues and findings of psychologists and survey methodologists on the cognitive, motivational, and communicative aspects of presentation and comprehension of items and scales (for overview, e.g., Schwarz 1999). Most items are composed of three parts, the test instruction (and presumably further information on the context of the surveillance), the question/task itself, and the response part which may be designed in open or closed format. Open formats would allow the respondent to give free answers, but in clinical research (except for questions on tolerability and adverse events), probably the most preferred approach, especially when asking for efficacy of treatment (as primary outcome measure), is the closed format. Items researchers will have to consider the specific cognitive, motivational, and communicative processes underlying task or question and the dimension and response format under clinical conditions. Unlike measurement in natural sciences, behavioral measurement is based on complex cognitive and emotional processing of the information given to the respondent. Patients, and also doctors, tend to draw information from the content and design of an item and may determine more or less explicitly what the best answer or solution would be. Experimental conditions and scales provide some information to the respondent which can result in unexpected outcomes. Therefore, the wording of the statements or expressions should always be comprehensible and readable to the intended respondent. In case of diagnostic scales, the items must be coherent to relevant criteria of the disease or the syndrome. Patient-reported scales should avoid as much as possible technical or medical terms and replace them by more daily speech terms. The input of patients from the target population is mandatory to find appropriate speech. In new fields structured interviews may be more informative than unstructured interviews. With regard to the response options and the response scaling, the response scale should reflect both temporal properties of the disease or symptom(s) as well as the temporal aspects of the hypothesized time and mode of action of the drug, all issues be adequately reflected in the study design and protocol. Typical response options used in clinical trials are on the intensity or frequency dimension with ratings related to time windows like "yesterday," "this morning," "the last week," "the last month," "since the last visit," and others. The choice of temporal anchoring of assessments must be in line with the nature of the symptom or the disease and must consider reasonable memorization of the construct, symptom, or process. Answers about the past are less reliable when they are very specific, because the questions may exceed the subjects' capacity to remember accurately. There is much literature about the functioning of autobiographic memory (e.g., Williams et al. 2008). Human judgment is context dependent by drawing on the information that is most accessible at the time point of measurement. In certain cases the respondent may tend to randomly guess or feel encourage to respond in direction of a presumed desired direction. As a

practical example, items asking for comparative judgment, for example, "How is your pain today," "no pain," "pain as usual," or "pain more severe than usual," are obsolete because of the undefined anchor "usual." When drafting items, the researcher must be focused on the wording (1) to match theoretical criteria or research object of interest and (2) be aware of item characteristics interpreted as relevant information by the respondent. Comprehension, interpretation, the type of memorizing the topic, and communicative information of the item (and the scale in total) are examples of cognitive processing of questions, statements, and tasks and deserve special attention.

Implementation of a scale into the clinical development plan and single studies. There is a common misconception that if a published validated instrument is chosen for a clinical study or evaluation program, one does not need to check the reliability and validity of the instrument in the target population. If content and response format do not fit into the study design, the planning team should consider modifications. The modified instrument has to be tested again for its scale properties in a separate sample taken from the target population before using it in explorative or pivotal clinical trials. If the investigated symptom has temporal properties in the way that its appearance or intensity is changing during daytime, or its appearance or intensity may vary from day to day, the investigator might consider the implementation of a patient's diary in the study design. FDA supports the use of diaries where appropriate, but "If a patient diary or some other form of unsupervised data entry is used, the FDA plans to review the protocol to determine what measures are taken to ensure that patients make entries according to the study design and not, for example, just before a clinic visit when their reports will be collected" (Federal Register Vol. 71, Nr. 23, pp. 10; 334–337). In recent years, systems for use of electronic questionnaires, PRO, and diaries frequently offered by specialized companies are used by sponsors to facilitate data selection. The eventual advantages of the administration of well-controlled electronic questionnaires may be among others (e.g., Kelly 2015)

- Help to counteract declining response rates
- Reduce collection costs
- Improved data quality
- Rapid and continuous access to data
- More detailed insight into subject's behavior regarding input of data

Validation, reliability, and compliance with Part 11 electronic data requirements (http://www.fda.gov/ora/compliance_ref/Part11/) are important issues that will be raised by the FDA and other authorities, and the sponsor will have to successfully address them. A typical question is often heard: Are paper questionnaires and electronic questionnaires equivalent? The answer most likely is no, because any modification in a scale (including method of administration) will certainly need some additional validity evidence.
Scale properties. A questionnaire or test as an instrument to measure entities or a construct should have specific properties in order to be accepted as a scale. The instrument should be standardized in the way and the circumstances it shall be presented and used, the evaluation of the ratings shall be defined and objective in a way that the results of this evaluation is independent of the evaluating person, and there should be important data available on its reliability and validity. In psychology, the concept of validity has in addition to the already discussed aspects two further fields of application. The first aspect addresses test validity, a concept that has evolved in psychometrics, dealing with theory and technique of psychological and educational measurement. The second is related to the study design, pointing to the fact that different types of studies are subject to different types of bias. For example, recall bias is likely to occur in cross-sectional or case-control studies where subjects are asked to recall exposure to life events or other special events. Subjects with the relevant condition (e.g., the degree of disease or syndrome to be investigated) may be more likely to recall relevant events that they had experienced than subjects who do not have the condition or have a lower degree of it.

In contrast to test validity, assessment of the validity of a research design does not involve data collection or statistical analysis but rather evaluation of the design in relation to the desired conclusion on the basis of prevailing standards and

theory of research design. This obviously is an issue in drug development programs, especially when designing several phase III studies in various regions and cultures.

Test validity, which is in focus here, can be assessed in a number of ways (APA 2014). Test validation typically involves more than one type of evidence in support of the validity of a measurement method (e.g., structured interview, questionnaire, test, etc.). The various types of validity include content-related, construct-related, and criterion-related evidence with the subtypes concurrent and predictive validity according to the timing of the data collection. In the following we will present and discuss some of the various aspects.

Construct validity evidence involves the empirical and theoretical support for the interpretation of the construct. A good construct validity has a theoretical basis which is translated through clear operational definitions involving measurable indicators. Construct validity evidence includes statistical analyses of the internal structure of the test including the relationships between responses to different test items. They also include relationships between the test and measures of other constructs. Researchers should establish both of the two main types of construct validity, *convergent* and *discriminant*, for their constructs.

Convergent validity is assessed by the correlation among items that make up the scale or instrument measuring a construct (internal consistency validity); by the correlation of the given scale with measures of the same construct using scales and instruments proposed by other researchers, if appropriate, with already accepted in the field (criterion validity); and by correlation of relationships involving the given scale across samples.

Internal consistency is one type of convergent validity that seeks to assure there is at least moderate correlation among the indicators for a concept. Cronbach Coefficient Alpha is commonly used to establish internal consistency (as well as an aspect of reliability and for evidence of construct validity) with at least Alpha of 0.60 considered acceptable for exploratory purposes, Alpha of 0.70 considered adequate for confirmatory purposes, and Alpha of 0.80 considered good for confirmatory purposes.

Simple factor structure is another test of internal consistency, seeking to demonstrate for a valid scale that indicator items for a given construct load unambiguously on their own factor. This tests both convergent and discriminant validity.

Rasch models, one-parameter logistic models (1PL), are also internal consistency tests used in item response theory for binary items. Rasch models for polytomous items are also available. They are generalizations of 1PL Rasch model. Like Guttman scales, Rasch models test that the included items which are measuring a construct will form an ordered relationship (see Rasch 1960). A set of items may have ordered internal consistency even though they do not highly correlate (additive internal consistency as tested by Cronbach Alpha or factor structure). Ordered internal consistency reflects the difficulty factor, which means that correct response to a more difficult item will predict the response on less difficult items but not vice versa.

When factor analysis is used to validate the inclusion of a set of indicator variables in the scale for a construct, the researcher is assuming a linear, additive model. Linearity is assumed as part of correlation, which is the basis for clustering indicator variables into factors. With additivity is meant that items will be judged to be internally consistent if they are mutually highly correlated. However, items may lack high intercorrelation but have a strong ordered relationship. For this reason, many researchers prefer to use a Rasch model for scale construction, in preference to additive models like Cronbach Alpha or factor analysis.

Discriminant validity, the second major type of construct validity, refers to the principle that the indicators for different constructs should not be highly correlated. Discriminant validity analysis refers to testing statistically whether two constructs differ as opposed to testing convergent validity by measuring the internal consistency within one construct. In constructing scales, some researchers reject an indicator if it correlates more highly with a construct different from the one which was intended to be measured. Some researchers use $r = 0.85$ as a rule-of-thumb cutoff value for this assessment. Construct validity is not

distinct from the support for the substantive theory of the construct that the test is designed to measure, which is an issue for measurement models in drug development. Experiments designed to reveal aspects of the causal role of the construct may contribute to construct validity.

Content validity evidence involves the degree to which the content of the test matches a content domain associated with the construct. Content-related evidence typically involves subject matter experts evaluating test items against the test specifications. Content validity is also called *face validity* and has to do with items seeming to measure what they claim to do. In content validity, one is also concerned with whether the items measure the full domain implied by their label. Failure of the researcher to establish credible content validity may easily lead to rejection of his or her findings. For help one should consider the use of surveys of panels or experts and/or additional focus groups of representative subjects. In case of PROs, asking patients, are ways in which content validity may be established.

It is a challenging task to make sure that the measures operationalized by experts or common sense sufficiently address the concept of the later scale. There could also be a *naming fallacy*. Indicator items may display construct validity, yet the label attached to the concept may be inappropriate.

Criterion validity evidence involves the correlation between the test and a criterion variable (or several variables) taken as representative of the construct. The correlation with known and accepted standard measures or criteria is of interest. Ideally these criteria are direct objective measures of what is being measured. Where direct objective measures are unavailable, the criteria may be merely closely associated. For example, employee selection tests are often validated against measures of job performance. If the test data and criterion data are collected at the same time, this is referred to as *concurrent validity* evidence. If the test data is collected first in order to predict criterion data which is collected at a later point in time, then this is referred to as *predictive validity*.

Reliability. According to classical test theory, reliability is not a fixed property of a test but a property of test scores that is relative to a particular population. A reliability coefficient is computed for a sample. This is because test scores will not be equally reliable in every population or even every sample. For instance, as is the case for any correlation, the reliability of test scores will be lowered by restriction of range. Also note that test scores are perfectly unreliable for any given individual *i*, because, as has been noted above, the true score is a constant at the level of the individual, which implies it has zero variance, so that the ratio of true score variance to observed score variance, and hence reliability, is zero. The reason for this is that, in the classical test theory model, all observed variability in *i*'s scores is random error by definition (see above). Classical test theory is relevant only at the level of populations and samples, not at the level of individuals. Reliability cannot be estimated directly since that would require one to know the true scores, which according to classical test theory is impossible. Estimates of reliability can be obtained by various means. However, there is no one standard method. The method of assessing reliability must reflect the medical use of the instruments. Some of the statistical and psychometrical methods are as follows.

Frequently the *p*-value is cited as evidence of reliability: a significant Pearson correlation means a correlation significantly different from 0. But one should scatterplot the data and check for biased values. The concordance correlation coefficient addresses the concept of agreement. However, it can be misleading in that it summarizes the fit around the line of identity and, therefore, like the Pearson correlation, a value close to one may not denote lack of variability around the line.

If a cut point is to be used to classify patients, agreement of the classifications could be examined, using Kappa indices. Kappa is commonly used to measure reliability or agreement for nominal or ordinal variables; however, it also has limitations. If one method is a gold standard, then predictivity (sensitivity, specificity, or similar statistics) should be determined. Receiver operating characteristic (ROC)-type analyses have much to offer; also it can be argued that paying attention to the misclassifications, rather than the consequences of misclassification, may not result an appropriate comparison (Obuchowski 2005).

The intraclass correlation coefficient or its analogs (Bland and Altman 1996) is another class of models to test the consistency of ratings made by different observers when rating the same entity. However, the value of this method depends heavily on the sample used, and without repeated measurements, estimates of precision are impossible.

Stability of the response. The same form of a test is given on two or more separate occasions to the same group of examinees (test–retest). On many occasions, this approach is not practical because repeated measurements are likely to make changes within the rater (patient or observer). For example, the rater could adapt the test format and thus tend to score higher in later tests. A careful implementation of the test–retest approach is recommended. If appropriate and possible, parallel-test forms of the scale will be of great help in case of repeated measurement, which is the rule in most clinical trials, to control, for instance, memory and/or training bias. Extensive training of observers before entering the clinical trial is another method of reducing this kind of potential bias.

An aspect of reliability of special interest in drug development is the instrument's *sensitivity to change (responsiveness)*. In more general terms, it is that the measured scores are changing in direct correspondence to actual changes in the entity under treatment. There is a growing recognition that assessing the effect of an intervention should not only focus at the statistical significance of the differences in outcome measures between the experimental and the control group but should also focus at the relevance or importance of these outcomes. Estimating the magnitude of the difference between change scores in both groups, the difference between mean change scores may be expressed in standard deviation units with the effect size index (ES). One of the possible definitions has been developed by Cohen. Unfortunately, there is no agreed standard method for the estimation and the interpretation of the magnitude of intervention-related change over time or responsiveness assessed with outcome measures. For further details, see Middel and van Sonderen (2002) who are discussing advantage and limitations of several ES proposals.

Form equivalence is related to two or more different forms of test or questionnaire (sometimes called parallel version) based on the same content and administered in an identical way to the respondent. The presentation of a test (or questionnaire) one time as a paper–pencil test version and the next time as a computer-based test version is not regarded as being parallel versions and cannot be exchanged in a setting assuming equally valid and reliable. After alternate/parallel forms have been developed, they can be used for different persons or for several measurement occasions with the same person in a trial. This method is, for instance, very common in educational examinations to prevent communication between participating people. A person who took form A earlier could not share the test items with another person who might take form B later, because the two forms have different items. We should always consider the use of parallel-test versions in trials with intraindividual repeated measurements when we cannot exclude considerable training effect or change in the strategy of responding to the items caused by experience with the test.

Internal consistency is defined as the association of responses to a set of questions designed to measure the same concept. It is normally expressed by the coefficient of test scores obtained from a single test or survey. Usually, internal consistency is measured with Cronbach Coefficient Alpha, or its algebraically equivalent, the Kuder-Richardson Formula 20, when the data are dichotomous, or the Spilt-half method based on the assumption that two halves of a test is parallel except for having different variances. Cronbach Alpha, which is the most frequently used and easily available procedure in nearly every commercial, statistical software package, is defined by:

$$\alpha = \frac{n\bar{r}}{(1 + \bar{r}(n-1))}.$$

Here the Coefficient Alpha is based on the average size of item-to-total score correlations, sometimes named standardized Alpha. One could also use the item-to-total score covariances that may be more informative when the items have different variances.

To describe the logic of internal consistency more vivid, assume patients participating in

a postmarketing survey about drug D indicated for treatment of symptom S. They are asked to rate statements about their satisfaction with the treatment. One statement is "Drug D helped me very much in getting rid of the symptom." A second statement is "After intake of the drug I frequently experienced unusual headache." A third statement is "If the symptom will come back, I will use drug D again." People who strongly agree with the first statement would most probably agree with the third statement and vice versa. Patients, who agree with the second statement, will most probably disagree with statement one and, depending on the anticipated need for future treatment, or the availability of alternative therapy, will more or less disagree with the third statement. If the rating of the statements is patternless high and low among the participants of the survey, the responses are said to be inconsistent. When no pattern can be found in the patients' responses, probably the test is too "difficult" and patients just guess the answers randomly. Of course, different conclusions could be drawn from inconsistent results, like, the items may be reworded or items addressing similar aspects of patient's satisfaction may be added to the survey to capture the intended construct in a more reliable way. Internal consistency is a measure based on the correlations (or covariances) between different items or statements on the same questionnaire or test. It measures whether several items that presumably measuring the same construct are producing similar scores. The procedure of Cronbach Alpha is a statistic calculated from the pairwise correlations between items. The coefficient ranges between 0 and 1. In case where some or many items are negatively correlated with the total score, the coefficient can take on negative values even less than −1.0. One can check the effect of those items by reversing the item scoring and run the procedure again. As a rule of thumb, Alpha of 0.6–0.7 indicates acceptable reliability and Alpha of 0.8 or higher indicates good reliability. High reliabilities (0.95 or higher) are not necessarily desirable, as this indicates that the items may be entirely redundant. The goal in designing a reliable instrument is for scores on similar items to be related (internally consistent), but for each to contribute to some part a unique information.

In 2004, Lee Cronbach, the inventor of Coefficient Alpha as a way of measuring reliability, reviewed the historical development of Alpha: "I no longer regard the formula as the most appropriate way to examine most data. Over the years, my associates and I developed the complex generaliability (G) theory" (Cronbach 2004, p. 403). Discussion of the G theory is beyond the scope of this contribution. Cronbach did not object the use of Coefficient Alpha, but he recommended that researchers should take the following into consideration while employing this approach:

- Standard error of measurement is the most important piece of information to report regarding the instrument, not a coefficient.
- Independence of sampling.
- Heterogeneity of content.
- How the measurement will be used: Decide whether future uses of the instrument are likely to be exclusively for absolute decisions, for differential decisions, or both.
- Number of conditions for the test.

Ratings repeatedly performed on the same entity or object by one observer ask for reliability of the rater's judgment, which is called intra-observer or *intra-rater reliability*. The comparison between the rating of several raters on the identical entities (objects, persons, etc.) is called the interobserver or *inter-rater reliability*. Statistical methods for measuring agreement between categorical outcomes are well established. Cohen (1960) developed the kappa statistic as an agreement index for two binary variables. It has an appealing interpretation as a measure of chance-corrected agreement. Later, Cohen (1968) generalized the original kappa to the weighted kappa coefficient for ordinal discrete outcomes. Since its development, kappa with its extensions (Cohen 1960, 1968; Fleiss 1971, 1981; Fleiss and Cohen 1973 and others) has been well studied in the literature and broadly applied in many areas.

From the previous presentation and discussion, one may draw the conclusion that after considerable discussion about the scientific

value of validity evidence and the relation between reliability and validity, the message is that reliability is a *necessary* but not *sufficient* condition for validity.

Modification of the Method

The psychometrical methods of development, evaluation, and application of measurement tools may be even more useful for drug development if they are integrated part in the conceptualization of treatment, target population, and measurement in an early phase of drug development. Scales may then be more specifically selected, modified, or developed to depict essential aspects of person, disease, and interventional properties and are no longer restricted to the role of patient selection criteria or endpoint definitions for single clinical studies.

The commitment to plan drug development programs instead of isolated experiments and/or trials and start modeling of interventional effects right from the beginning of the program lets one to also think of big data analytics. Big data analytics (BDA) is defined as the application of advanced (exploratory) analytic techniques to very big data. The term big data is used for data of high volume with diverse data types and/or velocity (streaming data) (Gartner IT Glossary; http://www.webopedia.com/TERM/B/big_data_analytics.html). The analytical techniques comprise, e.g., predictive analytics, data mining, artificial intelligence algorithms, and language processing. The potential of an IT approach to drug development programs make installation of platforms suited to BDA necessary.

Advanced IT has influenced the way scales are presented and processed in experiments and clinical trials. As said before, "scale" should no longer be defined in the traditional way as a question or task that is physically presented as paper–pencil–tool or established on a wired apparatus with a monitor and keyboard. Scales are increasingly frequently presented to the responders as computerized survey on smartphones, tablets, or WEB-based on computers. The computerized versions of scales need special investigation in the validity, reliability, and eligibility of the tool and demand new concepts of statistical data analysis. Integrated measurement models are in line with new concepts programs of modern drug development (e.g., EMEA/127318/2007).

The view and opinion of patients (and relatives or health carers) on treatment outcomes will increasingly contribute to the development of new, effective, and safe medicinal drugs and translate pharmacological and clinical study outcomes into meaningful information for physicians, patients, and their relatives in everyday practice. Measurement concepts extended beyond the clinical disease models, and related target populations may then enable links from concepts of preclinical experiments to cost–benefit quantifications of marketed drugs.

References and Further Reading

Acquadro C, Jambon B, Ellis D, Marquis P (1996) Language and translation issues. In: Spilker B (ed) Quality of life and pharmacoeconomics in clinical trials. Lippincott-Raven, Philadelphia, pp 575–585

Agresti A (1996) An introduction to categorical data analysis. Wiley, New York

AMDP&CIPS (1990) Rating scales for psychiatry, European edn. Beltz Test GmbH, Weinheim

APA (American Psychiatric Association) (2013) Diagnostic and statistical manual of mental disorders: DSM-5, 5th edn. American Psychiatric Association, Washington, DC

APA (American Psychological Association) (2014) Standards for educational and psychological testing. American Educational Research Association, Washington, DC. http://www.apa.org/science/programs/testing/standards.aspx

Asadi-Lari M, Tamburini M, Gray D (2004) Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. Health Qual Life Outcomes 2:32

Biel V. Big data analytics. http://www.webopedia.com/TERM/B/big_data_analytics.html

Bland JM, Altman DG (1996) Measurement error and correlation coefficients. Br J Med 313:41–42

Borg I, Groenen P (2005) Modern multidimensional scaling: theory and applications, 2nd edn. Springer, New York. 2nd edn. Washington, DC: American Council on Education

Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20(1):37–46

Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 70:213–220

Cohen J (1977) Statistical power analysis for the behavioural sciences, revth edn. Academic, New York

Committee for Medicinal Products for Human Use (2005) Reflection paper on the regulatory guidance for the use of Health-Related Quality of Life (HRQL) measures in the evaluation of medicinal products. EMEA, London

Cronbach LJ (1971) Test validation. In: Thorndike RL (ed) Educational measurement. American Council on Education, Washington, DC, pp 443–507

Cronbach LJ (2004) My current thoughts on coefficient alpha and successor procedures. Educ Psychol Meas 64:391–418

Cronbach LJ, Meehl PE (1955) Construct validity in psychological tests. Psychol Bull 52:281–302

EMEA (2007) Innovative drug development approaches, EMEA/127318/2007. EMEA, London

Fleiss JL (1971) Measuring nominal scale agreement among many raters. Psychol Bull 76(5):378–382

Fleiss JL (1981) Statistical methods for rates and proportions, 2nd edn. Wiley, New York, pp 38–46

Fleiss JL, Cohen J (1973) The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 33:613–619

Food and Drug Administration (2000) Title 21 Code of Federal Regulations (21 CFR Part 11) Electronic Records; Electronic Signatures. http://www.fda.gov/ora/compliance_ref/Part11/

Food and Drug Administration (2006). Guidance for industry, patient-reported outcome measures: use in medical product development to support labeling claims, Silver Spring

Food and Drug Administration (2009) Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. Available from: http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf. Accessed 3 Jan 2017

Food and Drug Administration (2016) Clinical outcome assessment (COA): glossary of terms. Available from: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm. Accessed 3 Jan 2017

Food and Drug Administration (2017) Clinical outcome assessment qualification program. https://www.fda.gov/drugs/developmentapprovalprocess/drugdevelopmenttoolsqualificationprogram/ucm284077.htm. Page Last Updated: 06 Aug 2017

Gartner IT. Glossary. https://research.gartner.com/definition-whatis-big-data?resId=3002918&srcId=1-8163325102

Guilford JP (1946) New standards for test evaluation. Educ Psychol Meas 6(5):427–439

Hamilton M (1960) A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62

Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York

Kelly P (2015) Electronic questionnaires at statistics Canady. https://wwwn.cdc.gov/qbank/QUest/2015/s423-paul_kelly_quest2015.pdf

Kendell R, Jablensky A (2003) Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 160(1):4–12

Kundel HL, Nodine CF, Conant EF, Weinstein SP (2007) Holistic component of image perception in mammogram interpretation: gaze-tracking study. Radiology 242:396–402

Lienert GA, Raatz V (1998) Testaufbau und Testanalyse. Weinheim, Psychologie Verlags Union 1998 (6. Aufl)

Linden WJ, van der Hambleton RK (1997) Handbook of modern item response theory. Springer, New York

Little RJA, Rubin DB (1987) Statistical analysis with missing data. Wiley, New York

Messick S (1989) Validity. In: Linn R (ed) Educational measurement, 3rd edn. American Council on Education and Macmillan, New York, pp 13–103

Middel B, van Sonderen E (2002) Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integ Care 2.:2002:e15

NASA Report. ftp://ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf

National Institute of Mental Health. Research priorities. NIMH page on RDoC. http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml

Obuchowski NA (2005) ROC analysis. Am J Roentgenol 184:364–372

Osgood CE, Suci G, Tannenbaum P (1957) The measurement of meaning. University of Illinois Press, Urbana

Patrick DL et al (2007) Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value Health 10(Supplement 2):S125–S137

Powers JH, Patrick DL, Walton MK, Marquis P, Cano S, Hobart J, Isaac M, Vamvakas S, Slagle A, Molsen E, Burke LB (2017) Clinician-reported outcome assessments of treatment benefit: report of the ISPOR clinical outcome assessment emerging good practices task force. Value Health 20(1):2–14. https://doi.org/10.1016/j.jval.2016.11.005

Rasch G (1960) Probabilistic models for some intelligence and achievement tests. Danish Institute for Educational Research (expanded edition), 1980, Copenhagen. University of Chicago Press, Chicago

Rost J (2004) Lehrbuch Testtheorie – Testkonstruktion. Huber, Bern (2. Aufl)

Schadow G, McDonald CJ (2014) The unified code for units of measure. Regenstrief Institute, Inc. and the UCUM Organization 1988–2014. http://unitsofmeasure.org/ucum.html

Schwarz N (1999) Self-reports. How questions shape the answer. Am Psychol 54(2):93–105

Stevens SS (1951) Mathematics, measurement and psychophysics. In: Stevens SS (ed) Handbook of experimental psychology. Wiley, New York, pp 1–49

Thompson B (ed) (2003) Score reliability: contemporary thinking on reliability issues. Sage, Thousand Oaks

Walton MK, Powers JH III, Hobart J et al (2015) Clinical outcome assessments: conceptual foundation: report of the ISPOR clinical outcomes assessment – emerging good practices for outcomes research task force. Value Health 18:741–752

WHO. http://www.who.int/classifications/icd/en/

Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, Hareendran A, Wiklund I, Chong LY, von Maltzahn R, Cohen L, Molsen E (2009) Multinational trials – recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR patient-reported outcomes translation and linguistic validation good research practices task force report. Value Health 12(4.) 2009):430

Wilke RJ, Burke LB, Erickson P (2004) Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved labels. Control Clin Trials 25:535–552

William HL, Conway MA, Cohen G (2008) Autobiographical memory. In: Cohen G, Conway M (eds) Memory in the real world, 3rd edn. Psychology Press, Hove

Methods in Clinical Pharmacology

30

Lorraine M. Rusch and Clayton Dehn

Contents

Abstract

Early clinical development is often referred to as the "Valley of Death" in clinical trials because of the challenges involved in the translation from preclinical research to First in Human (FIH). The high attrition rate of drugs is widely attributable to the lack of data from the very basic and mechanistic assumptions made during the earliest phases of computational PK and screening, the limited preclinical studies executed in two species, and then the animal model factors used for allometric scaling.

Phase 1 is the first time that a novel compound is dosed to healthy normal volunteers, or in some specific cases, to patients with a targeted disease. The objective during this phase is to evaluate the risk to benefit to humans and to gain understanding of the administered doses and systemic drug concentrations, correlated to any observed adverse events via the route of administration intended for the product (IV, oral, subcutaneous, etc.). Increasingly, this phase is also concerned with pharmacodynamics, or physiological effects exerted by the novel drug over time.

This chapter is divided in two parts. The first part will elucidate the considerations which go into the design of early clinical research studies including drug disposition, pharmacokinetics, and pharmacodynamics and then clinical pharmacology clinical trial design, elements, and endpoints are examined in the second part.

L. M. Rusch (✉) · C. Dehn
Highpoint Clinical Trials Center, High Point, NC, USA
e-mail: lorrainemrusch@gmail.com; lrusch@highpointctc.com; cdehn@highpointctc.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_28

Part I: Analytical Approaches in Clinical Pharmacology

Pharmacokinetics (PK) Considerations in Early Clinical Development

Pharmacokinetics describes the body's effect on, or the fate of, an administered drug. It is mathematically defined based on the systemic (i.e., plasma concentration) absorption, bioavailability, distribution, metabolism, and elimination of the compound.

In general PK studies are designed to answer the questions related to the exposure, site activation of the drug, metabolism of the compound, and the proportionality between concentration and dose so that later modeling can effectively plan a true product dosing regimen. The major parameters are defined and demonstrated in equations as shown below in Table 1:

Nonclinical PK and biological data are used to demonstrate the scientific validity of the mechanistic hypothesis and that the novel drug can address an unmet medical need. The PK data from nonclinical species are designed to demonstrate that sufficient exposure that results in efficacy via a clinically relevant route in a basic formulation. Preclinical animal PK results are used to scale results to the hypothetical human parameters to estimate the first and lowest dose that will be administered to humans in the experimental First in Human (FIH) study. This first study is considered a milestone in drug development and has the potential to be quite dangerous, considering that safety studies in preclinical species are not entirely predictive of the human experience due to interspecies differences in physiology, genetics, diversity, size, immunology, etc. These differences can result in unexpected adverse events ranging from minor, transient observations up to and including death. For this reason, the calculation of the first administered dose to humans is critical and deserves specific attention.

The estimations based on preclinical models include the evaluation of multiple species and use the appropriate parameters (Cl_i, CL_s, $T_{1/2}$, Vd, etc.) as a factor to calculate the predicted PK of the drug in humans. The methodology includes the review of all in vitro data, especially the binding of drug to plasma proteins to understand the free drug concentration. Any special characteristics such as high Vd, nonlinear kinetics, compartmentalization, and metabolic reactions are considered across similar drug structures and used, via allometric scaling, to estimate the human plasma concentration and the $T_{1/2}$. These factors will form the basic understanding of risk for human exposure.

For example, from:

CL (human) = 40 CL (Rat) (L/hr)
Vdss (human) = 200 Vdss (Rat) (L)
$T_{1/2}$ (human) =4 $T_{1/2}$ (Rat) (hr)

It should be noted, however, that renally or biliary cleared drugs are more predictable than those that undergo metabolism and indeed

Table 1 Basic PK parameters, definitions, and equations

Category	Parameter	Definitions/Concept
Metabolism, elimination, and excretion	Half-life ($T_{1/2}$)	Time (hours) that it takes for the drug plasma concentration to decline by 50%
Distribution	Area under the curve (AUC)	The cumulative area under a curve constructed from a graph of plasma drug concentration versus time (h)
Distribution	Apparent volume of distribution (Vd/Vdss)	The conceptual volume that the entire amount of drug would occupy if it were evenly distributed in the plasma. It is the amount of drug in body divided by the plasma drug concentration
Metabolism, elimination, and excretion	Clearance (CL)	The volume of plasma from which the drug is completely removed over a unit of time (i.e., l/h). Rate of drug elimination/plasma drug concentration

those that are mainly excreted unchanged commonly have human half-lives 3–4X that of rats. Drugs which are mainly metabolized are estimated to have human half-lives 8–12x of rats, assumed because of differential enzyme expression.

Performance of a package of studies appropriate by length of exposure and route of administration intended for humans in at least two animal species, one of which must be nonrodent, is required to support the First in Human (FIH) exposure. These animal studies establish reasonable scientific evidence of safety and efficacy as well as a better understanding of the investigational compound's pharmacokinetics including proportionality and bioavailability. This will inform the FIH Maximum Recommended Starting Dose (MSRD).

First in Human Dose Selection

There are multiple empirical and mechanistic methodologies for selecting a starting dose in humans. The elimination or clearance of the drug under study is of prime concern to avoid accumulation to a toxic level. Allometric methods are commonly used to estimate these data from nonclinical species, which also contributes to the understanding of the volume of distribution (V). Early studies that are model-based such as physiologically based pharmacokinetic (PBPK) method can fortify this information. The FDA Guidance for Industry: Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers Appendix E summarizes the starting dose selection decision making strategy with a multistep decision tree which is recreated in Fig. 1.

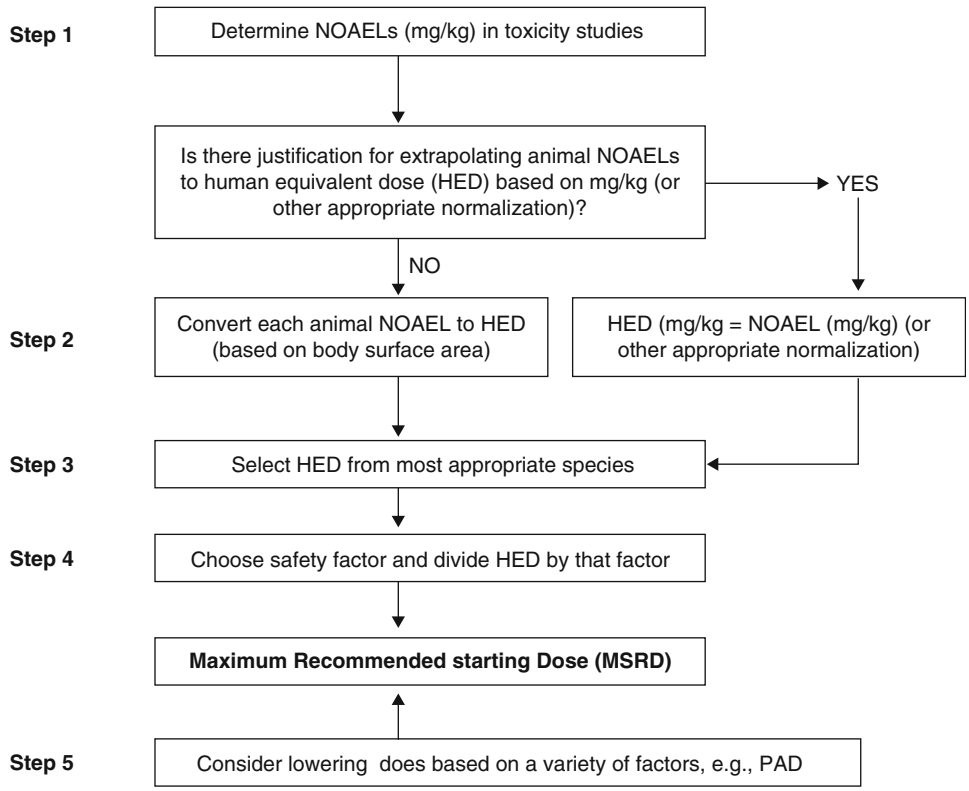


Fig. 1 Selection of maximum recommended starting dose for drugs administered systemically to normal volunteers (United States of America. Center for Drug Evaluation and Research, Food and Drug Administration 2005)

Identifying the Maximum Recommended Starting Dose (MRSD) for a FIH study involves the application of a "safety factor." The simplest approach to determining MRSD is 1/100th of the No Observed Adverse Effect Level (NOAEL) of the preclinical species. This represents the aggregate of an initial reduction to 1/10th of the NOAEL to account for interindividual variability within the preclinical species of interest, and a further reduction to 1/10th of the 1/10th (or 1/100th total) NOAEL to account for interspecies variability between the most relevant preclinical species and man (United States of America. Center for Drug Evaluation and Research, Food and Drug Administration 2005; Gad 2010). Another approach involves converting the preclinical NOAEL to a "humanized" NOAEL based on a body surface area conversion factor (CF) into a Human Equivalent Dose (HED): NOAEL (mg/kg) in preclinical species × CF in preclinical species = HED (mg/kg). Commonly a factor of 10 is applied to the HED to come to a Maximum Recommended Starting Dose (MRSD), but this is based primarily on convention. One of the most common methodologies, allometric scaling, uses the body weight of animals and respective clearance: $LC - a *BW^b$, Log CL = log a + b* log (BW) where BW is body weight.

These MRSD doses usually results in drug concentrations that are generally far below the level of bioanalytical sensitivity and detectable biological activity. Measurable levels in terms of concentration and activity may be as high as 1/10th the NOAEL level, depending on the anticipated concentration required at the site of activity. Putative plasma concentrations in humans can be back calculated from the body weight: plasma clearance relationship from multiple species to estimate the dosing which should result in a targeted plasma concentration in humans.

There are restrictive caveats to the utility of this most common approach. Principally, there is a significant need for data from multiple species and consideration for the clearance of unbound free drug which may contribute to an incorrect estimation. In those cases, only the unbound clearance should be used in calculations. Further, if the drug is subject to active transport, this may also skew results and in this methodology the route of elimination must be nonmetabolic for good estimation. As an alternative, a similar analysis can be conducted using volume of distribution versus body weight.

Metabolism in Early Clinical Pharmacology Studies

Novel drugs submitted to the FDA as an Investigational Drug Application (IND) for the goal of dosing a FIH study must evaluate nonclinical and clinical investigation of absorption, distribution, metabolism, and elimination (ADME). These studies qualitatively and quantitatively describe how the body processes a drug. The major concept around metabolism is understanding the bioavailability of a drug, meaning specifically for an oral drug how much of that drug is systemically available after it undergoes administration. The processes of disintegration, dissolution, absorption, and then first pass metabolism via the liver all occur prior to the drug entering systemic circulation. Even then some drugs are extensively bound to plasma proteins and thus unavailable for activation at the site of interest.

The confounding factor in the traditional list of ADME studies is the appropriateness of specific studies for large molecules versus the traditional small molecule drugs see Table 2 below as a guide. The guiding principle is that large molecules are readily degraded in the oral administration and not subject to the same types of metabolism as a small molecule compound. Being large also means they are not readily permeable for cell assays, do not bind to proteins, and are not subject to the traditional metabolism pathways such as cytochrome P450 enzymes.

Absorption and Distribution

Understanding the absorption of drug in humans is one of the objectives in a FIH study. Preferably this would be an excellent opportunity to include an IV arm as well as the expected route of administration.

Distribution of the drug is typically a function of the extent to which the tissues are perfused with blood and how much of that represents unbound

Table 2 ADME studies for small molecules versus large molecules

Study type	Small molecule	Large molecule
Absorption	Evaluate according to route (mostly oral) via in vitro assays	Not applicable; no in vitro permeability
Distribution	Plasma protein binding, target tissue distribution	Protein binding not applicable; sometimes target tissue distribution assessed
Metabolism	Metabolite identification, Cytochrome P450 phenotyping, inhibition and induction	Not applicable
Excretion	Sometimes	Not applicable
Bioanalytical	LC/MS, drug and metabolites	LC/MS analysis of drug and anti-drug antibody (ADA)

or free drug (i.e., not bound to plasma proteins) as well as how permeable the tissues are to the drug. V, or the volume of distribution (human total body water is typically assumed to be 42 L) is a depiction of the extent of that distribution (Table 1). The preclinical species studies define these parameters in the nonclinical species, and then applying allometric scaling to anticipate the human scenario can provide a decision gate for drug development. V may be the most important parameter, reflecting the drug's polarity (i.e., can it cross cell membranes?), its tissue binding (higher binding may mean higher distribution) and plasma protein binding (higher binding limits the distribution of free drug).

Metabolism

Metabolism refers to what the body does to a drug, how it is transformed in the body, and by what systems to evaluate systemic exposure as a safety indication. These studies provide insight to understand potential safety issues raised by differential clearance or inhibition of the drug or other potentially concomitant drugs. There are multiple experimental technologies for these studies, including well-characterized liver microsomes, hepatocytes, and intestinal preparations). Traditional IND packages include in vitro and in vivo enzymatic phenotyping (i.e., what enzymatic reactions cause biotransformation), cytochrome 450 inhibition and induction (which cytochrome P450 enzyme families may be inhibited or induced, causing other drugs to not be cleared systemically), stability plasma studies and preliminary LC/MS major, quantitative and qualitative metabolite identification from PK and TK samples.

Elimination

How a drug is eliminated from the body is as important as the other aspects of drug metabolism; if elimination is impaired or poorly understood, patient safety can ultimately be compromised. Elimination studies elucidate the major organs of elimination and how that potentially effects both targeted patients and special populations. The main study objectives include how elimination could affect concomitant drugs which might be metabolized via the same pathways and/or if there are resultant organ toxicities. Clearance (CL) is a constant that in proportion of the elimination of a compound quantitatively (ml/min) related to the actual loss from the site as a function of plasma concentration. Typically, the major organs for elimination are major and minor, mostly via renal and biliary excretion. In the case of the hepatic clearance, the elimination takes place via biotransformation of compound to metabolites, including biliary excretion. Renal clearance is the other major form of elimination where the compound is excreted untransformed via urine. Renal elimination is a function of clearance from the kidney into the urine/AUC_{plasma}, which depends on the kidneys' Glomerular Filtration Rate (GFR): CL_{renal} = fu* GRF-reabsorption + secretion. Biliary elimination, or the secretion into the bile from the bile duct, is an active process where the concentration is generally higher in the bile than the blood. There are many transport systems that work to specifically eliminate compound via biliary elimination.

Minor clearance routes include exhalation or other bodily fluids (sweat, saliva, tears, etc.).

ADME/AMES Studies

ADME (preclinical species) and AMEs (human) studies involve the administration of 100 μCi dose of a high specific activity ^{14}C or ^{3}H labeled drug to characterize the absorption, distribution (preclinical only), metabolism, and elimination of a compound administered via the intended route of administration. This study allows for quantitative analysis via sample processing and liquid scintillation or imaging analysis of the radioactivity associated with the drug or drug-related metabolites in urine, feces, plasma, and organs in terminal animal studies, providing full ADME data over time.

Metabolite Profiling Approaches in Clinical Pharmacology

Early metabolism profiling in humans can be tremendously advantageous to understand translational toxicities from nonclinical to clinical species or de-risk the drug development program. Approaches during first in human studies generally include pooled urine and plasma samples to scan for expected metabolites to understand the major and minor metabolites via LC-MS/MS, generally. Following identification and chemical synthesis of the putative metabolites and bioanalytical methods development, metabolites can be quantitated on an individual subject basis. Generally speaking, any metabolite >5% of the total drug radioactivity should be identified, quantified, and potentially monitored in future clinical studies. Specifically, each program is individual, and the requirements should be discussed with the appropriate regulatory authorities with the respect to the appropriate guidelines. The US FDA utilizes the Metabolites in Safety Testing (MIST) Guidance to outline metabolite identification strategies.

Part II: Human Studies in Clinical Pharmacology

For purposes of this book, clinical pharmacology will be considered in the context of Phase I and II clinical development. Phase I represents the first evaluation of safety, PK and PD in generally healthy normal volunteers. Phase II studies represent therapeutic exploration of the drug's efficacy, putative dosing regimen, and utility in specific medical indications; finally Phase III is the therapeutic confirmation of the principle. This chapter focuses on early clinical development in Phase I. Designs, logistics, and objectives generally include (Table 3):

General Phase I Study Design

A traditional scenario in a FIH study is a Single Ascending Dose (SAD) study where individual cohorts of volunteers (n = 6 active, n = 2 placebo) are administered drug in a blinded fashion over time and escalated from dose to dose based on safety, adverse events, and PK. In each cohort, the dose is administered, and blood sampling is progressed in a geometric series. Review of the

Table 3 Phase I/II study designs

Phase	Design	N	Subjects	Objectives	Logistics
1a	Single ascending dose (SAD)	40–60	Healthy normal volunteers, patients (special cases, i.e., oncology)	Safety, tolerability, PK, PD	4 months, $1 M minimum
1b	Multiple ascending dose (MAD)	30–50	Healthy normal volunteers, patients (special cases, i.e., oncology)	Safety, tolerability, PK, PD	6 months, >$1.5 M
2a	POC	50–200	Patients with targeted disease, 1 or 2 exploratory doses	Confirm POC hypothesis	>4 weeks, >$1.5 M to undefined
2b	Dose-finding	>200	Patients in targeted indication	Confirm dose response to optimize dose/tolerability/efficacy	>6 months, >$2 m to undefined

medical safety data, occasionally supported with PK results, allows the progression to higher doses in a stepwise approach. These FIH studies are costly and time-consuming, ultimately yielding a limited amount of safety and PK data in healthy normal volunteers. When exposure of healthy volunteers to the investigational drug is considered unethical, due to toxicities or therapeutic effect for example, the first administration may be made to patients. This is very common in the oncology fields where there is an expectation that the therapy may be beneficial. In these cases, the initial dose level may be higher than those that would have been used in healthy normal subjects. Dose escalated would continue to be based on safety and would additionally consider efficacy.

Clinical Protocol Essential Elements

Clinical protocol design can appear overly simplistic but should be carefully considered for safety and efficacy goals of the present and future. The major issue in Phase I and 2 is truly safety and that the subjects are not exposed to unnecessary risk, including overly invasive procedures.

The number of subjects, the preclinical safety information, inclusion and exclusion criteria for subjects, and the proposed dosing regimen are the primary considerations to begin the formation of the protocol. The sample size for Phase I studies are typically not of statistical significance due to their early development status; in Phase II and beyond, the question of superiority versus noninferiority/equivalence may drive the need to power to significance. The following design components are defined, and their consideration explained (Table 4):

Table 4 Design components of early clinical pharmacology studies

Primary endpoints	Primary endpoints are the main objectives of the study and secondary endpoints are generally exploratory in nature
Parallel or cross-over design	Parallel studies include subjects that are randomly assigned to one group for the entire trial. Cross-over studies include subjects that receive each of the treatments in a random order
Controlled or uncontrolled	A controlled study assumes a highly detailed I/E criteria and a control or comparator treatment potentially with a population
Placebo/active comparator	Placebo and active comparators allow for the true response to the drug to be monitored and "placebo effect" ruled out as well as comparator used to evaluate superiority
Open-label/blind	Open label studies mean that all parties involved in the trial can identify the trial drug. Blinded studies hide the identity of the trial drug by overencapsulation, identical formulations, and other means of third party blinding
Randomized	Simple or block randomization may be used, in addition assignments may be stratified by such factors as gender, risk factor, age, race, etc.

Early Proofs-Of-Concept, Pharmacodynamics Biomarkers, and Adaptive Study Design

Pharmacodynamics is a branch of pharmacology that describes the biochemical, physiological, and molecular effects on the body of an organism to which a drug is administered. This includes the relationship between the drug concentration at the site of action as well as the time course and intensity of the resulting therapeutic and adverse effects. Detecting a clinical therapeutic effect in an early phase setting is challenging if not entirely impossible. This is due in part to the small sample sizes, and relatively short duration of exposure achievable in this segment of clinical research. Sequential progression through the phases of clinical development with early phase being solely concerned with assessing safety and tolerability in normal healthy volunteers is, however, hopelessly inefficient. In response, the drug development industry has embraced pharmacodynamic proofs-of-concept (POC) supported with biomarkers as signals of efficacy in early phase drug development as "fail-early / fail-fast" cost-containing and de-risking strategies.

The phrase "biological marker" had been in use from the 1950s. During the 1970s, the first linguistic blending of the two words created the term "biomarker," which came into widespread use in the 1980s.

A biomarker reflects the interaction between a biological system and environmental challenges. It is a characteristic, substance, structure, or process that can be objectively measured in the body or its products and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to therapeutic intervention (Strimbu and Tavel 2010). A biomarker might be a soluble analyte measurable in otherwise unperturbed bodily fluids, a medical sign such as blood pressure, or it might be more conceptual in the form of in vivo pharmacodynamic testing involving the administration of perturbations to induce excursions which may be interpreted as signals of therapeutic response.

Improving morbidity, the disease condition, and mortality, the patient's risk of death, are classic clinical endpoints. These are often subjective representations from the patient's perspective of their health and well-being and how they feel, function, or survive. The frequency with which primary clinical endpoints such as survival can be measured makes them impractical and unethical for use in early phase POC research. In addition to supporting POC objectives and reducing the sample size required to meet them, biomarker applications in early clinical development may, in special circumstances, serve as surrogate endpoints supporting provisional registration. A surrogate endpoint serves as a substitute for a clinically meaningful one, and only a subset of well-established biomarkers that are consistently predictive of either positive or negative clinical outcomes will be fit for use as surrogate endpoints (Strimbu and Tavel 2010).

Interestingly, it is possible that the modulation of biomarkers within a patient might not correlate to the patient's disease experience or reflect any outward clinical manifestation. In fact, it is possible for a responsive, validated, surrogate-suitable biomarker to lack a fundamental relationship to the disease pathology. Such relationships may be better described and correlated than causal and indirect or adjacent to the disease process. Therefore, a biomarker's clinical relevance may be entirely decoupled from its validity as a surrogate endpoint (Strimbu and Tavel 2010).

Strategizing drug development using biomarkers and surrogate endpoints presents the risk of abandoning therapies that may be effective at improving clinical outcomes, but that fail biomarker-based efficacy tests. On the other end of the spectrum, there is risk of over-extrapolating successful surrogate biomarker endpoints into presumed clinical outcomes. For example, intense glycemic control has paradoxically been demonstrated to reduce survivability in at least some patients with diabetes (Skyler et al. 2009). Interventional marketing approval with surrogate endpoint demonstrated efficacy can only be provisional. In these instances, phase IV follow-up trials are required to correlate the surrogate marker findings to clinical endpoint anchors.

Early Proof of Concept and Adaptive Study Design

Adaptive proof of concept (POC) studies has been an increasingly popular concept in early clinical development. Specifically, for early clinical pharmacology studies, POC studies often incorporate FIM/SAD, MAD, and an arm of patient cohorts under one "superprotocol" in order to quickly understand safety and tolerability and potentially obtain early indication of efficacy. The options to explore early allow for early termination of a futile development program, re-estimate sample size, or select best dose for future studies. Additional arms can be incorporated for parameters to identify early issues such as cardiovascular safety, different routes of administration, effect of food on oral products, for example (Fig. 2).

The ideal candidate for such a study has putative endpoints that are measurable via surrogate biomarkers (i.e., viral load, specific protein expression, specific physiological responses). Additionally, these compounds should have enough of a therapeutic window to allow that patient safety across a wider range of concentrations should be assured.

These protocols are built with "if/and/or" scenarios with flexibility to adjust dose, patient number, and other points as needed. For example, sample size is generally calculated on the current

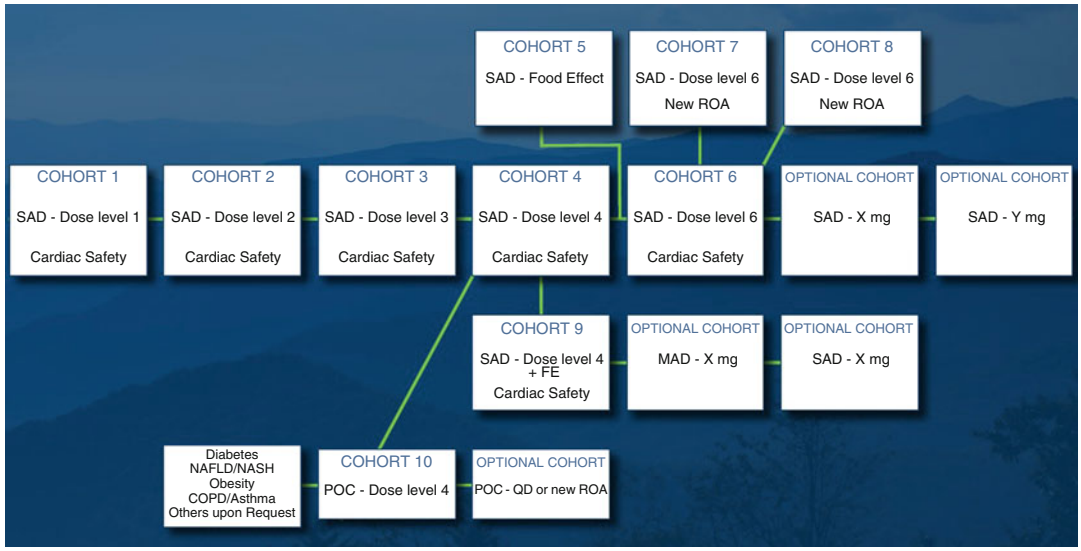


Fig. 2 Adaptive design FIM/SAD/MAD and POC in HNV and patient populations, respectively

information available, but in an adaptive protocol at a predetermined point this can be re-examined. At this point assumptions can be re-adjusted and sample size could be adjusted accordingly. This flexibility allows for studies to be appropriately powered and funds conserved (generally >10% of the costs of the separate studies due to overlap). Finally, if the early data from patients do not indicate success (lack of effect, lack of separation from placebo, etc.), then significant financial, scientific, and patient resources can be conserved.

The caveats around this protocol approach include the extensive preplanning requirements, consideration of stopping criteria and data review periods, the applicability of the patient population to the biomarker or endpoint readout, and the general flexibility required to manage such studies in terms of the clinical pharmacology facility, medical writers, and scientific input needed.

What Could Possibly Go Wrong?

As a word of caution, many drug development teams are very (understandably) eager to demonstrate Proof of Concept (POC) in patients and may view the FIH study as simply a checklist study which must be completed to proceed into more compelling patient studies. One of the most historically quoted missteps in drug development during a FIH trial was the Tegenaro incident. This excerpt from Wikipedia best describes the situation:

Phase I FIH trial of TGN1412 resulted in hospitalization of all six volunteers administered the drug, at least four of whom suffered multiple organ dysfunction. The trial was a double-blind, randomized, placebo-controlled study, with two of the eight subjects receiving a placebo, and six receiving 1/500th of the highest intravenous dose used in previous experiments with cynomolgus macaques. All the men were reported to have experienced cytokine release syndrome resulting in angioedema, swelling of skin, and mucous membranes, similar to the complement cascade in severe allergic reaction. The patients were treated with corticosteroids to reduce inflammation, and plasma exchange to attempt to remove TGN1412 from their circulation. The treating doctors confirmed in August 2006 that all six men had suffered from a cytokine storm, paradoxically, the men's white blood cells had vanished almost completely several hours after administration of TGN1412. The company claims that preclinical studies did not indicate any safety issues, but the effects were due to "on target" effects of the drug. Criticism has been raised that six participants were given the

drug in such a short time, which is against the recommendations of standard literature.

The case study outlined above created new industry guidelines and industry norms, though they are not necessarily appropriate for all cases. The careful evaluation of safety is crucial as the study initiates, and early data need to be reviewed carefully in real time as ultimately the animal data have limited translation. The following items should be considered during dose escalation:

- PK Considerations: is there any nonlinear, nonproportional PK observed? Is the bioavailability highly variable?
- Toxicity Observations: are there severe or potentially nonmonitorable toxicities? Are there any irreversible toxicities lacking warning signs? These of course become stopping points for the study depending on how the protocol is written.

Stopping Criteria in Clinical Protocols

Stopping rules are written into the protocol prospectively based on preclinical observations and any available clinical safety data in humans. These rules generally prescribe a dose at which the final dose is administered which leads to termination of the trial, the stopping of the dose administration to subjects within a cohort, stopping of the advancement to the next cohort, or any other part of the study. The criteria are generally set around C_{max} or AUC-driven exposure based on individual (not mean) data.

How PK and PD Inform Dose Finding Studies in Target Indications

In an ideal drug development scenario, the clinical development plan is formulated from the final market profile as the starting point and the entire program of studies designed to support that route of administration, expected dosing regimen, and nonclinical and clinical studies all aligned to the final vision of the drug in the intended patient population.

The appropriate dosing regimens must be established in the putative patient populations as part of the Phase 2 development program to demonstrate the efficacy at a pharmacologically effective dose. To reach this therapeutic dose (based on CL) in patients at the target site (i.e., absorption and distribution) at the right rate (based on $T_{1/2}$) via the appropriate route of administration (bioavailability), it must effect the receptor correctly (PD) and eventually be eliminated (metabolism). Additionally, the safety of medically compromised individuals (i.e., patients with renal and/or hepatic insufficiency) must be demonstrated via PK parameters in comparison to demographically matched healthy normal subjects, as outlined in the FDA Guidance entitled "Guidance for Industry: Pharmacokinetics in Patients with Impaired Renal Function-Study Design, Data Analysis, and Impact on Dosing and Labeling."

References and Further Reading

Gad SC (ed) (2010) Pharmaceutical sciences encyclopedia: drug discovery, development, and manufacturing. Wiley, New York. Copyright ©

Skyler JS, Bergenstal R, Bonow RO et al (2009) Intensive glycemic control and the prevention of cardiovascular events: implications of the ACCORD, ADVANCE, and VA diabetes trials: a position statement of the American Diabetes Association and a scientific statement of the American College of Cardiology Foundation and the American Heart Association. Diabetes Care 32(1):187–192. https://doi.org/10.2337/dc08-9026

Strimbu K, Tavel JA (2010) What are Biomarkers. Curr Opin HIV AIDS 5(6):463–466

United States of America. Center for Drug Evaluation and Research, Food and Drug Administration (2005) Guidance for industry: estimating the maximum safe starting dose in clinical trials for therapeutics in adult volunteers. Office of Training and Communications Division of Drug Information, Rockville

Pharmacodynamic Drug–Drug Interactions

31

Ming Zheng

Contents

Abstract

Concomitant medications with similar or opposite pharmacological effects can cause pharmacodynamic drug interactions. Pharmacodynamic interactions generally fall into three categories: additive, antagonistic, and synergistic. This chapter describes the

M. Zheng (✉)
Clinical Pharmacology and Pharmacometrics, Bristol-Myers Squibb Company, Princeton, NJ, USA
e-mail: ming.zheng@bms.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_29

commonly used empirical methodologies to evaluate pharmacodynamic interactions. Due to knowledge and data gaps in the systems involved, pharmacodynamic interactions are difficult to predict. Quantitative systems pharmacology models are emerging recently as promising approaches that integrate knowledge from multiple disciplines including drug pharmacology, systems biology, physiology, mathematics, and biochemistry. Readers are referred to other sources for more detailed discussions on such novel methodologies.

General Considerations

Drug–drug interactions (DDIs) can occur during polypharmacy therapies. DDIs can be the results of pharmacokinetic interactions, e.g., metabolic or transporter-mediated, or pharmacodynamic interactions. Pharmacokinetic interactions are the topic of discussion in another chapter in this book. Pharmacodynamic interactions can occur between coadministered drugs with similar or opposite pharmacological effects from direct interactions at drug receptors or targets. Pharmacodynamic interactions due to drug exposure changes as the result of pharmacokinetic interactions are not discussed here.

The resulting pharmacodynamic interactions can be generally classified into three categories: additive, antagonistic, and synergistic. Additive DDIs describe the resulting effect of the two coadministered drugs that is greater than the effect of each drug given alone. Examples of additive DDIs include sleeping aid medicines taken with alcohol which can result in greater drowsiness than caused by either sleeping aid or alcohol taken alone, or aspirin (antiplatelet) plus heparin (anticoagulant) which may increase bleeding risk. In antagonistic DDIs, one drug reduces or eliminates the effect of the other coadministered drug. Antagonistic DDIs can be beneficial to reverse dangerous drug effects. For example, vitamin K is a reversal agent for anticoagulant warfarin, and naloxone is used an antidote for narcotic overdose.

Synergistic DDIs describe the situations when the combined effect of two drugs is greater than the sum of the effects of each drug given alone. Synergistic DDIs are commonly employed in drug therapies. Drug cocktails have been developed and often used for treatment of diseases from HIV to cancer. Anti-HIV drugs are almost always combined to minimize the development of drug-resistant HIV virus. Harvoni® combines two HCV direct-acting agents, ledipasvir and sofosbuvir, that have significantly reduced the length of treatment compared to the previous interferon-based therapy. In the more recent exciting areas of immuno-oncology, immune check point inhibitors for different immune targets are combined to enhance the immune responses against tumors (Allard et al. 2018). For example, in patients with metastatic melanoma and less than 1% PD-L1 expression, exploratory subgroup analysis showed that the combination of PD-1 inhibitor nivolumab and CTLA-4 inhibitor ipilimumab had higher progression-free survival than either agent given alone (Opdivo® Package Insert, March 2018).

Unlike pharmacokinetic interactions where one may be able to predict the direction or magnitude of the DDIs based on the known disposition mechanisms of the drugs involved using mechanistic static or dynamic models (e.g., physiologically-based pharmacokinetic models) (FDA 2017; Jones et al. 2015; EMA 2012), it is far more difficult to predict pharmacodynamic interactions due to knowledge and data gaps in the systems involved. Quantitative systems pharmacology models are emerging recently as promising approaches that integrate knowledge from multiple disciplines including drug pharmacology, systems biology, physiology, mathematics, and biochemistry (Abernethy et al. 2011; Leil and Ermakov 2015; Gadkar et al. 2016). These models allow in silico hypothesis testing that would otherwise need to be evaluated experimentally and potentially prospective predictions following drug therapies either as a single agent or combination. The readers may refer to the above references for more detailed discussion of these approaches.

Dose- or Concentration-Response Curve Analysis

Purpose and Rationale

Analysis of the dose- or concentration-response curves and comparison of the curves from Drug A alone and Drug A and B combined allow exploration of pharmacodynamic interactions of the two drugs.

Procedure

Observed responses are plotted on the y-axis against the respective doses or concentrations where the responses were obtained on the x-axis. Response can be expressed as a percentage of the maximum response or change from baseline. The doses or concentrations on the x-axis are usually log-transformed so the corresponding dose- or concentration-response curve becomes sigmoidal (Fig. 1a). Log-transformation is helpful; especially, the dose or concentration range is relatively wide. Comparison of the resulting curves from Drug A alone and Drug A and B combined can yield useful insights on the nature of the pharmacodynamic interaction.

Evaluation

The dose- or concentration-response curve allows estimation of the maximum effect (E_{max}) and the dose or concentration at which half of the maximum effect is observed, ED50 or EC50, respectively. If the E_{max} of combined Drug A and B is greater than Drug A alone, an additive or synergistic effect can be assumed (Fig. 1b). However, unless the dose- or concentration-response curve is known for Drug B alone, additive or synergistic effect from the combination cannot be differentiated.

In the case of competitive antagonistic effect, the curve shifts to the right and Drug A needs to be

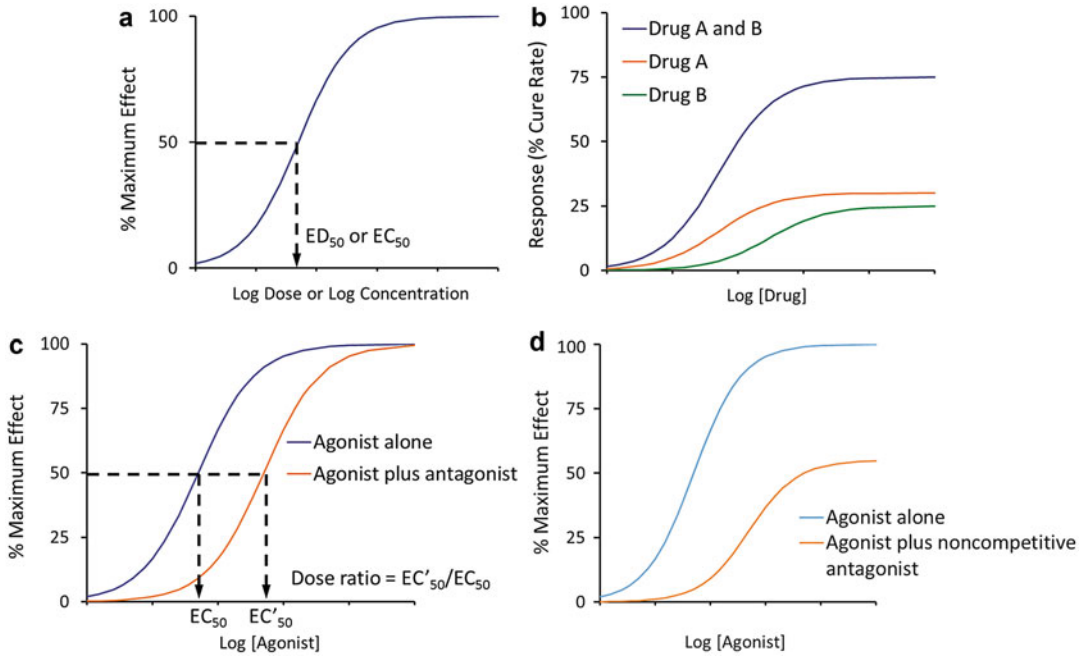


Fig. 1 (**a**) Semilogarithmic dose- or concentration-response curve. Log dose or concentration scale is for illustration purpose only and does not represent any actual values. (**b**) Concentration-response curves when Drug A or B is administered alone or combined. (**c**) Right shift of a concentration-response curve in presence of a competitive antagonist. (**d**) Change in concentration-response curve in presence of a noncompetitive antagonist

at higher doses or concentrations to reach the E_{max} (Fig. 1c). In the presence of noncompetitive antagonism after combining Drug A and B, not only the curve shifts to the right but the E_{max} is also decreased (Fig. 1d).

The linear part of the log-transformed concentration-response curve, between 20% and 80% of the maximum effect, can be described by the following equation (Gibaldi and Perrier 1982):

$$E = m \times \log C + b$$

where E is the response, m the slope of the linear portion of the plot, b the intercept.

Critical Assessment of the Method

Comparison of dose- or concentration-response curves is an empirical method to explore pharmacodynamic interactions. The ability to successfully conduct such analysis relies on the completeness of the available data. To be able to establish a good dose- or concentration-response relationship, it is important to conduct the study over a wide range of doses or concentrations. Due to obvious practical concerns over cost and time, studies meeting such criteria were not widely conducted. However, the advent of model-informed drug development in recent years predicts that more better quality studies will be conducted as well as more mechanistic modeling approaches as mentioned earlier.

Isobolograms

Purpose and Rationale

Direct comparison of the dose- or concentration-response curves from Drug A or B alone versus combination is not able to readily distinguish additive from synergistic effect. Isobologram allows a graphical analysis whether the pharmacodynamic interaction is additive, synergistic/supra-additive, or sub-additive. The basis of this analysis was based on the concept of dose equivalency where an equally effective dose of Drug A would add to the dose of Drug B. When the potency ratio of the two drugs is constant, the isobole is linear and shows an additive effect. Varying potency ratio leads to nonlinear additive isobole indicating synergistic or sub-additive effect (Chou 2006; Tallarida 2006, 2011, 2016).

Procedure

First the ED50 or EC50 is determined from the dose- or concentration-response curve for both Drug A and B when administered alone (Fig. 2a). The ED50 or EC50 values for the two drugs on the x- and y-axis are connected by a linear line when the potency ratio is constant (Fig. 2b). Any dose combination of Drug A and B based on the linear line can be tested to determine the combined effect.

In addition to 50% effect, other desired effect levels, e.g., 30%, 40%, 70%, etc., can also be studied and the resulting lines will be parallel to the 50% effect line.

Evaluation

The linear line in the isobologram (Fig. 2c) implies an additive effect from the combination. When different combinations of Drug A and B along the linear line in Fig. 2b are tested, if the combination results in synergism, then the 50% effect will fall below the linear line of additivity. On the other hand, the 50% effect from sub-additivity will fall above the linear line.

The linear additive isobole is commonly expressed as (Tallarida 2006):

$$\frac{a}{Ai} + \frac{b}{Bi} = 1$$

$$(0 \le a \le Ai, 0 \le b \le Bi)$$

where a and b are the dose pairs of Drug A and B, respectively, along the isobole, and Ai and Bi are intercept with the x- and y-axis and equivalent to the ED50 or EC50 of Drug A and B. As mentioned earlier, the additive isobole can also be drawn for other effective levels and the above equation still applies. The resulting isoboles will be parallel to the 50% effect line.

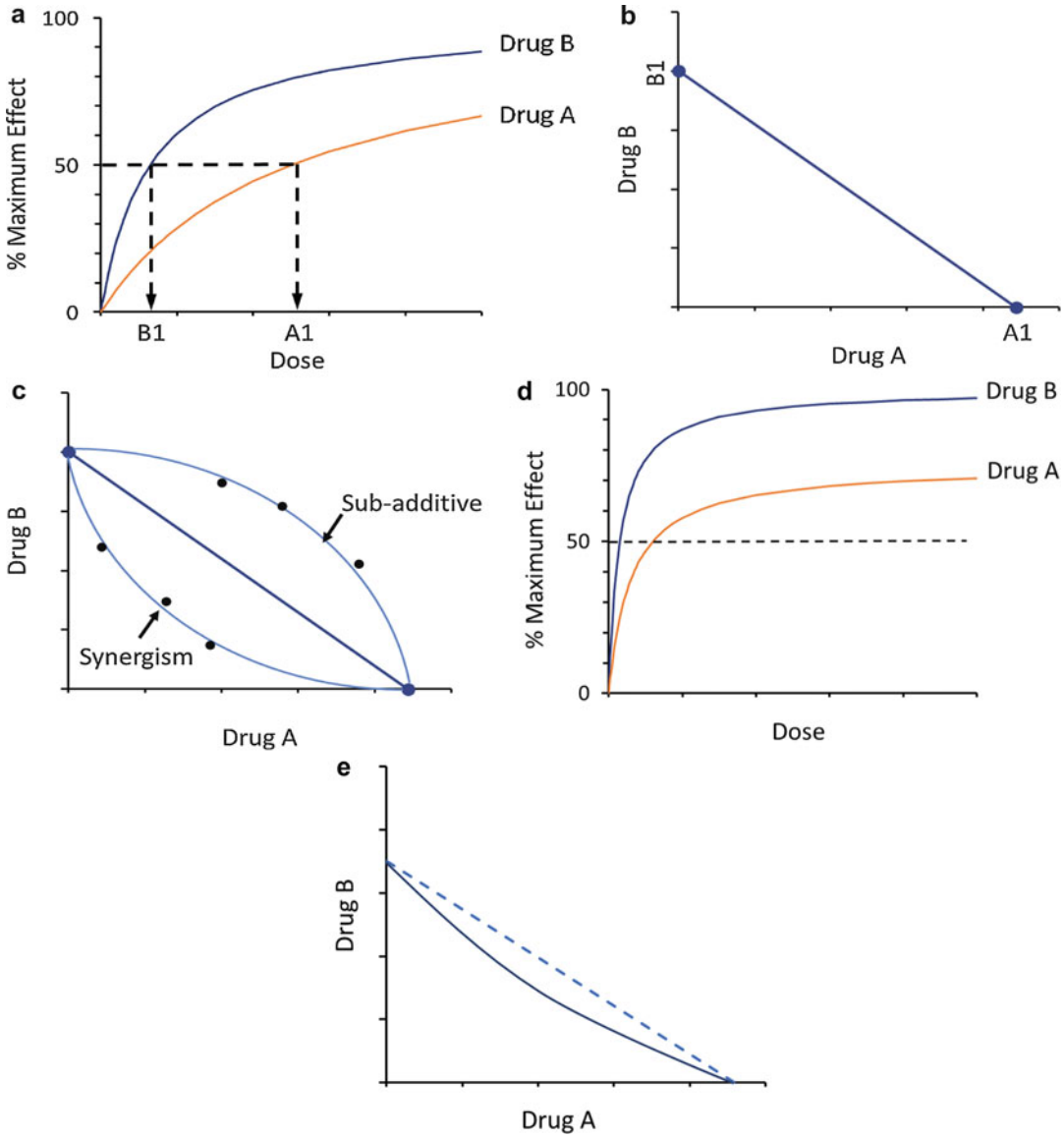


Fig. 2 (**a**) Determination of ED50 from dose-response curves of Drug A and B when administered alone. (**b**) Linear additive isobole when Drug A and B have a constant potency ratio. (**c**) Isobologram showing additive, synergism, and sub-additive effect from drug combinations. The black dots depict experimental data. (**d**) Dose-response curves from Drug A and B with different maximum effects. (**e**) A curvature isobole deviates from linearity when potency ratio of the two drugs is not constant

Critical Assessment of the Method

An isobologram has only two dimensions for two-drug combinations. For combinations of three or more drugs, it is not convenient to construct multidimensional isobologram. Secondly, accurate determination of synergism or subadditivity requires accurate determinations of ED50 or EC50 for each drug as well as the effects after combinations. Large variabilities in these determinations can lead to uncertainties in drawing the correct conclusion. Finally, there is also no general rule as to how far the combined effect should be from the isobole to conclude nonadditivity.

Modifications of the Method

As the actual determined ED50 or EC50 values from studies always have a variance, there is also a variance for the total additive dose. Consequently, every point on the isobole has an error. Calculation of variance or confidence intervals for isoboles can be incorporated for proper statistical comparison with the observed data points (Tallarida 2006, 2016).

As mentioned earlier, the additive isobole is linear if the potency ratio of the two drugs is constant. When two drugs have different maximum effect (Fig. 2d), the potency ratio is no longer constant. In this case, the additive isobole will show a curvature (Fig. 2e). Therefore, it is very important to first establish the relationship of the potency ratio of the two drugs before drawing conclusion from the shape of the isobole. The equations and calculations to derive the additive isobole when the two drugs have variable potency ratios were discussed in detail by Tallarida (2011).

Schild Plots

Purpose and Rationale

In the presence of an antagonist, it can be expected that there will be a rightward shift of the dose- or concentration-response curve of an agonist. In the case of competitive antagonism, the rightward shift is parallel to the original curve (Fig. 3a). For noncompetitive antagonism, there is a rightward shift as well as a decrease of the maximum effect (Fig. 1d). By testing with the antagonist over a wide range of doses or concentrations, one can quantify the potency of the antagonist (Schild 1957; Arunlakshana and Schild 1959).

Procedure

First a series of separate experiments are conducted, one without the antagonist, the others with the antagonist and increasing doses or concentrations of the agonist (Fig. 3a). The log-transformed dose- or concentration-effect curves are plotted. If the curves are parallel, competitive antagonism can be presumed. Next the constants of the following equation describing competitive antagonism are determined (Schild 1957; Arunlakshana and Schild 1959):

$$\log(x-1) = \log K2 - n\,pAx$$

where x is the dose/concentration ratio, K_2 the dissociation equilibrium constant for the antagonist, n the slope, and pAx the negative logarithmic molar concentration of the antagonist that produces the fold rightward shift of the agonist dose- or concentration-response curves. For a two-fold shift, pAx is expressed as pA_2. For competitive antagonism, $n = 1$. The dose/concentration ratio x is the ratio of agonist dose/concentration needed to achieve the same effect, e.g., half maximum effect, in the presence and absence of the antagonist. It is worth noting that the above relationship describing competitive antagonism is independent of the fraction of active receptors, i.e., receptors bound to the agonist.

The Schild plot is constructed by plotting log (x-1) against the negative logarithmic molar concentrations of the antagonist, which is pAx by definition (Fig. 3b).

Evaluation

The Schild plot intersects with –log[molar antagonist] when dose ratio is 2, i.e., $\log(x\text{-}1) = 0$. This value corresponds to pA2 which is a measure of the potency of the antagonist. When $n = 1$, $pA2 = \log K_2$, and $pA2 - pA10 = 0.95$. The pA2−pA10 difference of 0.95 can be used as a simple test for competitive antagonism. Several sources of errors may lead to underestimation of the pA2−pA10 difference: depression by a maximum agonist dose, a change in effect and receptor relationship (e.g., due to desensitization), and failure or delay of the antagonist to reach equilibrium (Schild 1957).

Critical Assessment of the Method

The application of Schild plots requires exhaustive tests of administering the agonist and

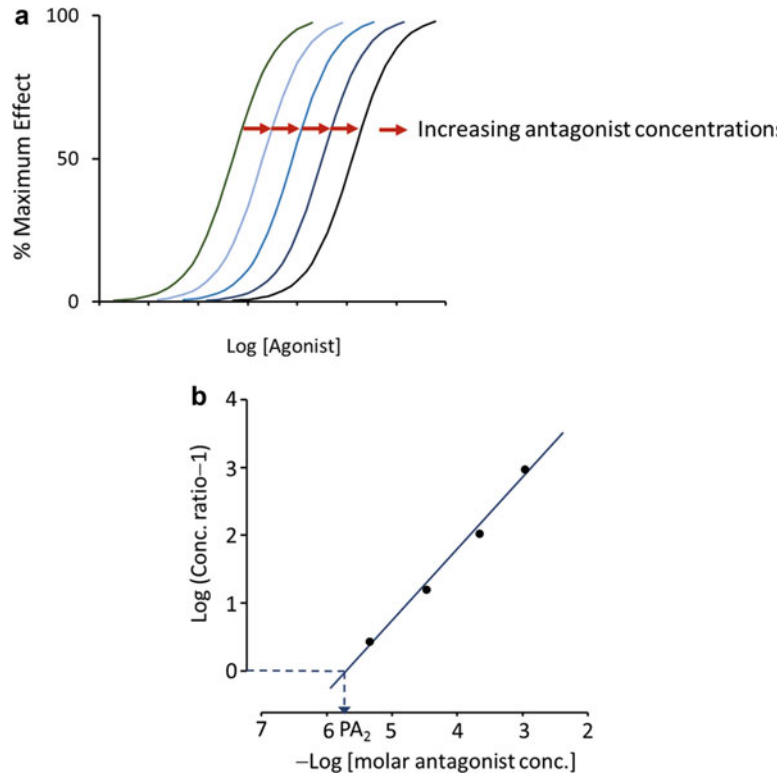


Fig. 3 (**a**) Right shift of concentration-response curves with increasing concentrations of competitive antagonism. The green curve is in the absence of an antagonist. (**b**) Schild plot

antagonist at various combinations of doses. This is possible under in vitro experimental settings but impractical under clinical settings due to ethical, safety, cost, and time restrictions.

Modifications of the Method

pAx is suitable for determining and comparing activities of antagonists which do not alter the slopes of the logarithmic dose- or concentration-effect curves of an agonist. In the case of noncompetitive antagonism where the curves are not parallel but become gradually flattened and the maximum effect declines with increasing doses of the antagonist, the pAx values vary with the concentrations of the agonist. In addition, both pAx and pAx differences, e.g., pA2–pA10, depend on the fraction of active receptors. One way is to describe the noncompetitive antagonistic activity in terms of the concentration required to produce a given reduction of the maximum effect, e.g., the concentration which reduces the maximum effect by half (Schild 1957). pAh was introduced to describe the negative logarithm of this molar concentration.

Factorial Design Trials

Purpose and Rationale

Factorial design trials are used to assess two or more interventions (i.e., independent variables) on the same endpoint in the same study. In a factorial design, two or more interventions are tested simultaneously using varying combinations. The main advantage of a factorial design is its efficiency which potentially reduces the sample size of the study, up to one-half, compared to two separate two-arm parallel studies (Pandis et al. 2014). Factorial design may allow evaluation of interactions only when the study is properly powered (Montgomery et al. 2003; Pandis et al. 2014; Pocock et al. 2015).

Procedure

The simplest example is a two-treatment 2×2 factorial design. In this design, subjects are randomly allocated to one of the four possible combinations of Intervention A and B: A alone, B alone, combination of both A and B, and neither A nor B (placebo or control group). In a 2×2 factorial design (Table 1), subjects are randomized twice: once to either the experimental or control arm for Intervention A and again to either the experimental or control arm for Intervention B. Alternatively, subjects may be randomized simultaneously to the four groups (Pandis et al. 2014).

There are three hypotheses one may want to test in a 2×2 factorial design: (1) the hypothesis on the effect of Intervention A, (2) hypothesis on the effect of Intervention B, (3) interaction hypothesis that the effect of Intervention A depends on the level of Intervention B, or vice versa. Factorial design trials are most commonly powered to detect the main effect of each intervention, on the assumption that there is no interaction between the interventions (Montgomery et al. 2003). A main effect refers to the effect of one intervention, either Intervention A or B, on the endpoint, regardless of the effect of the other intervention. As factorial design trials are powered to detect the main effects, they are under powered to detect interactions (Montgomery et al. 2003).

Table 1 An example of a 2×2 factorial design

		Intervention B	
		Yes	No
Intervention A	Yes	A + B	A only
	No	B only	Neither A nor B

A real-world 2×2 factorial design study can be illustrated with the CURRENT-OASIS 7 study (CURRENT–OASIS 7 Investigators 2010). The CURRENT-OASIS 7 study randomized 25,086 patients with acute coronary syndrome to either double-dose clopidogrel or standard-dose clopidogrel and either higher-dose aspirin or lower-dose aspirin. The primary outcome was cardiovascular death, myocardial infarction, or stroke at 30 days. The study was powered assuming no interaction between the two study-drug comparisons. The patients were first randomly assigned in a double-blind fashion to a double-dose clopidogrel or standard-dose clopidogrel regimen. This was followed by a second randomization in an open-label fashion to higher-dose aspirin or lower-dose aspirin. As summarized in Table 2, main effect analyses showed that there was no significant difference between a double-dose clopidogrel regimen and the standard-dose regimen, or between higher-dose aspirin and lower-dose aspirin, with respect to the primary outcome of the study. There was a nominally significant interaction between the clopidogrel and aspirin dose comparisons (Table 3). Since the interaction was not expected based on a known biological mechanism, the authors suggested that it was due to the play of chance. However, had there been a known mechanism, as the study was powered to only detect the main effects, the authors would not be able to draw conclusion on the interaction.

Interaction from a 2×2 factorial design may be visualized in a line graph. Using the CURRENT-OASIS 7 results, the interactions may be represented in Fig. 4. If the two lines are parallel, it may be interpreted that there was no interaction, whereas if the lines are not

Table 2 Results from CURRENT-OASIS 7 study

Main effect of clopidogrel dose on primary outcome			
	Double dose clopidogrel (percent rate)	Standard dose clopidogrel (percent rate)	P value
Primary outcome	4.2	4.4	0.30
Main effect of aspirin dose on primary outcome			
	Higher dose aspirin (percent rate)	Lower dose aspirin (percent rate)	P value
Primary outcome	4.2	4.4	0.61

Table 3 Interaction test of clopidogrel dose regimen according to aspirin dose regimen

	Primary outcome rates		
	Double dose clopidogrel	Standard dose clopidogrel	P value for interaction
Higher dose aspirin	3.8	4.6	0.04
Lower dose aspirin	4.5	4.2	

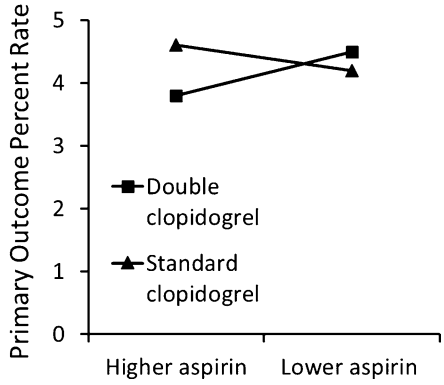


Fig. 4 Graphical interaction analysis of clopidogrel and aspirin

parallel, there may be an interaction. In this particular, as explained previously, since there is not a plausible biological explanation, it may be due to the play of chance.

Critical Assessment of the Method

The efficiency of reduced sample size from factorial designs is only realized if there is no interaction between the interventions. A study that is designed to specifically test interactions requires a much larger sample size, thus has no advantage compared to a multiarm parallel trial (Montgomery et al. 2003; Pandis et al. 2014; Pocock et al. 2015). It is therefore critical during the design stage to ascertain the assumption about interaction. If absence of interaction cannot be ascertained and the study is powered to detect main effects, then it is not possible to draw conclusion on interaction. Because factorial design trials are most commonly powered to detect main effects, it has limited power to detect interaction.

Modifications of the Method

In a 2×2 factorial design, two independent variables and two levels (two levels per independent variable) are investigated. In the above CURRENT-OASIS 7 study, the two independent variables were clopidogrel and aspirin, and each had two dose levels. Factorial designs are also able to investigate multiple independent variable and multiple levels. For example, a 3×3×2 factorial design is a design with three independent variables, the first two independent variables with three levels and the last with two levels.

References

Abernethy DR, Altman RB, Brouwer KLR et al (2011) Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms. An NIH White Paper by the QSP Workshop Group. October, 2011

Allard B, Aspeslagh S, Garaud S et al (2018) Immuno-oncology-101: overview of major concepts and translational perspectives. Semin Cancer Biol. https://doi.org/10.1016/j.semcancer.2018.02.005

Arunlakshana O, Schild HO (1959) Some quantitative uses of drug antagonists. Br J Pharmacol 14:48–58

Chou (2006) Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. Pharmacol Rev 58:621–681

CURRENT–OASIS 7 Investigators (2010) Dose comparisons of clopidogrel and aspirin in acute coronary syndromes. N Engl J Med 363:930–942

EMA (2012) Guideline on the investigation of drug interactions

FDA (2017) Guidance for industry, in vitro metabolism- and transporter-mediated drug-drug interaction studies

Gadkar K, Kirouac D, Parrott N et al (2016) Quantitative systems pharmacology: a promising approach for translational pharmacology. Drug Discov Today Technol 21–22:57–65

Gibaldi M, Perrier D (1982) Kinetics of pharmacologic response. In: pharmacokinetics, 2nd edn. Marcel Dekker, New York, p 222

Jones HM, Chen Y, Gibson C et al (2015) Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. Clin Pharmacol Ther 97(3):247–262

Leil TA, Ermakov S (2015) Editorial: the emerging discipline of quantitative systems pharmacology. Front Pharmacol 30(6):129

Montgomery AA, Peters TJ, Little P (2003) Design, analysis and presentation of factorial randomized controlled trials. BMC Med Res Methodol 24(3):26

Pandis N, Walsh T, Polychronopoulou A et al (2014) Factorial designs: an overview with applications to orthodontic clinical trials. Eur J Orthod 36(3):314–320

Pocock SJ, Clayton TC, Stone GW (2015) Challenging issues in clinical trial design. J Am Coll Cardiol 66 (25):2886–2898

Schild HO (1957) Drug antagonism and pA2. Pharmacol Rev 9:242–246

Tallarida RJ (2006) An overview of drug combination analysis with isobolograms. J Pharmacol Exp Ther 319(1):1–7

Tallarida RJ (2011) Quantitative methods for assessing drug synergism. Gene Cancer 2(11):1003–1008

Tallarida RJ (2016) Drug combinations: tests and analysis with isoboles. Curr Protoc Pharmacol 72:9.19.1–9.19.19

Medical Devices: Definition and Clinical Testing

32

Lea Wettlaufer and Daniela Penn

Contents

Abstract

Similar to drugs, medical devices need to be approved before being CE-marketed. Clinical data, so-called preclinical data, have to be obtained and gathered, and a technical file has to be compiled including a clinical evaluation. In some cases, clinical trials have to be conducted before the approval of the medical device due to lacking of clinical data. This chapter defines medical devices: it describes the regulatory context of approval, the most important parts of the technical file, and the ways and methods of clinical testing required to get the medical device approved and on the market.

L. Wettlaufer (✉) · D. Penn
Medisis, Neustadt, Germany
e-mail: lea.wettlaufer@gmx.net; daniela.penn@medisis.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_70

Regulatory Requirements

In Europe, medical devices and in vitro diagnostics (IVDs) are regulated by the EU guidelines Medical Device Directive (MDD, 93/42/EEC), the IVD 98/79/EEC, and the AIMD 90/385/EEC. The MDD and the AIMD are now replaced by the Medical Device Regulation (MDR) 2017/745 from May 25, 2017. The MDR will be in force from May 26, 2020 on. The same applies for the IVD; it will then be replaced by the IVDR.

National law is governed by the EU Directives. In order to regulate national differences and laws, national regulations differing from European member state to member state are in force (Fig. 1).

Objective of the MDD being still in force during the transition phase and the newly established MDR is to ensure a free movement of goods through a standardized regulation of medical devices in Europe (MDR, Articles 2 and 4). Furthermore, the safety of medical devices shall be fulfilled by standardized EU requirements (MDR, Article 3 and Annex I). At the moment, during the transition phase regarding the MDR, the MDD is still in force. Therefore both, the MDD and MDR, will be referenced in this chapter.

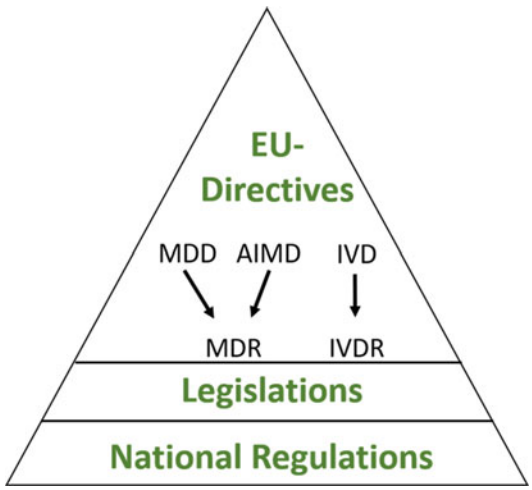


Fig. 1 Dependencies of the European laws to regulate medical devices

Medical Device Definition

In order to ensure a standardized understanding for the definition of medical devices, the MDR defines a medical device (MDR, Article 2). According to the MDR, medical devices are instruments, devices, implants, reagents, materials, and software that are intended to fulfill a medical purpose as defined by the manufacturer or by their intended use. This can be a single or a combined device.

The medical purpose includes diagnoses, monitoring, therapy or relief, prophylaxis, and prognosis of diseases. Furthermore, treatment of injuries or disabilities is also addressed as medical purpose.

This includes the treatment or relief of pain. For example, an active cupping device may contribute to pain reduction and to an improved physical effect (Teut et al. 2012; Lauche et al. 2013; Cramer et al. 2011). At the same time, it fulfills its medical purpose to concomitantly improve and treat an injury (in the sense of pain). The skin is dragged into the cupping device by means of the mechanical impact, and this increases the blood circulation and supply with nutrients of the affected tissue (Li et al. 2017; Cramer et al. 2011; Larsson et al. 1999). During this procedure, the effect is not produced by medications or immunological medicines since only mechanical energy activates endogenous functions resulting in a medical and therapeutic effect.

There is a legal distinction between the effect of drugs and medical devices. Drugs have a pharmacological, pharmacodynamics, immunologic, or metabolic effect, while medical devices act in a mechanic, physical, chemical, or physicochemical way.

In order to define a medical device, it is important to distinguish between the medical impacts of a product. Like a drug, the main effect of a medical device can be in- or outside of the human body. It is important that the main effect on the body is not caused by pharmacological, immunological, or metabolic drugs (MDR, Article 2).

For a medical impact on the human body, we already had the example of a cupping device. Medical devices in the human body, which are used for treatment, relief of pain, or therapy of diseases, can be implants (e.g., cardiac

pacemakers, hip or dental implants, etc.) or substance-based medical devices. It is possible to compare them with liquid drugs. Nevertheless, they fulfill the legal definition and effect of a medical device. Liquid herbal cough syrup, for example, can easily be mistaken for a drug. Herbal cough syrup, which lays a protective film over the irritated mucous membranes, carried out by its components, achieves the mitigation of the irritation of the throat by its mechanical impact. Through this protective film, the natural protective function of the mucous membranes is improved, and, thereby, the irritation of the throat and hoarseness is mitigated.

Therefore, herbal cough syrup, with its components, acts in a mechanical way and has no pharmacological, metabolic, or immunological effects.

In order to achieve their effect and clinical performance, medical devices can be supported by pharmacological, immunological, or metabolic drugs (MDR, Article 2). An example of a combined product is bone cement where an antibiotic has been added (drug and medical device). The target of that kind of bone cement is to anchor implants or prostheses directly in the bone (Kühn et al. 2017). This results in the compensation for disabilities or injuries, so the bone cement is defined as a medical device. The addition of the drug (antibiotic) to the bone cement makes the patient feel better and prevents infection (Kühn et al. 2017).

If a product supports the investigation, replacement, or modification of the anatomy or a physiological, respectively, pathological process or condition, it is also defined as a medical device (MDR, Article 2). In addition to the clearly defined medical devices, products controlling, supporting, assisting, cleansing, disinfecting, or sterilizing other medical devices are also defined as such. Examples are sterilizers for instruments used during surgeries or disinfectors of bedpans.

In Vitro Diagnostics as Medical Devices

In vitro diagnostics (IVD) are also defined as medical devices (MDR, Article 2). Often, they are laboratory systems, e.g., devices, kits, etc., supporting the provision of information through in vitro investigation. They investigate samples externally, i.e., outside of the human body (urine, blood, tissue, etc.). IVDs support physical or mental interferences, predispositions of health condition, detect (in-)compatibility, effects of a therapy, or the estimated reaction, and they can assist in the detection or observation of therapeutic measures by obtaining information of the physiological or pathological conditions (IVDR, Article 2).

Simple IVDs, which can be used by patients themselves, are pregnancy tests or blood glucose level tests, for example.

This list of examples and definitions shows that the form and application options of medical devices are extensive. The definition of a medical device according to the MDR means that medical devices must always have a clinical claim to support the patient. During the application of the medical device, the benefit must always outweigh the risk of harm.

Essential Requirements for Medical Devices and the Technical File

It is only possible to place medical devices on the European market if they comply with the strict safety requirements of the European Union. The affixation of the CE conformity mark is the result of fulfilling these requirements (French-Mowat and Burnett 2012).

In Europe, notified bodies (NB) and authorized representatives are responsible for the CE marking process and conformity assessment procedures. The NBs are designated by the respective competent authorities of the respective European member state. They monitor and ensure compliance with regulatory provisions.

It is up to the manufacturer to ensure that their medical device complies with the essential requirements of the relevant EU legislation. The CE marking process comprises the following aspects:

- Checking of applicable directives and annexes
- Selection of the respective conformity assessment procedure/route

- Registration of the medical device at the national competent authority (in Germany DIMDI authorities, e.g., in Switzerland Swissmedic)
- Preparation of the technical file
- Fulfilling the essential requirements
- Preparation of the declaration of conformity
- Submitting it to the NB or to the national competent authority (if applicable in case of Class I medical devices) for certification
- Application of the CE mark and marketing of the medical device
- Implementing vigilance and post-market surveillance actions by monitoring safety, efficiency, and reviewing experience of use and any action required during the product life cycle

Irrespective of the class of the device, all medical devices shall:

- Comply with the essential requirements
- Be object of an evaluation of clinical performance and safety by means of a clinical evaluation and, if applicable, already by a preclinical evaluation or clinical trial
- Fulfill the reporting requirements under the medical device vigilance system
- Have a CE mark

The essential requirements (ERs) can be found in Annex I of the MDD, and the essential performance and safety requirements are detailed in Annex I of the MDR. The latter replace the previous MDD ERs and also define the key aspects to address within the technical file.

Classification and Inherent Risk

All medical devices are placed into one of the four categories using the classification rules listed in MDD Annex IX. They are also listed in the MDR in Annex VIII.

The categories are:

- Class I (including Is and Im)
- Class IIa
- Class IIb
- Class III

with Class III being the class of medical devices with the highest risk. The classification of the medical device is defined by its intended purpose. Several aspects are important for the classification of the medical device including the duration of contact with the body and application, degree of invasiveness, and local versus systemic effect. The following table provides details for each class (Council of the European Communities. 2007) (Table 1).

Technical File (Technical Documentation)

The essential requirements (ERs) are evidence for the conformity of the medical device with the respective guidelines. These guidelines include detailed requirements for the medical device that must be fulfilled and verified as well as validated in the technical file. Each essential requirement in the checklist of the ERs of a technical file contains each requirement of the MDR. Here, insight in the standards, references, and documented evidence (test reports, drawings, descriptions, etc.) for every single requirement shall be provided by the manufacturers.

In the following, the most important parts of the technical file are described in more detail.

Risk Management

Risk management is an integral part of the ERs and, therefore, of the technical file. The fundamentals of the so-called integrated security are explicitly claimed in the essential requirements of the MDD still in force and the MDR in its transition phase.

A well-working risk management system, similar to a quality management system, is the standard for manufacturers to legally launch the medical device. Therefore, several parts of the risk management annexes are included in the MDD or the MDR.

The standard ISO 14971 “Medical devices - Application of risk management to medical devices” has been enforced as an equal evaluation

Table 1 Examples of product classification

	Classification	Risk	Description	Examples
	General controls			Hospital beds, bedpans
I	Sterile (Is)	Low	Most noninvasive devices that do not interact with the body	Sterile plasters
	Measuring (Im)			Thermometers, weighing scales
IIa	Special controls required: may include special labeling, mandatory performance standards, and post-market surveillance	Medium	Exchange energy with a patient in a therapeutic manner or are used to diagnose or monitor medical conditions. Generally invasive but limited to natural orifices, if hazardous to a patient then it becomes a Class IIb	Powered wheelchairs, hearing aids, ultrasonic diagnostic equipment
IIb	Special controls (as IIa)	Medium	Most surgically invasive/ active devices partially or totally implantable in the body. May modify composition of body fluids	Infusion pumps, ventilators, surgical lasers
III	Pre-market approval is the required process of scientific review to ensure the safety and effectiveness of these	High	Support or sustain human life and are of substantial importance in preventing impairment of human health, or which present a potential, unreasonable risk of illness or injury. Device that connects directly with the central circulatory system or CNS or contains a medicinal product	Many implants: vascular and neurological, replacement heart valves, silicone gel-filled breast implants, and implanted cerebella stimulators

of the risks for the manufacturers of medical devices for the authorities and notified bodies.

The requirements of the ISO 14971 are valid for all medical devices, IVDs, and in all phases of the product life cycle. The requirements include the following aspects:

- Risk estimation
- Risk evaluation
- Risk control
- Observation efficiency

Risks can occur anytime in the product life cycle. They are determined at a certain point of the cycle and can be minimized by a measure at a different moment of the product life cycle.

It is necessary to evaluate the residual risks versus the expected benefit before the decision to apply a medical device with a specific intended use in the defined indications for the respective patient population by the defined users (risk–benefit analysis).

The manufacturer is responsible for the risk definition. It needs to be well proved and state of the art.

The decision is based on:

- Public law
- Certain habits
- Values and awareness of the risk

The monitored risks are primarily evaluated based on the relationship with the affected patients or users. Risks which can harm the user or third parties need to be monitored. Furthermore, damage to animals, buildings, belongings, or the environment also needs to be monitored.

Depending on the implemented safety measure, the value of the measure will depend on the minimization of the respective risk. The implementation of a safety measure will reduce the occurrence of the risk a lot more than just a warning sign. On the other side, inherent safety can

eliminate the risk completely or even reduce the severity of the anticipated harm.

The medical device directive requires the manufacturers to proceed in the following order:

1. Inherent safety: making medical devices inherently safe.
2. Precautions: if this is not possible, risk-minimizing measures have to be implemented.
3. Notes: if this is not possible or if there is insufficient information, then label the remaining risks.

The order of these risk-minimizing measures must be adhered to. It is only allowed to use the next lower measure if it is impossible to use a superior one.

Usability

Usability is a fundamental requirement for medical devices according to Annex 1 of the MDD or MDR. Medical devices need to be manufactured to work under normal conditions of use and fulfill their intended purpose. Medical devices need to be safe and effective and shall not harm the patients or affect the health of the user or third parties (MDR, Annex I).

According to the European standard EN 62366, the usability is defined as interface between the medical device and user. The property of the usability can comprise the efficiency, learnability, and satisfaction of the user (IEC 62366, Section 5.3).

The interface between the medical device and user often results in use errors (incomprehensible operating, false operating, wrong application, etc.). These errors can influence the performance, safety, and complete benefit of the medical device.

Frequent operating functions or the intended context of use are related to a high potential of use error. Deficiencies such as inadequate labeling or ambiguous images can lead to an incorrect application of the medical device (IEC 62366). The context of use of the medical device, e.g., in combination with accessories, can have a significant impact on usability (IEC 62366, Section 5.3).

A laser device for the treatment of pain and wounds can serve as an example for a use error of a medical device. Laser therapy devices are available with different power levels.

The example laser has a power of 30 W and cures skin tissue with a transmitting power of 2 mW. With an additional tip focusing the laser beam, a transmitting power of 22 mW can be achieved. For the application of a laser with a power of 100 W, a transmitting power on the tissue of approximately 10 W is needed. With an additional tip, the laser would be so powerful that the tissue would suffer an extensive damage. The use of such a tip and the focus of the laser beam would bear a potential risk for the patient. By means of usability testings being part of the validation of the medical device, use errors can be identified and corrected, and this finally results in the safe use of the medical device. The use of such a tip in the example above would have been prevented.

Biocompatibility

According to the MDR, biocompatibility of a medical device is defined as biological safety of the device shown by preclinical and general clinical data (MDR, Annex II, Section 6.1). Biocompatibility needs to be tested for all materials of the medical device, which have direct or indirect contact with the human skin or with body fluids of the patient or user. For the evaluation of the medical device's biological safety, this clinical data is analyzed in form of a biocompatibility report and within the clinical evaluation.

Standards, such as the international ISO standards, describe harmonizing and acknowledged procedures which need to show and fulfill specific regulations. The ISO 10993 is an internationally acknowledged standard for the evaluation of the biological safety of medical devices. In Section 3.1 of the ISO 10993-1, the biocompatibility of a medical device or a material is defined as the ability to perform well in a specific application with an appropriate host reaction.

Biological testings of the material of the medical device serve as evidence. With these testings,

the characteristics of the material of the medical device, which are able to influence the biological reaction, are evaluated. The evaluation of the biological safety also considers the impact of soluble chemicals or morphological characteristics, if applicable (ISO 10993, Section 3.1).

According to ISO 10993, biological safety describes the freedom of non-acceptable risks in context with the foreseen use. It can be proved by means of verification testings of the material being in direct contact with human skin or body fluids (ISO 10993-1, Section 3.3).

The final biological evaluation of medical devices comprises the testing of several relevant aspects. This includes the analysis of the material, predicted additives, production-related contamination, packing material, soluble ingredients, and metabolites in the medical device (ISO 10993, Section 4.3).

A complete biological testing of a medical device can be supplemented by additional tests. An additional evaluation like cytotoxicity testings (ISO 10993-5), sensitivity and irritation testings (ISO 10993-10) or testings on hemocompatibility (ISO 10993-4), etc. can become necessary. This depends on the impact, respectively, application, and site of application of the medical device. A cooling pack, which is applied on intact skin, for example, should be tested on cytotoxicity according to IOS 10993-5. Implants that come into contact with damaged skin, tissue, blood, or even mucosa need to be tested on sensitivity and irritation according to IOS 10993-10. Furthermore, a chemical evaluation (ISO 10993-18) is necessary (ISO 10993, Section 6.1.3).

In the clinical evaluation, the clinical data on the biological safety of the medical device are also used for the evaluation of the clinical safety of the medical device.

Clinical Evaluation of a Medical Device Based on Clinical Data

According to the MDR, the clinical evaluation is a systematic planned process to continuously collect, analyze, and evaluate clinical data of a medical device. This process has the objective to show the clinical safety, clinical performance, and the clinical benefit of the medical device. The intended use of the medical device, as defined by the manufacturer, is confirmed (MDR, Article 44).

The intended use includes information on the application of the product, the user group, patient population, and the application environment. In the clinical evaluation, this information is verified by means of clinical data pertaining to the evaluated medical device or to an equivalent medical device. Furthermore, the respective indication(s) or the field of application of the medical device shall also be verified by means of clinical data. The main focus of the clinical evaluation of medical devices is the evaluation of the risk–benefit–relation of the medical device.

In regard to the risk–benefit ratio, the clinical benefit is defined as the positive effect of a medical device on human health. Such a positive effect shall be confirmed and must be measurable and patient-relevant, including the results of a diagnosis. A positive effect on the treatment or on public health is also defined as a clinical benefit (MDR, Article 52).

The applicable guideline for clinical evaluations is the MEDDEV 2.7/1 Revision 4. If the clinical evaluation report (CER) is written according to this guideline, the manufacturer also complies with the MDD/MDR requirements and the ERs.

Equivalent Medical Device (according to MEDDEV 2.7/1 Rev. 4)

Equivalent medical devices are comparator products to the evaluated medical device. They are equivalent to each other and in some aspects similar. The evaluation of equivalence is done by comparing the following aspects of the medical devices:

clinical/technological/biological

The evaluated and possibly equivalent medical device must be equal in terms of the clinical (intended use, indication, application) and biological (material) aspects. They must be similar in terms of technical parameters. If all of the three aspects are fulfilled, the medical devices can be considered to be equivalent.

In order to show the clinical benefit of a medical device, a cooling pack may serve as an example. The intended use of the cooling pack is the physical transmission of cold on human skin. The

treatment is part of cryotherapy. Cooling packs are often used for the relief of pain, swelling, or sports injuries.

Within these indications, the clinical benefit of the medical device may be assumed. In case of pain or swelling, the benefit of a cool pack can be pain reduction or a decrease of the swelling. If the manufacturer wants to advertise that kind of benefits, these must first be proved by clinical data within the clinical evaluation. Clinical data collected after the marketing of the medical device resulting from post-market surveillance and clinical follow-ups (PMCF), if applicable, are also defined as clinical data (MDR, Article 48).

The clinical performance and safety of the medical device shows its clinical benefit. The clinical performance is defined as the ability to fulfill the intended use of the medical device by means of its technical, functional, and diagnostic characteristics (MDR, Article 52). Clinical safety implies the safe use and application of the medical device.

In our example of the cooling pack, the medical device is able to cool down a specific body area of the patient by its therapeutic effect. The performance specifications and the application conditions of the cooling pack are decisive.

An analgesic effect precipitated by cold starts from the moment the surface of the skin reaches a temperature of 13.6 °C (Bugaj 1975). After the application, cooling packs filled with ice need 20 min to cool the skin surface to 10.2 °C (Bitton et al. 2016).

With these data, the clinical performance and benefit of the medical device can be shown. The literature reveals that by means of cooling down the skin and then reducing the pain. This is the cooling pack's clinical benefit on the patient's skin.

PMS data, data collected from regional safety databases like the BfArM, Swissmedic, or MHRA or data from international databases like the FDA MAUDE database and also the clinical data in related scientific and evident publications, all give insight into the safety of medical devices. In the clinical evaluation, safety databases are used to analyze reportable events, incidents, and recalls related to the medical device. Furthermore, recalls and complaints from the internal manufacturer's PMS data are used for the evaluation of clinical safety. The number of recalls, reportable events, and incidents within a fixed period of time compared to the sales figures of the medical device shows the ratio between events and number of sold devices: if the number of reportable events, incidents, recalls, etc. is low versus a high number of sold devices, the safety of the respective medical device can be assumed.

The objective of the clinical evaluation is the comparison of the residual risk(s) resulting from the risk analysis and the clinical benefit of the medical device. The benefits and risks of a medical device can only be assessed in interrelation by "weighing any benefit to health from the use of the device against any probable risk of injury or illness from such use" (Guidance for Industry and Food and Drug Administration Staff 2012). Hence, they are to be understood as relative terms: a balanced consideration based on valid scientific evidence in making risk and benefit determinations, including the critical issue of identifying benefits and residual risks, is essential.

Based on the findings from literature, clinical data, as well as risk analysis, it can be inferred that the probability of a patient experiencing a substantial benefit when using the evaluated medical device outweighs the probability of suffering harm due to a residual risk of the device significantly.

Clinical Trials with Medical Devices

A clinical trial also called clinical investigation is defined as "any systematic investigation or study in or on one or more human subjects, undertaken to assess the safety and/or performance of a medical device" (SG5/N1:2007).

Conducting a clinical trial is a scientific process that represents one method of generating clinical data. A clinical trial can be conducted with the medical device in the preclinical stage before the CE marking and marketing of the device if clinical data to the medical device are insufficient to show its clinical safety,

performance, and benefit. To achieve CE approval for specific medical devices, such as long-term invasive applications or implants, clinical trials must be conducted unless the use of existing clinical data can be sufficiently justified. During the pre-approval phase of a medical device and depending on the goal of the clinical trial, various study designs could be useful, ranging from a single-center non-randomized trial to a multicenter, multinational, randomized, controlled trial.

Furthermore, clinical trials can be conducted during the post-market phase of the medical device during the post-market surveillance process. In this case, further clinical data pertaining to the medical device need to be collected to further prove its clinical safety, performance, and benefit or new claims in this context. These clinical trials are conducted as so-called post-market clinical follow-up trials and can include clinical trials according to ISO 14155 or non-interventional studies (NIS) in form of observational studies. Which form or way of collecting further clinical data is chosen belongs on the result of the clinical evaluation indicating which further clinical data is needed.

The overall objective of a clinical trial is the evaluation of the clinical safety and clinical performance of the medical device in question and whether the device is suitable for the purpose(s), user(s), and the population(s) for which it is intended within the claimed indications (ISO 14155).

ISO 14155-1:2009 Clinical Investigation of Medical Devices for Human Subjects – General Requirements details the general requirements for the conduct of clinical trials, and ISO 14155-2:2009 Clinical Investigation of Medical Devices for Human Subjects – Clinical Investigation Plan contains detailed information about the procedure and contents of a clinical trial plan. Clinical trials (also referred to as clinical studies whereas this term rather belongs to drugs than to medical devices) must take into account the respective scientific principles underlying the collection of clinical data along with accepted ethical standards for the use of human subjects. The clinical trial objectives and design are documented in a clinical trial plan.

The following table reveals the differences between clinical studies with pharmaceuticals and medicines (drugs) and clinical trials with medical devices (Chittester 2014) (Table 2):

ISO 14155 also defines the good clinical practice (GCP) standard for clinical trials with medical device. European legislation also explicitly requires compliance with the Declaration of Helsinki defining the ethical principles to be respected when conducting clinical studies or trials on human subjects.

As a rule, all clinical trials need to be approved by Ethics Committees and notified to the competent authorities of involved countries. Other regulatory institutions may need to be involved in the

Table 2 Clinical studies with drugs and clinical trials with medical devices

Drugs	Medical devices
Phase 1:	Pilot:
Aimed at safety and tolerance	Smaller population with disease or condition (10–30 subjects)
Healthy volunteers (20–100 subjects)	Determine preliminary safety and performance information
Determine dosing and major adverse effects	
Phase 2:	Pivotal:
Aimed at safety and effectiveness	Larger population with disease or condition (150–300 subjects)
Small population with disease or condition (50–200 subjects)	Determine effectiveness and adverse effects
Confirm dosing and major adverse effects	
Phase 3:	
Aimed at safety and effectiveness	
Large population with disease or condition (hundreds to thousands of subjects)	
Determine drug–drug interactions and minor adverse effects	
Phase 4:	Post-approval study:
Post approval study	Collect long-term data and adverse effects
Collect long-term data and adverse effects	

regulatory process depending on national law. Only in case of NIS, notification to the competent authority is sufficient, and an Ethics Committee approval is not required.

The essential documents for a clinical trial with a medical device are similar to the ones required for a pharmaceutical study. The term Clinical Investigation Plan or Clinical Study Plan is generally used to refer to the clinical study protocol in case of a clinical trial with a medical device. Inclusion of a section on risk management in the Clinical Investigation Plan is required. Also the inclusion of preclinical data on the medical device is necessary.

Regulatory requirements for clinical trials with medical devices are different to pharmaceuticals, and this has an impact on the design of their clinical trial (www.hra.nhs.uk). There is no legal requirement to demonstrate the efficacy of the device to obtain CE marking. The objective of the clinical trial is to demonstrate the safety and performance (conformity with claims which is the clinical benefit) of a medical device. In a pharmaceutical study, the objective is to demonstrate the safety and efficacy of the medicinal product. One consequence is that case numbers in a medical device trial are usually lower than in pharmaceutical studies. The stage of a clinical investigation which needs to be satisfactorily completed for CE marking may therefore be likened to Phase II in drug development, where evidence of clinical activity of a drug is sought, rather than Phase III. Since efficacy does not need to be demonstrated, randomized controlled trial designs for medical devices are rarely necessary, and, therefore, proof of statistical significance may not be necessary. Interim analysis of study data may be feasible, provided it has been written into the investigation plan.

In comparative pharmaceutical studies, the most robust comparator is a placebo control, which is often applied and generally required by authorities. In a medical device trial, a placebo control is usually not feasible. This is particularly the case with implantable devices, where placebo control groups are simply not possible. However, studies comparing a medical device with standard therapy are possible, although in some cases there may be no standard therapy available which is similar enough to warrant comparison, especially for innovative devices. In addition, the user (usually a healthcare professional) often cannot be blinded to the study intervention.

A specific feature of a clinical trial with a medical device is that product performance may be influenced by the user. Furthermore, the use of a medical device may sometimes be associated with a learning curve for the user, where the outcomes improve with experience.

Another aspect is that adverse events, in particular adverse device effects, may not only concern the study subjects but also third parties, such as users of the device. In contrast, adverse events in pharmaceutical studies are only assessed and monitored for the clinical study subjects.

Due to the wide range of types of device, testing methodologies vary widely. Some performance data might simply require user handling feedback; other data might be more analytical. Medical devices often create large amounts of data that are transmitted, processed, and stored via specific software interfaces. For such data sets, specific monitoring rules have to be established focusing on supervising data processing rather than individual data points.

Conclusion

Medical devices are subject to regulatory requirements in order to ensure their safety and performance. The relation between the clinical benefit experienced by the patient and the residual risk of the medical device is most important. In order to evaluate the benefit and risk of a medical device, several evaluations and appraisals claimed by the European guidelines are necessary.

The ERs applicable for medical devices (MDD/MDR) result in a uniformly accepted system which is necessary for a safe application of medical devices and enabling international trade.

The form of evaluation of clinical data pertaining to a specific medical device depends on its classification. If the ERs for a medical device are fulfilled, a Class I medical device can be marketed without the involvement of a notified

body, and the clinical evaluation is sufficient. The documentation and technical file of Class IIa, IIb, or III medical devices needs to be assessed and approved by a NB. Compliance with applicable standards is essentially required. Each medical device must fulfill these requirements independent on its classification.

The clinical evaluation of medical devices evaluates the clinical safety and performance as well as benefit of the medical device and assesses if any residual risks exist. The ratio between benefit and residual risk is evaluated, and the medical device is only assumed to be safe when the benefit overweighs any residual risk. In this context, scientific evident publications, scientific research, clinical trials and investigations, medical knowledge, PMS data, safety reports, and clinical evaluations of the medical device contribute to the compliance with the ERs. In case of specific (high-risk) medical devices or if no clinical data is available for the evaluated medical device, the manufacturer needs to collect own clinical data in form of clinical trials or later in form of PMCF activities.

References and Further Reading

Bitton E, Lacroix Z, Leger S (2016) In-vivo heat retention comparison of eyelid warming masks. Cont Lens Anterior Eye 39(4):311–315

Bugaj R (1975) The cooling, analgesic, and rewarming effects of ice massage on localized skin. Phys Ther 55:11–19

Chittester B (2014) Medical device clinical trials – how do they compare with drug trials? https://cdn2.hubspot.net/hub/149400/file-646377456-pdf/docs/mc-n-med-dev-trials-compare-with-drug-trials.pdf. Accessed 21 Jan 2018

Council of the European Communities (2007) Council Directive of 14th June 1993 concerning medical devices. 93/42/EEC

Cramer H, Lauche R, Hohmann C, Choi KE, Rampp T, Musial F, Langhorst J, Dobos G (2011) Randomized controlled trial of pulsating cupping (pneumatic pulsation therapy) for chronic neck pain. Forsch Komplementmed 18(6):327–334. https://doi.org/10.1159/000335294

French-Mowat E, Burnett J (2012) How are medical devices regulated in the European Union? J R Soc Med 105(Suppl. 1):S22–S28

Guidance for Industry and Food and Drug Administration Staff (2012) Factors to consider when making benefit–risk determinations in medical device premarket approval and de novo classifications

IEC 62366 (2018) Medical devices – application of usability engineering to medical devices

In Vitro Diagnostic Regulation (IVDR) (2018)

ISO 14155 (2018) Clinical investigation of medical devices for human subjects – good clinical practice

ISO 14971 (2018) Risk management procedures for medical devices

Kühn K-D, Renz N, Trampuz A (2017) Lokale Antibiotikatherapie. Der Unfallchirug 120(7):561–572

Larsson R, Oberg P, Larsson S (1999) Changes of trapezius muscle blood flow and electromyography in chronic neck pain due to trapezius myalgia. Pain 79(1):45–50

Lauche R, Cramer H, Langhorst J, Dobos G (2013) Cupping for chronic nonspecific neck pain: a 2-year follow-up. Forsch Komplementmed 20(5):328–333. https://doi.org/10.1159/000355634

Li T, Li Y, Lin Y, Li K (2017) Significant and sustaining elevation of blood oxygen induced by Chinese cupping therapy as assessed by near-infrared spectroscopy. Biomed Opt Express 8(1):223–229. https://doi.org/10.1364/BOE.8.000223

Medical Device Directive (MDD) 93/42/EEC (2018)

Medical Device Regulation (MDR) 2017/745 (2018)

Teut M, Kaiser S, Ortiz M, Roll S, Binting S, Willich SN, Brinkhaus B (2012) Pulsatile dry cupping in patients with osteoarthritis of the knee – a randomized controlled exploratory trial. BMC Complement Altern Med 12:184. https://doi.org/10.1186/1472-6882-12-184

Food Supplements: Definition and Classification

33

Evelyn Breitweg-Lehmann, Birgit Liebscher, and Carolin Bendadani

Contents

Abstract

Food Supplements are de jure foodstuffs the purpose of which is to supplement the normal diet. According to Directive 2002/46/EC, food supplements are concentrated sources of nutrients or other substances with a nutritional or physiological effect, alone or in combination, marketed in dose form, designed to be taken in measured small unit quantities. The labelling, presentation, and advertising of food supplements shall not include any mention stating or implying that a balanced and varied diet cannot provide appropriate quantities of nutrients in general. Accordingly, the labelling, presentation, and advertising must not attribute to food supplements the property of preventing, treating, or curing a human disease or refer to such properties. Preventing, treating, or curing human disease is subject to medicinal products.

E. Breitweg-Lehmann (✉) · B. Liebscher · C. Bendadani
Federal Office of Consumer Protection and Food Safety, Berlin, Germany
e-mail: evelyn.breitweg-lehmann@bvl.bund.de; Birgit.Liebscher@bvl.bund.de; Carolin.Bendadani@bvl.bund.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_67

As they are usually marketed in dose forms as well and in several cases contain low-dosed drugs, delimitation questions occur frequently.

Hence the issue, whether some of these products are to be classified as medicinal products and how they are to be differentiated.

Subsequently, we would like to introduce the basic requirements to put those food supplements legally on the market.

Background

In 2002, the European Union re-organized in Europe food law and harmonized nearly all relevant provisions. Until then the use of food supplements was not very widespread in Germany. Products containing small batches of concentrates, such as capsules or tablets were regarded as typically medicinal products, regardless if their ingredients had pharmacological properties or not.

With the introduction of the new Food Supplement Regulation (Nahrungsergaenzungsmittelverordnung, NemV) (Regulation on dietary supplements) of 2004, which implemented the European Directive 2002/46/EC on Food Supplements into German law, this changed.

To facilitate efficient monitoring of food supplements, Germany and other Member States require the manufacturer or the person placing the product on the market in their territory to notify the competent authority of that placing on the market by forwarding it a model of the label used for the product. Notifications increased from about 3,000 in 2007 to more than 10,000 in 2018 (Fig. 1). The Federal Office of Consumer Protection and Food Safety (Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, BVL) is the competent authority to receive the notifications. BVL sends the notifications to the competent authorities for food control within the German Länder.

Food supplements are foodstuff containing a wide range of different substances such as vitamins, minerals, trace elements, amino acids, fibers, bacteria, fungi, plant extracts, or other substances with a nutritional or physiological effect like Coenzyme Q10, Lutein, Carnitine, or concentrates of nutrients.

Food supplements are mainly placed on the market in tablet or capsule form. Their purpose is to supplement the normal diet, thereby unfolding a nutritional or physiological effect, which is supposed to maintain health.

The prevention (prophylaxis), treating, or healing of human diseases is strictly reserved to medicinal products. Medicinal products and foodstuffs exclude each other mutually.

Definition of Food Supplements

Food supplements are legally defined in Germany in the NemV (Regulation on dietary supplements) based on Directive 2002/46/EC. According to § 1 NemV and Article 2 of the Directive, the expression "food supplement" means foodstuffs in compliance with the following conditions:

1. They supplement the normal diet.
2. They contain concentrated sources of nutrients or other substances with a nutritional or physiological effect, either on their own or in combination with other substances.
3. They are marketed in dose form as capsules, pastilles, tablets, pills, powder sachets, drop-dispensing bottles, and other ways of presenting liquids and powders designed to be taken in measured doses.

General Food Law Provisions

Hence, a food supplement is food. Therefore, basically all legal regulations applying to food also apply to food supplements. These are mainly the Regulation (EC) No 178/2002 (Basic Regulation) together with the general rules for the production and marketing of food.

The liability for complying of the products to be offered on the market to the food law provisions lies – as it is normal in food regulations – completely in the hand of the business operator according to Articles 17 (1) and 19 of the Basic Regulation.

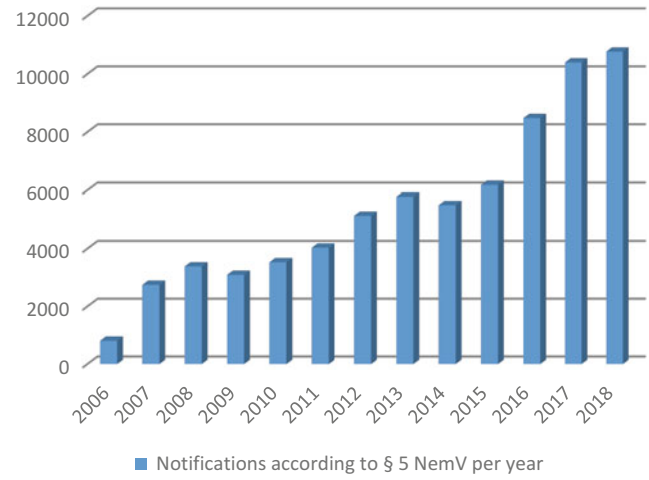


Fig. 1 Notifications of food supplements in Germany per year

Table 1 Food supplements versus medicinal products. Modified according to (Federal Office of Consumer Protection and Food Safety 2019)

Food supplements	Medicinal products
According to § 1 NemV they are foodstuffs that are intended to supplement the general diet of healthy persons.	Pursuant to § 2 (1) of the AMG, they are intended for use in or on human body or attribute the property of preventing, treating or curing a human disease.
General food law is applicable (LFGB, NemV, Novel Food Regulation, Basis Regulation on food law (EC) Nr. 178/2002)	Pharmaceutical legislation is applicable (Medicines Act, pharmacopoeias etc.).
To be notified at the BVL according to § 5 NemV.	Authorization procedure in Germany by the Federal Institute for Drugs and Medical Devices (BfArM), the Paul Ehrlich Institute (PEI) or by the European Medicines Authority.
No preliminary regulatory proof of efficacy or safety is required. The business operator is responsible for the safety of the product. The federal states are the competent authorities responsible for monitoring compliance of the products with the regulations of food law and especially the NemV („Control of the self-inspection of the manufacturer and the distributor“).	Manufacturers must submit clinical trials during the authorization process that demonstrate the efficacy and safety of the drugs.
Quantities labelled on the packaging may vary by up to 50% from the actual quantity in the product.	Quantities labelled on the packaging must not vary by more than 5% from the actual dosage of the active ingredients.
Usually, no specific provisions and no maximum levels are set for vitamins, minerals and other substances.	The doses of all ingredients are tested and specified within the scope of the authorization procedure.

Furthermore, the national regulations based on the German Food and Feed Law (LFGB) are applicable to food supplements (Food, Commodities and Feed Code). Control and compliance to the appropriate regulations are also due to the competent national authorities for food control, in Germany the federal states (Table 1).

General labelling requirements are stated in the Regulation (EU) No 1169/2011 on the provision of food information to consumers , which also applies to food supplements. In addition, the Health Claims Regulation (EC) No 1924/2006 (HCV) also applies for health and nutrition claims on the label of food supplements. This regulation also includes a ban of unauthorized health claims for food supplements according to the Health Claim Regulation in advertisements.

Also applicable are, e.g.,

- Regulation (EC) No 1333/2008 on food additives for technological purposes
- Regulation (EC) No 1332/2008 on food enzymes
- Regulation (EC) No 1334/2008 on flavorings)
- Regulation (EU) 2015/2283 on novel foods
- Commission Regulation (EC) No 1881/2006 setting maximum levels for certain contaminants in foodstuffs
- Regulation (EC) No 396/2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin

Special-Law Provisions for Food Supplements

The abovementioned Food Supplements Directive and the German NemV (Regulation on dietary supplements) contain special-law provisions for food supplements. Special requirements for labelling of food supplements and the notification procedure are described there. According to § 5 NemV, food supplements must be notified to the BVL by the producer or distributor before the first placing on the market at the latest, together with a sample of the label to be used on the product. The BVL does not verify the received notification but forwards it to the competent authorities in the federal states responsible for food control. The competent authorities for food control make random checks of the notified products. During the notification procedure, the BVL only informs the competent food control authorities if new food supplements are put on the market so they may take action if there is a suspicion that a particular product does not comply with the legal regulations. So, there is neither an approval nor permission before a product is put on the market. As with other foods, food supplements can be marketed if they comply with all food law regulations (Noble 2017).

Therefore, it is primarily up to the food business operators to ensure the legal compliance and safety of their food supplements. This means, consumers can contact the competent authorities for food control should there be any doubts about the safety of certain food supplements.

Clinical Trials

Unlike the specific provisions for clinical trials of medicinal products, no specific regulations exist for the clinical trials of foodstuffs. Nevertheless, the Declaration of Helsinki as a statement of ethical principles for medical research involving human subjects is also applicable to food product studies.

The procedures are based on the "Guidelines for Good Clinical Practice" (ICH GCP) being the international ethical and scientific quality standards for designing, conduction, recording, and reporting of clinical trials on humans defining a clinical trial as "any investigation in human subjects intended to discover or verify the clinical, pharmacological and/or other pharmaco-dynamic effects of an investigational product(s)." Even food can be labelled as a preparation to be tested.

Unlike a clinical trial with medicinal products, a trial with foodstuffs does not require an authorization by the Federal Agencies. But as mentioned before, in Germany food supplements will need a notification before being put on the market according to § 5 NemV.

Basically, the procedure of studies with healthy volunteers is already considered as a "placing on the market" in the sense of regulation (EC) No 178/2002 according to Article 3, No. 8 ("...any other form of transfer, whether free of charge or not...").

How Is It Determined if a Particular Ingredient in a Food Supplement Can Be Placed on the Market?

For some substance classes such as food additives, flavors, or contaminants, the question is easy to answer. Their use is regulated specifically; it has clearly been stated in the respective EU regulations which quantities are permitted in which food. They have annexes of so-called positive or negative lists, in which the permissible quantities in food are described in detail. If the provisions of these regulations are not adhered to, the food and thus the food supplement may not be placed on the market.

Also, the permitted vitamin and mineral compounds that may be used in food supplements are standardized in Europe by Directive 2002/46/EC on Food Supplements, to which § 3 of the NemV (Regulation on dietary supplements) refers. These substance classifications clearly mark whether their use is legal or not. So far, different models for the quantitative derivation of maximum levels for vitamins and minerals have been developed and discussed in Germany and other European countries. However, to date no binding maximum levels for vitamins and minerals in food supplements have been set at EU level. In Germany, the Federal Institute for Risk Assessment (BfR) has developed proposals for maximum levels for vitamins and minerals in food supplements (Weißenborn et al. 2018). Although the recommendations are not legally binding, they provide a very useful orientation for food business operators and the competent authorities.

For all other substances than vitamins and minerals in food supplements listed in the NemV (Regulation on dietary supplements) as "other substances with a nutritional or physiological effect" and in Article 8 in conjunction with Article 2 (2) of the Food Enrichment Regulation (EC) 1925/2006 referred to as "substances other than vitamins and minerals with a nutritional or physiological effect," so far no EU-specific legal regulations exist. Therefore, the food business operator needs to check the ingredients of the food supplements in every individual case whether they comply with food law.

However, very often there is little scientific literature on these substances available to allow a well-defined classification. For this reason a test scheme has been developed by the BVL examining substances in food supplements according to the exclusion principle (Breitweg-Lehman 2017) (Fig. 2).

This procedure is explained in the following diagram:

Demarcation to Medicinal Products – It Must be Ensured that the Product Is Not a Medicinal Product in the Sense of § 2 (1) Medicines Act

First of all food supplements are to be clearly distinguished from medicinal products, which do not count as food according to article 2,sentence 4 d) of the Food Basic Regulation (EC) No 178/2002 (Regulation (EC) No 178/2002).

The manufacturer has to ensure that its food supplement does not contain any pharmacologically active substances according to § 2 (1) No. 2a) of the Medicines Act in Germany (Arzneimittelgesetz, AMG). It defines medicinal products as substances or preparations of substances

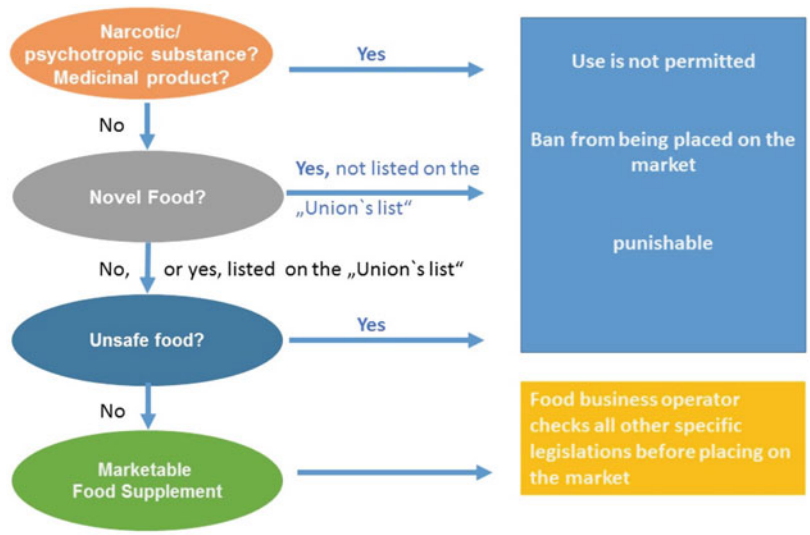


Fig. 2 Testing scheme on marketability of substances in food supplements according to the exclusion principle

- To be used in or on the human (or animal) body in order to heal, palliate, or prevent human or animal diseases or pathological complaints (known as medicinal products by presentation, because they are seen as such by the consumer solely by reason of their presentation, for example, their packaging or packaging leaflet),
- To be used in or on the human (or animal) body in order to
  - Restore, correct, or influence the physiological functions by means of a pharmacological, immunological, or metabolic effect
  - Or to make a medicinal diagnosis

Consequently, for the distinguishing of these products, the intended purpose is a relevant criterion.

While medicinal products are used to cure, alleviate, or prevent human diseases (also referred to as "drugs by presentation"), food supplements are designed to complement the diet of normal consumers.

For this, the predominant objective purpose is decisive. This means, how does the product appear to an average informed, attentive, and responsive consumer (Judgment of the German Federal Court 2008), which requires the objective decision, if this predominantly intended for use by physicians (therapeutically, prophylactically, etc.) or for food-specific purposes (supplementation of the general diet) (Kügel et al. 2016).

For this the package instructions, the enclosed instruction leaflets and the advertising statements are to be used.

If it is not a medicinal product by presentation, § 1 NemV (Regulation on dietary supplements) demands three specific characteristics in a food supplement, which include general dietary supplement, a concentrate of nutrients or other substances with nutrition-specific or physiological effect, and dose form.

The dose form (capsules, pastilles, tablets, etc.) as a characteristic has none to only little distinguishing effects as this also applies to most medicinal products. But it is crucial that the concentrate of nutrients or other substances has a predominantly nutrition-specific or nutritional-physiological effect and supplements the general diet.

The term "nutrition-specific" in this context covers only clearly defined "classic nutrient functions" as defined by "classic nutrients" in the sense of Article 2 (2) No. 2 of the Regulation on Health Claims (Regulation (EC) No 1924/2006) and Article 2 (2) (s) of the Regulation (EU) No 1169/2011 on the provision of food information to consumers, such as protein, carbohydrates, fat, fiber, sodium, vitamins, and minerals (Regulation on dietary supplements).

On the other hand, a "physiological effect" in the context of nutrition/nutritional supplementation exists if a substance or a concentrate of nutrients performs a specific function in the body for the maintenance of the physiological functions.

Should the addition of a substance cause functions beyond the physiological fluctuation range it can no longer be considered a "nutritional" or "physiological" effect and can no longer legally be called a food supplement.

A functional drug is defined by the existence of a pharmacological, metabolic, or immunological effect according to § 2 (1) No. 2a) of the German Medicines Act (AMG). The pharmacological effect is the essential element for the classification of functional drugs.

The term "pharmacological effect" is not legally defined. For the interpretation of the term "pharmacological effect," it is crucial that it "has a significant effect on the metabolism and thus really influence its functional conditions (...)."

In the judgment of the Lower Saxony Higher Administrative Court (OVG Lueneburg) (case 13 LB 31/14 dated 02.11.2017), the criteria for the existence of a pharmacological effect were systematically derived and can be adopted for corresponding issues.

According to this decision, a pharmacological effect exists when a product can influence

- The human physiological functions
- In a significant way
- By targeted control
- Restore, correct, or otherwise influence (Judgment of the OVG 2017), in which case

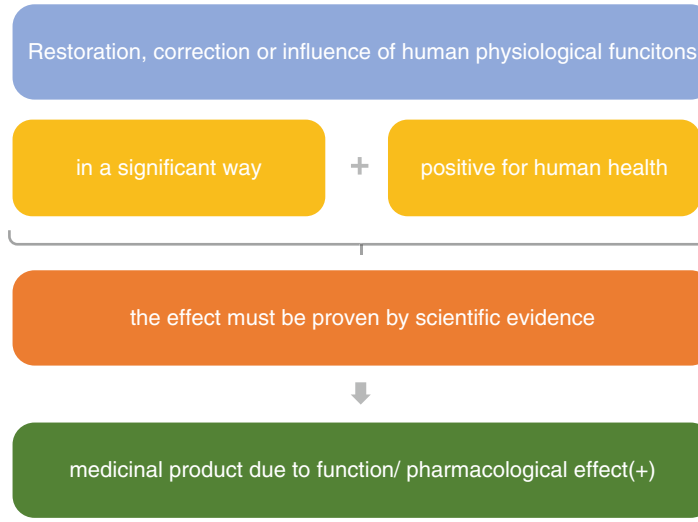


Fig. 3 Criteria of the German jurisdiction proving a pharmacological effect

the intended normal use of the product is relevant for the assessment (Fig. 3).

In a previous judgment of the European Court of Justice (ECJ) (Commission v Germany, case C-319/05 dated 15.11.2007) concerning a "garlic extract powder capsule," an important principle was stated. A substance whose effect on physiological functions is not greater than, or different from the effect, which a foodstuff taken as a part of the daily diet – consumed in a reasonable quantity – may have on those functions, does not have a significant effect on the metabolism and cannot be classified as a medicinal product. Since risks and contra-indications related to taking garlic preparations are limited and, more importantly, are not different from garlic ingested as a foodstuff, and because the criterion of the method of using the product concerned cannot be decisive, given that capsule form is not unique to medicinal products, such a preparation cannot be classified as a medicinal product by function (Judgment of the Court (First Chamber) 2007).

That means a significant effect on the physiological functions only exists if the effect of the substance goes beyond that can be achieved by consuming a foodstuff containing appropriate amounts of the same substance.

On the other hand, the criterion of restoring or correcting the course of normal life processes does not presuppose that a norm or standard range exists for the relevant physiological function. In many cases, a standard range was never defined although an effect is clearly measurable. According to the abovementioned judgment of the Lower Saxony Higher Administrative Court, it is sufficient that the product influences human physiological functions in such a way that it has a positive effect on human health (Judgment of the Court (Fourth Chamber) 2014; Judgment of the Court (Fifth Chamber) 2009). Otherwise, products with a significant effect on the human physiological functions would be excluded from the classification as medicinal products until a norm or standard range is set. This would confine the development of new medicines dramatically as in many cases those standards simply not exist or the exact mode of action is not entirely clear.

In addition, therapeutic efficacy is not a mandatory requirement for a pharmacological effect. If the therapeutic efficacy is lacking, it may nevertheless be classified as a functional drug.

Although therapeutic efficacy is an approval criterion for authorization of medicinal products, it cannot lead to the conclusion that products with a significant pharmacological effect on human physiological functions are to be considered as food until proof of therapeutic efficacy has been obtained and were therefore taken out of the drug regulatory framework.

The above-described pharmacological effect must be scientifically substantiated and appropriate evidence must be provided (Judgment of the OVG 2017). This is a mandatory criterion for the classification of a medicinal product.

The Possibility of Being a Novel Food Supplement Must Be Ruled Out

As food supplements are foodstuffs, the Novel Food Regulation (EU) 2015/2283 is also applicable. "Novel food" in the sense of the Regulation (EU) 2015/2283 defines any type of food that has not been consumed to a significant degree by humans in the European Union before 15 May 1997 and has not been listed in at least one of the categories defined in Article 3 of Regulation (EU) 2015/2283.

As a first indication for defining the novel food status the European Commission's Novel Food status catalogue may be consulted (The European Commission's 2019) which is publically available (https://ec.europa.eu/food/safety/novel_food/catalogue_en). In this catalogue, the member states have matched different substances and this catalogue represents the up-to-date status of the listed substances, whether they are considered novel or not.

However, it is not sufficient if only the special ingredient has been consumed to a significant degree before 15 May 1997. Relevant is the actually used extract in the final food product.

Already, the ECJ confirmed in a judgment of 15 January 2009 (ECJ document C38307 C-383/07) that it is not enough if each one of the ingredients of a food has been used for human consumption before 15 May 1997 to a significant degree in the territory of the European Community but that the final food product fulfils this condition. (see also BayVGH, Urt. v. 12.05.2009 case: B 09.199 9 B 09.199). Here, too, it depends on a specific case-by-case decision with regard to the specific product.

Novel food and novel food ingredients already approved are published in the so-called Union's list (Regulation (EU) 2017/2470 according to Article 10 of the Novel Food Regulation. It also contains the conditions of use of the respective food.

If the substance in question is not included there and cannot be proved to have been consumed to a significant degree in the EU before 15 May 1997, it is considered to be "novel." In this case placing on the market is only possible after prior approval. The procedure is regulated in the Novel Food Regulation. Responsible for the approval is the European Commission, which consults the member states. Upon completion of the procedure, the substance/food with the specific conditions of use may be included in the Union's list of the Implementing Regulation (EU) 2017/2470.

However, the Novel Food Regulation contains special rules for food supplements:

According to Article 3(2)a) x) of the Novel Food Regulation food that has solely been used in food supplements before 15 May 1997 and is supposed to be used only in food supplements afterward is considered as "not novel." If it is to be used in other food products, they would still be novel.

Anyway, the previously used term "NFS" (not novel in food supplements) has been legally codified and acknowledges a proven usage before 15 May 1997 only in food supplements as a separate category, which cannot be applied to other foods.

Food Supplements Must Be Safe – Unsafety Must Be Ruled Out

If the classification as a medicinal product and the Novel Food – status of the food or ingredients – can be ruled out, the safety aspects of the whole product remains to be assessed. According to Article 14 (1) of the Regulation (EC) No 178/2002, food shall not be placed on the market if it is unsafe. According to Article 14 (2) are those foods to be considered unsafe that are a considered to be injurious to health or unfit for human consumption.

Hereby, the normal conditions of use of the food by the consumer and to the information provided to the consumer shall be regarded, including information on the label, or other information generally available to the consumer.

In addition, any short, mid-term, and long-term effects of any food intake, possible cumulative

toxic effects, and the particular health sensitivities of certain consumer categories are to be considered when assessing health risks.

Infringements of EU food and feed legislation posing a direct or indirect risk to human health are reported through the Rapid Alert System for Food and Feed (RASFF) according to Article 50 of Regulation (EC) No 178/2002. It is a network of contact points in all member states and the European Commission managed by the latter who maintains a database with information on individual RASFF notifications. With this information, the member states are able to take immediate action, such as warning consumers or withdrawing products from the market.

During the recent years, several notifications and warnings concerning unsafe food supplements have been processed through the RASFF-system (The European Commission's priorities 2019).

In the most problematic cases, even (hepato-) toxic or pharmacologically active substances without any labelling, such as the appetite suppressant Sibutramine, have been found in products labelled as food supplements and were notified in the RASFF system.

In general, food business operators are responsible for ensuring that their products are safe. As soon as any competent authority is to become active, according to § 39 LFGB (Food, Commodities and Feed Code) in the national German law due to the assumption that the food is unsafe, the competent authority has to provide sufficient scientific evidence that this is the case, This means a reversal of the burden of proof from the food business operator to the competent authority of the European Member State. This may cause considerable difficulties in individual cases, because even if this regulation appears clear and unambiguous enough, it may be difficult to prove that a product is unsafe, e.g., due to overdosed plant extracts that may lead to undesirable long-term effects when taken.

Therefore, classification and further harmonization at EU level of further substances with a nutritional or physiological effect would be desirable. The legal foundations for the regulation of these substances have already been laid down in Regulation (EC) No 1925/2006.

Article 8 of this regulation provides procedures to address safety concerns of any substances other than vitamins and minerals, which are used for producing food supplements. The procedure provided for in this Article shall be followed where a substance other than vitamins or minerals, or an ingredient containing a substance other than vitamins or minerals, is added to foods or used in the manufacture of foods under conditions that would result in the ingestion of amounts of this substance greatly exceeding those reasonably expected to be ingested under normal conditions of consumption of a balanced and varied diet and/or would otherwise represent a potential risk to consumers.

On its own initiative or based on information provided by Member States, the European Commission may initiate the procedure in order to include a certain substance in a list to prohibit or restrict its use.

Annex III of Regulation (EC) No 1925/2006 contains three lists for this purpose, if a harmful effect on health has been identified:

- (1) Part A – Prohibited substances: addition to foods or its use in the manufacture of foods shall be prohibited.
- (2) Part B – Restricted substances: addition to foods or its use in the manufacture of foods shall only be allowed under the conditions specified therein.
- (3) Part C – Substances under Community scrutiny: if the possibility of harmful effects on health is identified but scientific uncertainty persists.

If a substance is put under Community scrutiny (Part C of Annex III), interested parties may at any time submit scientific data demonstrating the safety of a listed substance under the conditions of its use in a food or a category of foods and explaining the purpose of that use. Within four years of a substance being listed in Part C, the European Commission must decide, based on data provided by interested parties, to generally allow the use of the substance, or to list it in Part A or B of Annex III. Until the final decision is taken, existing national provisions are applicable.

Up to date, only three substances have been listed in Annex III, two of them following a request from Germany. According to Article 8, EFSA has published scientific opinions for the use of Ephedra species and Yohimbe (Pausinystalia yohimbe (K. Schum.) Pierre ex Beille). Based on this, the European Commission has decided to add both substances to the list of prohibited substances (list A) (Commission Regulation (EU) 2019/650 and Commission Regulation (EU) 2015/403. In 2019, trans fat other than trans fat naturally occurring in fat of animal origin was added to the list of restricted substances (Part B) (Commission Regulation (EU) 2019/649).

The Commission has initiated the procedure – by its own initiative – according to Article 8 of the Regulation (EC) No 1925/2006 for the intake of Monacolin K from red yeast rice. For botanicals containing Hydroxyanthracene derivatives and green tea catechines, concerns have been raised by the EU Member States regarding a potential risk to consumers when these substances are consumed. Decisions are still pending (The European Commission's priorities 2019).

Guidelines on the Assessment of Food Supplements

As there are currently no further bans or restrictions for "other substances" planned on EU level, except for a few substances, in member states national laws for the classification of substances used in food supplements still apply.

For a more systematic approach, Competent Federal Government and Federal State Authorities of Germany published a list of substances in 2014 including the classification of plants and plant parts. The classification is based on a decision tree worked out by a working group which consisted of representatives of the Federal Institute for Risk Assessment (BfR), the Federal Institute for Drugs and Medical Devices (BfArM), the Federal Office of Consumer Protection and Food Safety (BVL), and the Federal State Authorities (Chair: BVL) The list contains a classification of about 620 plants and plant parts (BVL 2014) including their novel food status. The classification of the plants assessed is based on the same principles applying to Article 8 Annex 3 of Regulation (EC) No 1925/2006. All substances posing a potential risk to the consumer are classified in three sub lists, when applicable:

List A: Substances not recommended for use in foods
List B: Substances for which restricted use in foods is recommended
List C: Substances for which assessment cannot be completed, yet

The classification in one of these categories is based on a decision tree.

The list is applicable for food supplements and fortified foods. The list is not legally binding but provides a very useful orientation for food business operators and the competent authorities. The list is on the website of the BVL: www.bvl.bund.de/stoffliste_eng. In the last 2 years, the list has been updated by an enlarged working group consisting of the abovementioned institutions and in cooperation with representatives from Austria and Switzerland. Furthermore, lists for fungi and algae have been created in the framework of this cooperation project, which will be published soon (Bendadani 2019).

To support the federal governmental and state authorities on queries regarding the evaluation and classification, the BVL and the BfArM have initiated in 2013 a "Joint Expert Commission for the Classification of Substances" (BVL 2016). The aim of this common expert commission as an approved, independent panel is to work out criteria catalogues, decision trees and written opinions for the classification of substances used in foods and food supplements. Topics include, for example, the identification of possible health risks or the classification as to whether a substance is qualified for the authorization as a medicinal product.

Further orientation support on the classification of comparable substances can be taken from decisions of national courts as well as statements of the working group of food chemistry experts from the federal states and the Federal Office of Consumer Protection and Food Safety (ALS - Working Group of food 2019).

Conclusion

The regulations mentioned in this chapter provide a legal framework for the placing of food supplements on the market. The rules for the labelling of food supplements and the most substances in food supplements, such as additives, flavorings, residues, or contaminants, have been harmonized – they are subject to specific regulations in the EU, such as maximum and minimum levels for specific food categories. For other substances with nutritional or physiological effects, the harmonization at EU level is still pending. It is desirable to increase the legal certainty for food business operators and the competent authorities for food control. For the time being, national laws still apply. In Germany, the "List of Substances of the Competent Federal Government and Federal State Authorities," the statements of the "Joint Expert Commission for the Classification of Substances" and of the "Working group of food chemistry experts from the federal states and the Federal Office of Consumer Protection and Food Safety" can serve as guidelines for food producers and traders and the competent authorities. In case of doubt, the courts have to decide.

References and Further Reading

Bendadani C (2019) Die Stoffliste des Bundes und der Bundesländer – Was gibt es Neues?. Food & Recht Praxis 02/2019, pp 13–17

Breitweg-Lehman E (2017) Kriterien zur Beurteilung von Nahrungsergaenzungsmitteln. Bundesgesundheitsblatt 60:277–285

BVL (2014) Stoffliste des Bundes und der Bundesländer – Kategorie „Pflanzen und Pflanzenteile". BVL Reporte, Bd. 8.4. Springer, Basel. http://www.bvl.bund.de/SharedDocs/Downloads/01_Lebensmittel/stoffliste/stoffliste_pflanzen_pflanzenteile.pdf?__blob=publicationFile&v=5. Last access 18 Jan 2019

BVL (2016) Expert commission for the classification of substances (outline of the commission). http://www.bvl.bund.de/DE/01_Lebensmittel/01_Aufgaben/07_Expertenkommission/lm_expertenkommission_node.html. Last access 18 Jan 2019

Commission Implementing Regulation (EU) 2017/2470 of 20 December 2017 establishing the Union list of novel foods in accordance with Regulation (EU) 2015/2283 of the European Parliament and of the Council on novel foods (OJ L 351, 30.12.2017, pp 72–201)

Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs (OJ L 364, 20.12.2006, pp 5–24)

Commission Regulation (EU) 2015/403 of 11 March 2015 amending Annex III to Regulation (EC) No 1925/2006 of the European Parliament and of the Council as regards Ephedra species and Yohimbe (Pausinystalia yohimbe (K. Schum) Pierre ex Beille) (OJ L 67, 12.3.2015, pp 4–5)

Commission Regulation (EU) 2019/649 of April 2019 amending Annex III to Regulation (EC) No 1925/2006 of the European Parliament and of the Council as regards trans fat, other than trans fat naturally occurring in fat of animal origin

Commission Regulation (EU) 2019/650 of 24 April 2019 amending Annex III to Regulation (EC) No 1925/2006 of the European Parliament and of the Council as regards Yohimbe (Pausinystalia yohimbe (K. Schum) Pierre ex Beille) (OJ L 110, 25.4.2019, pp 21–22)

Directive 2002/46/EC of the European Parliament and of the Council of 10 June 2002 on the approximation of the laws of the Member States relating to food supplements (OJ L 183, 12.7.2002, pp 51–57)

Drug law (Arzneimittelgesetz – AMG) in the version of the notice of 12 December 2005 (BGBl. I S. 3394), last changed by article 4 (11) of law of 18 July 2016 (BGBl. I S. 1666)

Federal Office of Consumer Protection and Food Safety: Food supplements versus medicines. http://www.bvl.bund.de/DE/01_Lebensmittel/03_Verbraucher/04_NEM/01_NEM_Arzneimittel/NEM_Arzneimittel_node.html. Last access 02 May 2019

Food, Commodities and Feed Code (Lebensmittel- und Futtermittelgesetzbuch – LFGB) in the version of the notice of 03 June 2013 (BGBl. I S. 1426), last changed by article 4 (19) of the Act of 18 July 2016 (BGBl. I S. 1666)

Judgment of the Court (First Chamber) of 15 November 2007, Case C-319/05 (garlic preparation in capsule form)

Judgment of the Court (Fourth Chamber) of 10 July 2014, Case-358/13 (Legal Highs)

Judgment of the Court (Fifth Chamber) of 30 April 2009, Case C-27/08 (incense tablets)

Judgment of the German Federal Supreme Court of 26 June 2008 – I ZR 112/05, NJW-RR 2008, 1255

Judgment of the OVG Lueneburg of 2 November 2017, Case 13 LB 31/14 (Ginkgo biloba)

Kügel, Müller, Hofmann (2016) Arzneimittelgesetz, 2. Auflage: Kommentar, Rn. 164

Noble P (2017) Nahrungsergänzungsmittel Rechtliche Grundlagen, Abgrenzung zu Arzneimitteln, sonstige Fragestellungen. Bundesgesundheitsblatt 60:260–267

Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. (OJ L 31, 1.2.2002, pp 1–24)

Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC (OJ L 70, 16.3.2005, pp 1–16)

Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods (OJ L 404, 30.12.2006, pp 9–25)

Regulation (EC) No 1925/2006 of the European Parliament and of the Council of 20 December 2006 on the addition of vitamins and minerals and of certain other substances to foods (OJ L 404, 30.12.2006, pp 26–38)

Regulation (EC) No 1332/2008 of the European Parliament and of the Council of 16 December 2008 on food enzymes and amending Council Directive 83/417/EEC, Council Regulation (EC) No 1493/1999, Directive 2000/13/EC, Council Directive 2001/112/EC and Regulation (EC) No 258/97 (OJ L 354, 31.12.2008, pp 7–15)

Regulation (EC) No 1333/2008 of the European Parliament and of the Council of 16 December 2008 on food additives (OJ L 354, 31.12.2008, pp 16–33)

Regulation (EC) No 1334/2008 of the European Parliament and of the Council of 16 December 2008 on flavourings and certain food ingredients with flavouring properties for use in and on foods and amending Council Regulation (EEC) No 1601/91, Regulations (EC) No 2232/96 and (EC) No 110/2008 and Directive 2000/13/EC (OJ L 354, 31.12.2008, pp 34–50)

Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004 (OJ L 304, 22.11.2011, pp 18–63)

Regulation (EU) 2015/2283 of the European Parliament and of the Council of 25 November 2015 on novel foods, amending Regulation (EU) No 1169/2011 of the European Parliament and of the Council and repealing Regulation (EC) No 258/97 of the European Parliament and of the Council and Commission Regulation (EC) No 1852/2001 (OJ L 327, 11.12.2015, pp 1–22)

Regulation on dietary supplements (Nahrungsergaenzungsmittelverordnung – NemV) of 24 May 2004 (BGBl. I S.1011), last changed by article 64 of the Regulation of 31 August 2015 (BGBl. I S.1474)

The European Commission's priorities: addition of vitamins and minerals. https://ec.europa.eu/food/safety/labelling_nutrition/vitamins_minerals_en. Last access 20 May 2019

The European Commission's priorities: EU Novel food catalogue. http://ec.europa.eu/food/safety/novel_food/catalogue/search/public/index.cfm. Last access 18 Jan 2019

The European Commission's priorities: reports and publications. https://ec.europa.eu/food/safety/rasff/reports_publications_en. Last access 20 May 2019

Weißenborn A, Bakhiya N, Demuth I et al (2018) Höchstmengen für Vitamine und Mineralstoffe in Nahrungsergänzungsmitteln. J Consum Protect Food Saf 13:25–38

Working Group of food chemistry experts from the federal states and the Federal Office of Consumer Protection and Food Safety (Arbeitskreis Lebensmittelchemischer Sachverstaendiger der Laender und des Bundesamtes für Verbraucherschutz und Lebensmittelsicherheit – ALS). www.bvl.bund.de/DE/01_Lebensmittel/01_Aufgaben/02_AmtlicheLebensmittelueberwachung/12_ALS/lm_ALS_node.html. Last access 18 Jan 2019

Stem Cell Therapy

34

Nina Graffmann, Lucas-Sebastian Spitzhorn, Soraia Martins, Md Shaifur Rahman, Lisa Nguyen, and James Adjaye

Contents

Abstract

Stem cell therapy and regenerative medicine have a tremendous potential for the treatment of a wide variety of currently fatal diseases. So far, routine stem cell based therapies are limited to the treatment of hematologic and dermatologic malignancies, but the field is continuously evolving. One research focus lies on mesenchymal stem cells (MSCs) which have an immunomodulatory potential and support regenerative processes throughout the body. Pluripotent stem cells, on the other hand, can be differentiated into every cell type and might be able to replace damaged tissue in the future. Currently, stem cell based clinical trials are ongoing for a broad range of diseases affecting every organ, and the number of

N. Graffmann · L.-S. Spitzhorn · S. Martins · M. S. Rahman · L. Nguyen · J. Adjaye (✉)
Institute for Stem Cell Research and Regenerative Medicine, Heinrich Heine University, Düsseldorf, Germany
e-mail: Nina.Graffmann@med.uni-duesseldorf.de; Lucas-Sebastian.Spitzhorn@med.uni-duesseldorf.de; Soraia.Martins@med.uni-duesseldorf.de; Shaifur.Rahman@med.uni-duesseldorf.de; Lisa.Nguyen@med.uni-duesseldorf.de; James.Adjaye@med.uni-duesseldorf.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_72

therapies that are obtaining approval, at least in certain countries, is continuously increasing. In this chapter we summarize the current state of research regarding stem cell based therapies. We introduce stem cell sources along with selected indications for their use and relevant clinical trials. We discuss options as well as limitations and risks of these treatments.

Introduction

It is a long-standing goal of regenerative medicine to cure diseases by replacing defective cells/tissues with healthy ones. Nowadays, with increasing improvements in stem cell technology, this dream comes closer to reality, although there are still many obstacles to overcome.

Stem cell therapy was initiated as a treatment for blood malignancies over 50 years ago and is now an integral part of handling these conditions. Preparations containing skin stem cell have been used to treat severe burns since 1981. However, for all other diseases, stem cell therapy is still at a stage of intense research with many ongoing clinical trials. New techniques for genome editing as, for example, CRISPR/cas9 (clustered regularly interspaced short palindromic repeat) or TALEN (transcription activator-like effector nucleases) have also found their way into regenerative medicine and augment hopes that genetic defects can be efficiently and routinely corrected in the future.

Stem Cell Sources

Adult and Pluripotent Stem Cells

Stem cells are undifferentiated cells capable of dividing symmetrically to maintain a stem cell pool or asymmetrically to produce terminally differentiated cells. They can be separated into pluripotent and multipotent cells according to their differentiation potential(Table 1). Pluripotent stem cells can generate every cell type of the human body, while multipotent stem cells are restricted to a few, closely related cell types. Because of their ability to replace damaged cells and restore tissues, stem cells are of major interest for the field of regenerative medicine (Chagastelles and Nardi 2011).

Stem cells can be obtained from many different sources, each associated with specific advantages and disadvantages. Usually, the disease that should be treated defines which stem cell sources are suitable and if autologous or allogeneic transplantation is needed. In the case of autologous transplantation, stem cells are obtained directly from the patient which obviates the risk of immune rejection. Cells for allogeneic transplantation are derived from an independent donor who is not necessarily related to the patient. Therefore, it is always associated with varying intensities of immune reactions and usually necessitates, at least temporarily, an immunosuppressive treatment. Although autologous stem cells are safe from an immunological point of view, they often cannot be applied due to other shortcomings. If the disease is caused by a genetic defect, all cells of the patient carry this mutation and hence are not suitable for restoring the function of the affected organ. In addition, in some conditions, there is simply not enough time for the sometimes lengthy process of isolation, purification, and potential biotechnological modification of autologous stem cells (Champlin 2003).

Pluripotent Stem Cells (PSCs)

Embryonic stem cells (ESCs), which are pluripotent, are isolated from the inner cell mass of the blastocyst during preimplantation development (Thomson et al. 1998). High expectations are evoked by these cells as they renew themselves indefinitely, do not age in culture, and can be directed in vitro to differentiate into every cell type of the human body (Chagastelles and Nardi 2011). However, their use is limited by major ethical considerations, as they can only be obtained by destroying the embryo. Regulations on the use of ESCs vary from country to country and influence research as well as cell therapy (Chagastelles and Nardi 2011; Elstner et al. 2009).

In 2006/2007 these ethical limitations were overcome, when the group of Shinya Yamanaka

Table 1 Definitions of stem cell-related terms

Term	Definition
Regenerative medicine	An area of medicine which uses cells or biological materials to treat diseases or improve organ/tissue functions
Advanced therapy medicinal product (ATMP)	Novel class of medicinal products comprising somatic cell therapeutics, gene therapeutics, and bioengineered tissue
Stem cells	Non-committed cells which have the ability to self-renew or generate higher differentiated daughter cells
Multipotent	Adult stem cells which differentiate only into a limited number of cell types, e.g., HSCs, MSCs, and NSCs
Pluripotent	Stem cells which can give rise to all cell types of the body. Naturally, only embryonic stem cells (ESCs), which reside in the inner cell mass of the blastocyst, are pluripotent. In culture, all cells can be coerced into pluripotency by overexpression of necessary factors (induced pluripotent stem cells, iPSCs)
Hematopoietic stem cells (HSCs)	Multipotent stem cells which reside in the bone marrow and form all cells of the blood
Mesenchymal stem cells (MSCs)	Multipotent stem cells which reside in almost all tissues and can differentiate into mesodermal cell types such as osteocytes, chondrocytes, and adipocytes. They secrete a variety of factors which reduce inflammation and promote regeneration
Neural stem cells (NSCs)	Multipotent stem cells which reside in the subventricular and the subgranular zone of the adult brain can differentiate into neurons, astrocytes, and oligodendrocytes. They secrete neurotrophic factors, cytokines, and growth factors
Holoclones	Subset of skin stem cells, with high proliferative and self-renewal potential
Gene editing	Molecular technique (e.g., zinc-finger nucleases, TALENs, and CRISPR/Cas9) for changing the DNA sequence in order to alter a gene's function or expression level
Autologous transplantation	Transplantation of cells that are taken from the patient
Allogeneic transplantation	Transplantation of cells which are taken from an independent donor
Good manufacturing practice (GMP)	(Inter)national rules to ensure high and consistent quality of medicinal products
Graft-versus-host disease (GvHD)	Potentially life-threatening conditions after HSC transplantation, where donor-derived immune cells attack cells from the patient

in Kyoto, Japan, managed to reprogram terminally differentiated murine and later also human cells into pluripotent stem cells. These so-called induced pluripotent stem cells (iPSCs) can be generated by overexpressing a cocktail of pluripotency-associated transcription factors in somatic cells which eventually resets these cells into an undifferentiated state (Takahashi et al. 2007). This was a major breakthrough in the field of stem cell research and was honored in 2012 with the Nobel Prize in Physiology or Medicine for Yamanaka (shared with John B. Gurdon) (Fig. 1).

Today, almost every cell type has been successfully reprogrammed, and the technique itself has been optimized in order to increase efficiency as well as safety. In the earliest experiments, retroviruses were employed to overexpress the pluripotency factors. However, as retroviruses randomly integrate into the genome, they can potentially induce cancer formation if, for example, their integration destroys a tumor-suppressor gene or activates an oncogene. Nowadays, reprogramming is usually performed by using non-integrating episomal plasmids or Sendai viruses (Abu-Dawud et al. 2018). Nonetheless, undifferentiated PSCs still might elicit tumors in the recipient, simply because of their proliferative

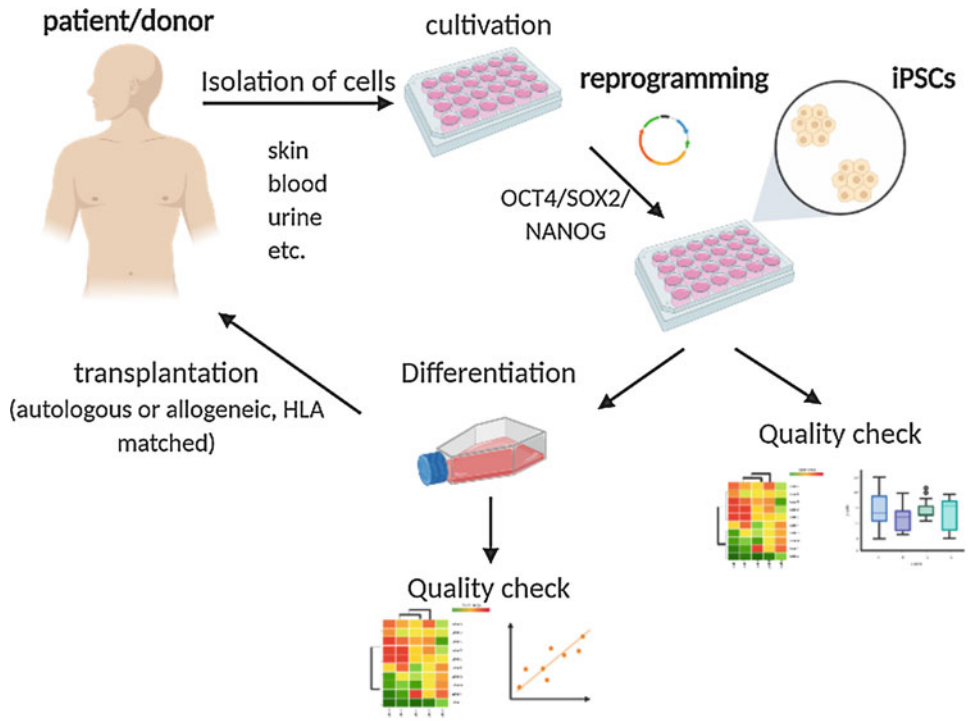


Fig. 1 Generation of iPSCs and their use as a potential cell therapeutics. Somatic cells are isolated from a donor, cultured, and reprogrammed into iPSCs. IPSCs are expanded and thoroughly checked, to ensure pluripotency and genomic integrity. Cells that pass the quality control are differentiated into the required cell type. After a quality assessment of the differentiated cells, ensuring successful differentiation and loss of pluripotent cells, the cells can be transplanted into the donor. Figure made in ©BioRender – biorender.com

potential. This is especially dangerous in the autologous setting where the transplanted cells are not recognized by the recipient's immune system. Therefore, sophisticated protocols have been established that allow efficient differentiation into all major cell types and include stringent quality controls to avoid transplantation of residual undifferentiated cells (Inoue et al. 2014).

iPSCs are associated with great anticipations for therapeutic employment as they can be generated from every patient; hence, their transplantation does not elicit immunological rejection. In addition, every given cell type can be reprogrammed which is especially interesting as it abolishes the need for invasive procedures, if, for instance, urine cells are taken as a source (Table 2). However, current data raise doubts regarding feasibility as reprogramming and all associated quality controls for excluding tumorigenic potential are not only very costly but, more importantly, time-intensive. Therefore, currently efforts are made to establish banks of iPSCs with homozygous HLA types in order to provide tolerable cells for a major part of the population. Japan with its very homogenous population is at the forefront of this and aims at having covered most of its population by 2030 (Kim et al. 2017).

Adult Multipotent Stem Cells

Hematopoietic Stem Cells (HSCs)

HSCs reside in specific niches in the bone marrow (BM) and can differentiate into all cells of the hematopoietic system. For transplantation, either they are isolated directly from BM aspirates, or they are collected by apheresis from peripheral blood after mobilization with granulocyte colony-stimulating factor (G-CSF). In both cases, the surface marker cluster of differentiation (CD)

Table 2 Ongoing clinical trials using PSCs listed at https://clinicaltrials.gov/ in August 2019

Disease	Therapy	Study ID	Sponsor	Country	Start
Spinal cord injury	ESC-derived OPCs	NCT01217008	Asterias Biotherapeutics	USA	2010
Stargardt macular dystrophy	hESC-derived RPE	NCT01345006	Astellas Institute for Regenerative Medicine	USA	2011
Dry AMD	hESC-derived RPE	NCT01344993	Astellas Institute for Regenerative Medicine	USA	2011
Stargardt macular dystrophy	hESC-derived RPE	NCT01469832	Astellas Institute for Regenerative Medicine	UK	2011
Dry AMD	hESC-derived RPE	NCT01674829	CHABiotech CO., Ltd	Korea	2012
Stargardt macular dystrophy	hESC-derived RPE	NCT01625559	CHABiotech CO., Ltd	Korea	2012
Exudative AMD	Autologous iPSC-derived RPE	UMIN000011929	RIKEN	Japan	2013
Spinal cord injury	ESC-derived OPCs	NCT02302157	Asterias Biotherapeutics	USA	2014
Ischemic heart disease	ESC derived $CD15^+$ $Isl\text{-}1^+$ progenitors	NCT02057900	Assistance Publique – Hôpitaux de Paris	France	2014
Type 1 diabetes mellitus	VC-01 combination product (device loaded with ESC-derived β-like cells)	NCT02239354	ViaCyte, California Institute for Regenerative Medicine (CIRM)	Canada, USA	2014
Dry AMD	hESC-derived RPE	NCT02286089	BioTime Inc.	Israel, USA	2015
Retinitis pigmentosa	hESC-derived RPE	ChiCTR-OCB-15007055	Institute of Zoology, Chinese Academy of Sciences	China	2015
Stargardt macular dystrophy, AMD, exudative AMD	hESC-derived RPE	NCT02903576	Federal University of São Paulo	Brazil	2015
AMD	hESC-derived RPE	NCT01691261	Pfizer	UK	2015
Dry AMD	hESC-derived RPE	NCT02590692	Regenerative Patch Technologies	USA	2015
AMD, Stargardt macular dystrophy	hESC-derived RPE	NCT02749734	Southwest Hospital, China	China	2015
PD	Human parthenogenetic neural stem cells	NCT02452723	Cyto Therapeutics Pty Limited	Australia	2016
Dry AMD	hESC-derived RPE	ChiCTR-OCB-15007054	Institute of Zoology, Chinese Academy of Sciences	China	2016
GvHD	CYP-001: iMSCs	NCT02923375	Cynata Therapeutics Ltd.	Australia, UK	2016
PD	ESC-derived NPC	NCT03119636	Chinese Academy of Sciences	China	2017
Dry AMD	hESC-derived RPE	NCT03046407	Chinese Academy of Sciences	China	2017
Neovascular AMD	Allogenic iPSC-derived RPE	UMIN000026003	Kobe City Medical Center General Hospital	Japan	2017

(continued)

Table 2 (continued)

Disease	Therapy	Study ID	Sponsor	Country	Start
Type 1 diabetes mellitus	VC-02 combination product (device loaded with ESC-derived β-like cells)	NCT03162926	ViaCyte	Canada	2017
Type 1 diabetes mellitus	VC-02 combination product (device loaded with ESC-derived β-like cells)	NCT03163511	ViaCyte, California Institute for Regenerative Medicine (CIRM)	Canada, USA	2017
PD	iPSC-derived dopaminergic progenitors	UMIN000033564	Kyoto University Hospital	Japan	2018
Dry AMD	hESC-derived RPE	NCT02755428	Chinese Academy of Sciences	China	2018
ALS	hESC-derived astrocytes	NCT03482050	Kadimastem	Israel	2018
Spinal cord injury	iPSC-derived NSCs	Approved	Keio University School of Medicine	Japan	2019
Retinitis pigmentosa	hESC-derived RPE	NCT03944239	Qi Zhou	China	2019
Retinitis pigmentosa due to monogenic mutation	hESC-derived RPE	NCT03963154	Centre d'Etude des Cellules Souches	France	2019
Heart failure	Allogenic iPSC-derived cardiomyocytes	NCT03763136	Help Therapeutics	China	2019
meniscus injury	hESC-derived MSC like cell	NCT03839238	Tongji Hospital	China	2019

34 can be used to separate HSCs from other cell types (Panch et al. 2017).

A third source of HSCs is the umbilical cord which contains stem cells in addition to very naïve immune cells. These cells can be collected directly after birth and are in contrast to HSCs from BM or apheresis frequently stored in liquid nitrogen for usage in the far future (Panch et al. 2017). Parents can choose to store the umbilical cord blood (UCB) either in public banks making it available for everyone or in private banks where it is kept exclusively for the donating family who also has to cover the costs. At the moment about six times more UCB units are stored in private banks than in public ones although the vast majority of UCB units used for transplantation are released by public banks (Ballen et al. 2015).

Routine allogeneic transplantation of HSCs only became possible after the human leukocyte antigen (HLA) system, which is fundamental for the function of the human immune system, was understood (Klein and Sato 2000; van Rood 1966). Nowadays, great efforts are undertaken to ensure that donor and recipient HLA type match in order to reduce the risk of graft-versus-host disease (GvHD) (Klein and Sato 2000). In this potentially life-threatening condition, donor-derived immune cells attack not only residual tumor cells in the recipient but also healthy tissue. Symptoms are ranging from rather mild skin reddening and itching toward massive damage of the gastrointestinal tract or the liver and always need special immunosuppressive treatments. If GvHD occurs in the first 100 days after transplantation, it is classified as acute and later as chronic. Over time, the new immune system is capable to adapt to the recipient's tissue surface antigens, and symptoms recede.

Mesenchymal Stem Cells (MSCs)

In 1924, Alexander A. Maximow identified cells residing within the bone marrow (BM) which had some similarities to fibroblasts. These cells were described to support hematopoiesis, the

generation of blood cells (Maximow 1924). By exploiting their ability of plastic adherence, Friedenstein and colleagues isolated these cells from BM (Friedenstein et al. 1970). The term "mesenchymal stem cells" (MSCs) was established by Arnold I. Caplan in the 1990s (Caplan 1991). Although in the beginning MSCs were exclusively isolated from the perivascular region of the BM, several other sources for MSCs have been established in recent years: placenta, umbilical cord blood (UCB), heart, skin, pancreas, lung, brain, kidney, adipose tissue, cartilage, tendon, as well as amniotic fluid and urine (Wu et al. 2018; Fitzsimmons et al. 2018; Spitzhorn et al. 2017; Rahman et al. 2018).

In addition to their extensive in vitro proliferation potential, MSCs exhibit a multipotential differentiation capacity which manifests in their ability to form various cell types from the mesodermal lineage including the cartilage, bone, and fat (Fitzsimmons et al. 2018). In vitro, they can also differentiate in cell types from other lineages such as pancreatic β cells or cardiomyocytes, although it is controversially discussed if this happens in vivo, too (Segers and Lee 2008; Cho et al. 2018). The fate-determining decision is triggered by various factors such as cytokines and growth factors present in the stem cell niche – a distinct microenvironment in which the stem cells are located (Fitzsimmons et al. 2018).

MSCs exhibit a fibroblast-like, spindle-shaped morphology. They are characterized by the expression of specific cell surface markers (also referred to as immunophenotype) such as CD73, CD90, and CD105 and the of lack hematopoietic markers. Together with their differentiation capacity and their plastic adherence, these features were set as the minimal criteria to identify MSCs by the international society for cellular therapy (ISCT) (Dominici et al. 2006). MSCs also have the ability to secrete various cytokines and growth factors. Thus, they can support hematopoiesis in the BM and are able to suppress reactions of the immune system. They support regeneration processes by the release of distinct molecules, for example, vascular endothelial growth factor A (VEGF-A), hepatic growth factor (HGF), platelet-derived growth factor (PDGF), and various interleukins (Fitzsimmons et al. 2018). Furthermore, MSCs can react to factors released by the environment. They migrate to sites of inflammation or injury to support the regeneration process. As MSCs express only very low levels of HLA, they can be transplanted even without matching their HLA type to that of the patient (Fitzsimmons et al. 2018).

To investigate their therapeutic potential, MSCs have been employed in over 740 clinical trials (ongoing or completed) for a broad range of diseases and organs affected which underlines their versatility (Fig. 2).

In most clinical trials, MSCs from adult BM are used followed by the adipose tissue, umbilical cord tissue, and placenta (Galipeau and Sensebe

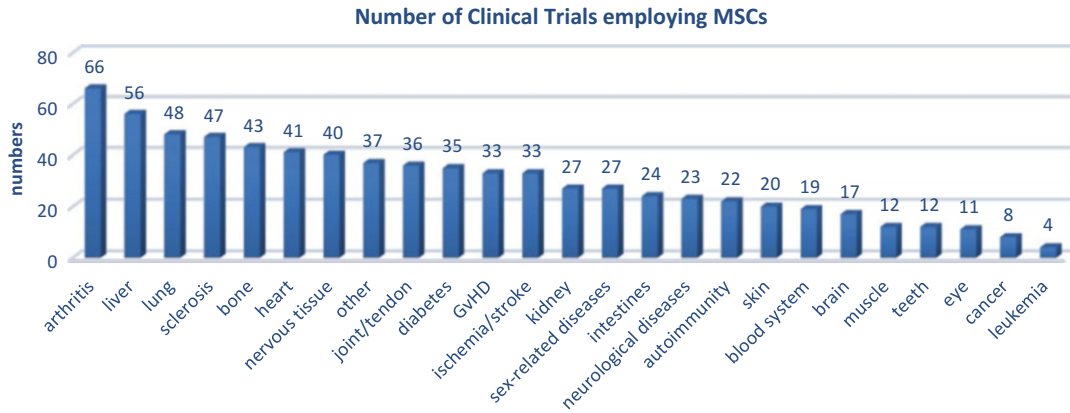


Fig. 2 Use of MSCs in clinical trials. Information on over 740 clinical trials using MSCs have been deposited at https://clinicaltrials.gov/ in March 2019

Table 3 Approved MSC-based therapies (aliancerm.org)

Product	Company	Indication	Treatment	Approval	Country
Queencell	Anterogen Co., Ltd.	Connective tissue disorders	Autologous adipose tissue MSCs	2010	South Korea
Cellgram-AMI	Pharmicell	AMI	Autologous BM-MSCs	2011	South Korea
Cupistem	Anterogen Co., Ltd.	Perianal fistula in Crohn's disease	Autologous adipose tissue MSCs	2012	South Korea
Cartistem	Medipost	Knee cartilage defects	Allogeneic UCB-MSC	2012	South Korea
Prochymal	Mesoblast Inc.	GvHD	Allogeneic BM-MSCs	2013	Canada, New Zealand
Neuronata-R	Corestem Inc.	ALS	Autologous BM-MSCs	2014	South Korea
Temcell HS	JCR Pharmaceuticals Co., Ltd.	GvHD	Allogeneic BM-MSCs	2015	Japan
Stempeucel	Stempeutics Research Pvt	Critical limb ischemia	Allogeneic BM-MSC	2017	India
Alofisel	TiGenix NV, Takeda Ltd.	Perianal fistulas in Crohn's disease	Allogeneic adipose tissue MSCs	2018	EU
Stemirac	Nipro Corp	Spinal cord injury	Autologous BM-MSC	2018	Japan
Neuronata-R	Corestem Inc.	ALS	Autologous BM-MSCs	2018	USA, orphan drug
Neuronata-R	Corestem Inc.	ALS	Autologous BM-MSCs	2019	EU, orphan drug

2018). The most prevalent way of administration is intravenously. This is very critical, because cell aggregation after infusion is one of the most prevalent adverse effects of MSC treatments (Caplan et al. 2019). MSCs are thawed immediately prior to transfusion. This is also a key step since it is well-known that the fitness of MSCs declines within the first 24 h after thawing which could impair their function (Caplan et al. 2019).

Despite the plethora of MSC-based clinical trials, so far worldwide only a few products based on MSCs have obtained clinical approval (Table 3).

MSC-Based Clinical Trials

In 1995 MSCs were the first cellular pharmaceutical tested in human. They were isolated from the BM of patients with hematologic malignancies in complete remission and reinfused intravenously (Lazarus et al. 1995). This study confirmed the safety of autologous MSC transplantation.

Later, MSCs were used in clinical trials to support hematopoietic recovery following high-dose myeloablative chemotherapy and accelerating hematopoiesis after BM transplantation (Fitzsimmons et al. 2018). The immunosuppressive capacity of MSCs is frequently used as a support in other transplantation settings, to reduce the risk of organ rejection or GvHD, especially if the HLA types of perfectly donor and recipient are not perfectly matching (Fitzsimmons et al. 2018).

In 2009, the first major industry-sponsored (Osiris Therapeutics (USA)) phase III trial of allogeneic MSCs for treatment of steroid-refractory GvHD (NCT00366145) was completed. Although it did not show an overall significant positive effect of the MSCs, it demonstrated that children with GvHD were responsive to this cellular treatment (Kurtzberg et al. 2014). In 2013, the therapeutic called Prochymal/remestemcel-L was conditionally approved in Canada and New Zealand for application in children with an obligation to prove its efficacy. Mesoblast Inc.

(Australia) took over the MSC therapeutic agents and focused on the prospective phase III trial (NCT00759018) (Kurtzberg et al. 2014). The overall response in the children 28 days after treatment was significantly improved in comparison to control patients. In Canada reimbursement issues prevented establishment as a common clinical treatment. In Japan on the other hand, the MSC therapy for GvHD and a proper reimbursement were approved and installed as Temcell in 2015 (JCR Pharmaceuticals Co. 2017).

For treating patients with complex perianal fistulas without inflammatory bowel disease, Cellerix S.A. sponsored a study of autologous adipose MSCs (NCT00475410). Forty-eight hours after thawing, the cells were injected in the area of the fistula. The outcome of the MSC group was not superior to the control group. After Cellerix S.A. was acquired by TiGenix NV in 2011, a new phase III clinical trial was initiated (NCT01541579), treating only Crohn's disease patients with an increased dose of allogeneic cells. The group receiving MSCs significantly improved their condition in comparison to the control group. The results disclosed in 2015 were reported to be the first completely successful outcome of MSCs in a clinical trial of an advanced state (Panés et al. 2016). The product, named Alofisel, was EMA approved first in 2009 as an orphan drug until it obtained full approval in 2018. A similar product, called Cupistem, got approval in South Korea.

The highest number of clinical trials with MSCs targets bone and cartilage diseases, for example, osteogenesis imperfecta, a genetic disorder which is characterized by mutations in the genes *collagen type I alpha 1 chain (COL1A)1* and *2*. This leads to osteopenia, multiple fractures, severe bone deformities, and considerably shortened stature. The first clinical trial to address this used allogeneic BM (containing MSCs) in children. It was shown that MSCs could migrate to the bone and improve bone structure (Horwitz et al. 1999). In a special case of such a treatment in 2005, allogeneic MSCs were transplanted in utero into a fetus with severe osteogenesis imperfecta. Within the first 2 years, only three fractures occurred, and correct growth tendencies as well as normal psychomotoric behavior were reported. At the age of 8 years, the same patient was treated again with MSCs of the same donor resulting in low-level beneficial outcome. This study has shown long-term safety, but still more work has to be done to unravel the effectiveness (Gotherstrom et al. 2014).

In patients suffering from osteoarthritis (OA), the articular cartilage is degenerated, and subchondral bone sclerosis and synovial inflammation lead to pain in the joints, local inflammations, and restricted movement. Intense research and many clinical trials are performed in this field, with limited positive outcome. In general, patients reported less pain and improved mobility after treatment, and improved cartilage quality could be observed, but due to study design, the evidence is not very reliable (Iijima et al. 2018). Nevertheless, South Korea approved Medipost's Cartistem in 2012, a product based on allogeneic UCB-MSCs which are plied to the joint during orthopedic surgery.

MSCs from Pluripotent Stem Cells (iMSCs)

The derivation of MSCs from BM or other sources is done routinely, but these invasive procedures are associated with certain risks (Sheyn et al. 2016). In the everyday clinical scenario with a high percentage of elderly people, the quality of MSCs is compromised due to cellular aging. This is manifested in higher levels of cellular senescence, DNA damage, oxidative stress, genomic instability, and immunogenicity which reduces their therapeutic potential (Yang 2018).

These hurdles can be bypassed by the generation of induced MSCs (iMSCs) from PSCs which have a rejuvenated phenotype (Frobel et al. 2014; Spitzhorn et al. 2019). Human iMSCs have been used in many preclinical studies for, e.g., liver and bone regeneration (Spitzhorn et al. 2018; Jungbluth et al. 2019, in press). A very important aspect for transplantation of iMSCs is the fact that although they are derived from PSCs, they are not tumorigenic and thus safe for transplantation. To date iMSCs have been successfully used in two clinical trials. Cynata Therapeutics Ltd. (Australia) is funding a phase I clinical study in which iMSCs coming from iPSCs are used for the treatment of steroid-resistant acute GvHD with preliminary

promising results (NCT02923375). The Tongji Hospital (China) in cooperation with the Chinese Academy of Sciences is running a phase I clinical trial for iMSCs generated from ESCs with the focus to evaluate the safety in treating meniscus injury patients (NCT03839238).

Neural Stem Cells (NSCs)

Neural stem cells (NSCs) are self-renewing multipotent cells with the potential to differentiate into neurons, astrocytes, and oligodendrocytes, the three main cell types in the central nervous system (CNS) (Seaberg and van der Kooy 2003). Additionally, NSCs are capable of secreting soluble molecules such as neurotrophic factors, cytokines, and growth factors. Due to these special characteristics, transplantation of NSCs is a promising treatment option for diseases associated with the CNS, for both regeneration of neural cells and restoration of the microenvironment at the injury site by providing trophic support.

First evidences of adult neurogenesis (the endogenous generation of new neurons) were described by Altman and Das in 1965 (Altman and Das 1965). In the adult brain, NSCs are confined primarily to two regions, the subventricular zone (SVZ) and the subgranular zone (SGZ). Adult neurogenesis is a dynamic but extremely coordinated process, where the proliferation, migration, and differentiation of the NSCs are controlled by microenvironmental stimuli. Although adult neurogenesis in the CNS is limited under normal physiological conditions, it can be induced after injury to recruit NSCs and reconstruct neural tissues. Unfortunately, in the case of severe injury, the response of activated NSCs is ineffective in keeping the balance between self-renewal and differentiation (Ming and Song 2011).

Cell therapy based on NSCs is now possible thanks to the comprehensive development of protocols based on growth factors to expand and differentiate these cells in vitro. NSCs can be derived from three major sources: directly extracted from primary CNS tissue (fetal and adult brain and spinal cord tissue), differentiated from pluripotent stem cells (PSCs), and through transdifferentiation from somatic cells (Tang et al. 2017).

NSCs derived from primary CNS: Human NSCs can be isolated directly from the ventricular zone of fetal brain tissue or from the SVZ of adult brain. After isolation, these cells can be propagated as single-cell suspensions that ultimately will form three-dimensional neurospheres in a medium supplemented with appropriate growth factors that allow proliferation, self-renewal, and expansion of NSCs in vitro (Gonzalez et al. 2016; Tang et al. 2017).

NSCs derived from PSCs: NSCs can be in vitro differentiated from ESCs and iPSCs via a process called neuroinduction. After this stage, NSCs can be maintained in culture in the same conditions as the NSCs directly isolated from the CNS. The use of iPSC-derived NSCs has advantages over ESC-derived NSCs since iPSCs can be generated from adult tissue, bypassing ethical issues. In addition, iPSC-derived NSCs can be obtained in a patient-specific manner, allowing autologous transplantation and thus overcoming the immune rejection (Tang et al. 2017; Gonzalez et al. 2016).

Transdifferentiation of somatic cells into NSCs: The generation of NSCs by transdifferentiation of somatic cells, a process that comprises the transformation of a somatic cell into another somatic cell type without undergoing the pluripotent stage, is a promising tool for therapeutic purposes. Although it was already shown that NPCs can be transdifferentiated from fibroblasts, urine cells, and MSCs, thus bypassing the tumorigenic potential of PSC-derived NSCs, it is a very recent field that needs further investigation (Tang et al. 2017).

Approval of Stem Cell Treatments

Stem cell therapy and regenerative medicine elicit high hopes – in patients who expect cures from a disease, as well as in industry that sees the great economic potential. This has led to the emergence of many clinics which perform unapproved stem cell treatments (Sipp et al. 2017).

The ethical and legal view on stem cells is highly divergent between countries. In the USA as well as in Europe, human cells and tissues are

mostly defined as advanced therapy technical medicinal products (ATMP). That means that they have to be approved by the US Food and Drug Administration (FDA) or the European Medicines Agency (EMA), respectively (Sakai et al. 2017). Usually this approval is only granted after thoroughly proving safety and efficacy of the treatment in controlled clinical trials. However, patients as well as industry express the need for a streamlined process which would reduce time and costs for approval. In 2014, Japan revised its law and introduced the Pharmaceuticals, Medical Devices, and Other Therapeutic Products (PMD) Act which provides conditional approval as long as the procedure has been proven to be safe. However, the new product has to prove its efficacy within 7 years. In the USA "the right to try" laws provide terminally ill patients the opportunity to test stem cell therapies before their final approval (Sakai et al. 2017). This provoked controversies as proper clinical trials are a highly valuable tool for determining safety and efficacy of novel treatments. On the other hand, many stem cell-based clinical trials (even in phase I/II) are only performed on terminally sick patients due to the high risks they are associated with, which decreases the likelihood of a positive outcome.

In parallel, a lot of unapproved stem cell therapies are on the market, many of them consisting in gaining MSCs, e.g., from liposuction and using these cells for a variety of treatments (Sipp et al. 2017; Goff et al. 2017). However, for these treatments neither safety nor efficacy has been studied in an accredited way, and in several cases there were dramatic outcomes such as severe infections or blinding (Marks et al. 2017). As clinics performing these unapproved treatments frequently advert them directly to patients with no independent physician explaining risks and benefits, they are very attractive for patients who are inclined to spend a lot of money and take incalculable risks (Sipp et al. 2017). Therefore, it is important to provide unbiased information to patients and educate the general population on risks and benefits of stem cell treatments.

Stem Cell-Based Regenerative Therapies

Hematopoietic System

From all areas of stem cell therapy, applications in the hematopoietic system have the longest history. The first allogeneic BM transplantation was performed in 1957 by E. Donnall Thomas, who paved the way for this kind of treatment and received the Nobel Prize in 1990 (Thomas et al. 1957). About 10 years after this first transplantation, techniques had developed that allowed for HLA typing, and the usage of stem cells from HLA-matched donors dramatically improved the outcomes of transplantations (Henig and Zuckerman 2014). As the HLA genotype limits the number of putative donors, international collaboration is mandatory. To this end the International Bone Marrow Transplant Registry (IBMTR) was established in 1972. The indications for HSC transplantation shifted with increasing knowledge about the diseases that are to be treated as well as with improved technologies that enable a broader spectrum of patients to be included (Little and Storb 2002).

In the beginning, predominantly hematopoietic cancers were treated with allogeneic stem cells. Later, also other blood malignancies were included, e.g., thalassemia or sickle cell anemia, and today also unrelated conditions as, e.g., solid tumors and autoimmune diseases are treated by HSC transplantation (Passweg et al. 2019).

Although therapy schemes have improved dramatically over time and the number of transplantations has increased immensely, HSC transplantation still is a very risky and stressful procedure for the patient. In a first step, the host's immune system and its stem cells have to be destroyed by irradiation and/or chemotherapy in order to remove the malignant cells. Nowadays there exist sophisticated protocols depending on the actual disease, which enable physicians to limit the destruction, especially of the stem cells, and thus allow also elder and fragile patients to undergo HSC transplantation (Gyurkocza et al. 2010). During the phase of engraftment, patients lack a functioning immune system and are very much at risk of developing

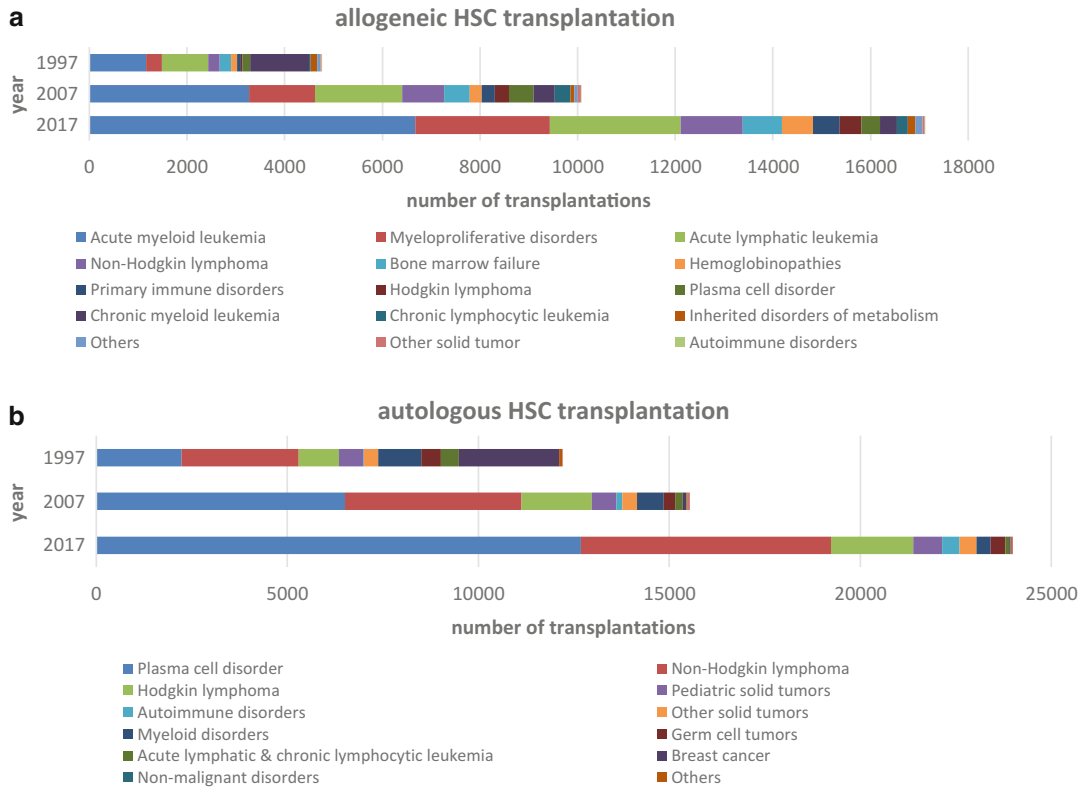


Fig. 3 Development of allogeneic (**a**) and autologous (**b**) HSC transplantation in Europe from 1997 to 2017. The number of HSC transplantations increased massively, with the frequency of distinct indications changing over time (Passweg et al. 2019)

infectious diseases. Only after the stem cells have homed into the BM niche, they efficiently produce new immune cells. However, in this phase the risk of GvHD is also augmented (Dowse and McLornan 2017) (Fig. 3).

Today, most HSCs for transplantation are collected via apheresis because this allows for a higher enrichment of stem cells, and the samples contain more T cells which are valuable for fighting the disease, called graft-versus-tumor/leukemia reaction. In addition, an increasing number of UCB transplants are performed. Although they contain less HSCs, which increases the risk of engraftment failure, their T cells are immature which reduces the risk of GvHD and allows for more HLA mismatches being present. Transplanting a combination of two independent UCB units seems to be the optimal solution in this case (Dowse and McLornan 2017; Panch et al. 2017).

Although HSC transplantation has been a routine treatment for many years, the mortality rates are still quite high. In acute leukemia it ranges between 7% and 27%, at 100 days post-allogeneic transplantation depending on disease severity. The leading causes are GvHD and infections. If patients survive for more than 2 years without relapse, their long-time survival rates are 80–92% (Henig and Zuckerman 2014).

Besides being a risky procedure, there are cases where allogeneic stem cell transplantation has additional benefits for the patients. Two patients who were infected with human immunodeficiency virus (HIV) reached remission for at least 20 and 18 months, respectively, after receiving allogeneic HSC transplants for treatment of leukemia or Hodgkin's lymphoma, respectively. Both patients were transplanted with HSCs from donors homozygous for a 32 base-pair deletion

within the *C-C chemokine receptor type 5 receptor* (*CCR5*) gene, which prevents expression of the receptor and thus precludes infection with HIV strains that need CCR5 as a co-receptor (Gupta et al. 2019).

Initial observations in the 1990s that HSC transplantation in cancer patients also improved coincidental autoimmune diseases (AD) developed this field as a new target for HSC transplantation. To date, several thousand patients have received HSC transplants (primarily autologous) to treat ADs, especially multiple sclerosis, Crohn's disease, and systemic sclerosis (Jessop et al. 2019). All ADs have in common that the patient's immune system attacks cells of its own body, leading to permanent inflammation followed by scar formation and impairment of the respective organ's function. HSC transplantation is capable of resetting the immune system, and it has been shown that after transplantation the amount of autoreactive immune cells can be reduced. However, HSC transplantation is still associated with many risks, and treatment outcome varies between patients and depends on disease severity. Nonetheless, many patients have long-term improvements, and for some ADs randomized controlled clinical trials have demonstrated the benefit of HSC transplantation compared to the best available conventional therapy (Jessop et al. 2019).

As HSC transplantation has been established a long time ago, it is also one of the fields where the most advanced techniques in terms of gene therapy are currently used. The first successful gene therapy trial was performed as early as 1990, although not targeting HSCs but T cells (Al-Saif 2019). Two patients who suffered from severe combined immune deficiency (SCID) caused by a mutation in the enzyme adenosine deaminase were included in this trial. The defective gene was exchanged for a functional copy in patients' T cells. Their blood values improved and they needed less medication after this treatment.

Current studies involve gene therapy directly in HSCs which can be performed in vivo as well as ex vivo. For in vivo gene therapy, genes are delivered to the HSCs by vectors or liposomes which can be injected into the bone or – after HSC mobilization – even intravenously. However, it has to be taken into account that these particles potentially can transfect also other cell types and frequently elicit strong immune reactions. On the other hand, they overcome transplantation-related issues as the need for chemotherapy in advance and the risk of poor engraftment. In ex vivo gene therapy, HSCs are collected from the patient and genes are modified in vitro. Besides simply overexpressing a healthy variant of the target genes, other possibilities include gene modification by applying, for example, TALENs or CRISPR/cas9. After the procedure, it is possible to selectively only transfer successfully modified cells back into the patient (Al-Saif 2019).

Currently, several clinical trials employing genetically engineered HSCs for transplantation are in progress. In sickle cell disease (SCD), a transversion of A to T within the coding region of the β-globin gene causes hemoglobin polymerization, which is associated with morphological changes of erythrocytes and impaired oxygen transport capacity. Standard treatment for SCD consists in applying hydroxyurea which increases the expression of the fetal y-globin, a molecule that prevents hemoglobin polymerization. Two gene therapies in clinical trials aim at increasing y-globin expression, either by overexpressing it directly (NCT02186418) or by inactivating a repressive factor (NCT03282656). In a different approach, a β-globin variant, which also prevents hemoglobin polymerization, is introduced into HSCs by a self-inactivating lentivirus (NCT02151526). Data from a first patient treated with this system indicate safety as well as efficacy (Al-Saif 2019).

In the future many therapy improvements are likely to be implied. One interesting novel approach concerns the field of chimeric antigen receptor (CAR) T-cell therapy. For this kind of therapy, until now, patient-derived T cells are genetically modified to express a chimeric antigen receptor specific for a tumor-related antigen. The engineered cells are then reinfused into the patient and supposed to eliminate the cancerous cells. As the whole procedure of engineering is very time-consuming, several groups are working on

protocols for differentiating iPSCs into T cells which can then be modified to express the chimeric receptor (Iriguchi and Kaneko 2019). A clinical trial with these cells has been already announced by a Japanese cooperation between the Center for iPS Cell Research and Application (CiRA) and Takeda (Kobayashi et al. 2019).

Skin Defects

The skin, the largest organ of the human body, is responsible for defending external stresses, regulating fluid retention and temperature, and mediating sensation. However, burns, infectious diseases, and inherited and acquired genetic defects of the skin affect about one billion people globally, leading to mortality and morbidity and requiring long-term hospitalization (Karimkhani et al. 2017). Keratinocyte cell-based autologous and allogeneic epithelial sheet transplantation is a clinically accepted technique, which has been used for more than 30 years to treat burned patients (Hernon et al. 2006). The clinical application of cultured epithelial allogeneic and autologous grafts to treat chronic ulcers, vitiligo, junctional epidermolysis bullosa (JEB)-non-Herlitz, JEB-Herlitz, recessive dystrophic epidermolysis bullosa (RDEB), epidermolysis bullosa (EB) simplex, and toxic epidermal necrolysis has also been reviewed (Petrof et al. 2014).

In 1975, for the first time, Rheinwald and Green isolated and cultivated human epidermal keratinocytes (the main basal cells of skin) obtained from a skin biopsy, which led to the development of cultured epithelial sheets (CEAs) (Rheinwald and Green 1975). Subsequently, in 1981 O'Connor et al. experimentally transplanted epidermal autografts to heal severe burn wounds using cultured keratinocyte clones (O'Connor and Mulliken 1981). These cells would nowadays be categorized – at least partly – as so-called holoclones, the undifferentiated colony-forming keratinocyte stem cells which have a higher growth and self-renewal potential compared to other colony-forming cell types (Barrandon and Green 1987). This type of skin equivalent graft is fabricated on a fibrin matrix where fibroblasts support the growth of an epithelial sheet containing basal and differentiated keratinocytes.

Beside the autologous keratinocyte-based therapy, promising findings were reported from a clinical trial (ten pediatric patients) using allogeneic BM-MSCs to manage second-degree burn wound healing (NCT02104713) and for the treatment of recessive dystrophic epidermolysis bullosa (RDEB) (EudraCT number 2012-001394-87, ISRCTN46615946). RDEB – a rare skin disease – results in the severe blistering of the skin, caused by a mutation in the *collagen type VII alpha 1 chain (COL7A1)* gene which impairs expression of the protein. In the RDEB clinical trial, patients were provided three infusions of allogeneic BM-MSCs. After 6 months of BM-MSC administration, improvement of wound healing, reduction of skin redness and blisters, and a decreased level of pain were noticed compared to the initial condition. Overall, the severity of the condition was reduced, and no substantial negative effects were reported (Petrof et al. 2015).

MSCs isolated from human UCB-MSCs are also widely studied and clinically accepted. They are in an early phase I clinical trial running until 2021 to evaluate the effect and safety for the treatment of psoriasis, a chronic and recurrent inflammatory skin disease (NCT03765957). In a case study published in 2018, injection of a MSC-containing cell preparation obtained from adipose tissue could significantly improve the condition. The mechanisms that are involved, particularly, whether and how these cells activate immune system associated growth factors and cytokines to support wound healing and reduce inflammation, are not well-understood (Comella et al. 2018).

Due to considerable progress in stem cell and genome editing technologies during the last decade, genetically engineered skin stem cell-based therapies for the treatment of fatal conditions such as Netherton syndrome (NS), junctional epidermolysis bullosa (JEB), and RDEB have found their way into experimental clinical treatments.

NS is a serious skin disorder caused by damage in the *serine peptidase inhibitor, kazal type 5 (SPINK5)* gene that encodes a protein called

lymphoepithelial kazal-type inhibitor (LEKTI) that is responsible for the function of the skin barrier. The skin of *SPINK5*-mutated patients becomes red and scaly (ichthyosiform erythroderma). Functionally, the skin leaks fluid and loses the capacity to bind water which causes skin dryness and hypernatremic dehydration. This and other complications, e.g., bronchopneumonia or sepsis, lead to a high mortality rate in early life (Di et al. 2019).

Currently, there are no approved treatments to cure this lethal condition. Based on the principle of ex vivo gene modification in stem cells, a clinical phase I trial of grafting autologous epithelial sheets harboring intact *SPINK5* for NS has been conducted (Di et al. 2019) (NCT01545323). It proved successful in terms of gene modification and engraftment, but only transient functional correction was observed, possibly due to a limited number of stem cell holoclones within the graft.

A groundbreaking example of skin stem cell therapy is that of a child having recovered from a fatal form of JEB in Germany in 2017 after engraftment with epidermal sheets consisting of autologous skin stem cells which were genetically modified to contain healthy copies of the *Laminin Subunit Beta 3 (LAMB3)* gene. Impressively, this *LAMB3* harboring epidermal autograft could restore dermal-epidermal adherence, and the long-lived skin stem cell holoclones could regenerate and renew the epidermis. Upon 1-month post-grafting, the new epidermis had formed without blisters, and the regenerated epidermis was strong and elastic alike healthy skin (Hirsch et al. 2017). Previously, in 2006, genetically engineered autologous keratinocyte grafts have been transplanted in a patient with laminin 332 deficiency (non-Herlitz junctional EB) (Mavilio et al. 2006). In a clinical study from Austria in 2014, genetically modified autologous epidermal sheets were grafted in an adult woman to cure from laminin 332-β3-dependent JEB with a large non-healing epidermal ulceration (Bauer et al. 2017). The transgenic epidermis expressed a normal amount of properly functional laminin 332 which is precisely located at the dermal-epidermal junction. These experimental results could benefit JEB patients with disrupted laminin 332.

The above-stated cases of JEB have demonstrated that it is possible to cure genetically damaged non-healing skin by using gene therapy in epidermal stem cells. However, for the EB simplex, which is not caused by the lack of the LAMB3 protein but by an active dysfunctional version of the protein, other approaches are necessary. Corrections with a gene-editing innovative tool like CRISPR/Cas9 might be possible. The feasibility of genome editing in skin stem cells using a self-inactivating (SIN) retroviral vector and CRISPR/Cas9 approach in a preclinical study in mice has already been reported (Hainzl et al. 2017; Izmiryan et al. 2018). Currently, this technique is also used in clinical trials in RDEB patients. The aim is to completely cure the disease using autologous skin grafts made of primary keratinocytes and fibroblasts with genetically edited type 7 collagen (NCT01263379, NCT02493816).

A special tissue composed of keratinocytes is the cornea of the eye. The integrity of this epithelial layer and the absence of blood vessels are crucial for proper vision. In a healthy setting, corneal cells are regenerated by limbal stem cells, located between the cornea and the bulbar conjunctiva. Thermal and chemical burns of the eye often destroy the limbal stem cells which results in vascularization of the cornea and ultimately vision loss (Pellegrini et al. 2009). Beginning in the late 1990s, the first patients were treated with transplants generated from limbal stem cells, and permanent restoration of the cornea was achieved in 76.6% of the cases. Importantly, a 10-year follow-up proved long-term stability of the regenerated tissue (Rama et al. 2010). The treatment relies on autologous cells which can be obtained if at least 1–2 mm intact limbal tissue is available. The cells are expanded and prepared for transplantation on a fibrin matrix. This product, marketed as Holoclar, was the first stem cell-based ATMP authorized by the European Medicines Agency in 2015.

Tremendous advancement of ESCs, iPSCs, and genome editing technology has been made during the last 10 years; now the major goal is to harness these technologies to generate unlimited amounts of genetically corrected keratinocytes not only from autologous but also from allogeneic donors.

Human PSCs can be differentiated into keratinocytes which are the basis for in vitro-generated functional epidermal sheets. These sheets could provide temporary skin substitutes for patients awaiting autografts (Guenou et al. 2009). Tolar and colleagues demonstrated that gene-corrected iPSCs could be generated from the skin of patients with a mosaic form of RDEB (Tolar et al. 2014). Compared to adult stem cells, iPSCs have a higher proliferation potential, allowing genetic manipulations and to overcome the autograft shortages. By generating iPSCs and, subsequently, iPSC-derived keratinocytes, they were able to provide proof of principle that iPSC technology can be used to generate essentially unlimited amounts of clinically normal epidermis from patients (Sebastiano et al. 2014; Wenzel et al. 2014).

Cardiac Diseases

The heart is the most important muscle in the human body. It pumps blood through the body and provides the essential oxygen and nutrient supply. Heart diseases can severely restrict the life of patients. Ischemic or coronary heart disease (CHD) is the most common form of heart diseases. It manifests in a reduced blood flow to the heart due to plaques which partially block its arteries. Complete blockage of the blood flow causes damage to the heart cells leading to myocardial infarction (MI) and even heart failure. Heart failure is defined as a syndrome that restricts the cardiac function of filling or ejecting blood in the ventricle, irrespective of cause. This condition has far-reaching consequences for the whole body and is not restricted to the heart only but can effect almost every other organ (Segers and Lee 2008).

Similar to other cases of organ failure, treatment of patients with end-stage heart failure includes organ transplantation, which greatly enhances the patients' quality of life as they regain physical endurance and are less hospitalized (Michler 2018). However, the demand of donor organs is much higher than the pool of available donors can manage. Many patients are also not eligible for transplantations depending on their conditions. They are particularly in the focus for the future use of stem cell therapy (Michler 2018).

The goal of stem cell treatment for heart diseases is the generation of cardiomyocytes and blood vessels, which support the regeneration of cardiac function (Michler 2018).

Some clinical trials already applied BM-MSCs. In an in vivo rodent model, differentiation of BM-MSCs to cardiomyocytes was reported when injected into the murine myocardium (Toma et al. 2002). However, it is debatable if these cells are able to generate cardiomyocytes in human patients, too. It is not proven that injected stem cells regenerate the heart tissue by differentiating into cardiomyocytes, but it is believed that secreted factors comprising growth factors, cytokines, and chemokines may improve the regeneration by activating reparative mechanisms and inhibiting apoptosis, fibrosis, and hypertrophy. Moreover, it seems that infused stem cells support the proliferation of cardiomyocytes and also recruit cardiac stem cells which regenerate the cardiac tissue (Segers and Lee 2008).

With regard to ischemic heart diseases, some clinical trials also showed that the injection of stem cells may lead to the formation of blood vessels which improved cardiac performance (Segers and Lee 2008).

In March 2019, more than 300 stem cell-based clinical studies for heart diseases were listed at https://clinicaltrials.gov/. Many of these are performed with BM-MSCs in various cell doses and varying infusion time points. Even though these cell therapies did not show significant improvements in the healing process of the heart diseases, there are some points which could be learned from these trials: (a) stem cell therapy is safe for patients, (b) therapy with cells is minimally effective, and (c) transdifferentiation of BM-MSCs does not occur in a frequency high enough to have an effect (Michler 2018).

In 2001/2002, the PERFECT phase II clinical trial, sponsored by Miltenyi Biotec GmbH, was initiated in order to study safety as well as efficiency of autologous $CD133^+$ BM-derived stem cells when injected into the myocardium during a coronary artery bypass graft (CABG). As primary

outcome, left ventricular contractility was measured and compared to CABG alone. Meanwhile, this study has developed into the first phase III clinical trial. Eighty-two patients with myocardial infarct were either injected with autologous $CD133^+$ BM-MSCs or placebo product while undergoing bypass surgery in a double-blinded and randomized fashion (Steinhoff et al. 2017). Analyses 180 days after the treatment showed that cardiac tissue repair and the improvement of left ventricular function were induced, possibly due to the presence of CD133 endothelial progenitor cells in the stem cell preparation (Steinhoff et al. 2017).

Between 2004 and 2005, 204 patients suffering from acute myocardial infarction (AMI) were treated in a phase III clinical trial called REPAIR-AMI (NCT00279175). The patients were randomly injected with autologous BM-MSCs or placebo medium into the infarct artery 3–7 days after infarct reperfusion therapy. The 12-month follow-up results showed a reduced number of patients who met the endpoint death, needed a revascularization, or had another MI in the BM-MSC-treated group than in the placebo group (Schachinger et al. 2006).

Another clinical trial combined the use of left ventricular assist devices (LVAD) with the injection of allogeneic MSCs (Ascheim et al. 2014) (NCT01442129). Thirty randomized patients suffering from end-stage heart failure were injected 25 million MSCs each during LVAD implantation. After 90 days, the functionality of the ventricle was checked while temporarily weaned from the LVAD. Even though an effect of MSCs was not clearly observed, the treated patients were able to wean of the LVAD more often and for longer periods.

Alternative to the injection of single cells, sphere-derived cells can be transplanted into the diseased tissue. Spheres are three-dimensional cell aggregates composed of cell types specific for one tissue or organ. In the clinical trial CADUCEUS (NCT00893360), patients with ischemic LV dysfunction were randomly injected with autologous cardiosphere-derived cells (CDC) generated from myocardial biopsies into the infarct-related artery 90 days after the myocardial infarct. The control group received standard treatment (Malliaras et al. 2014). Follow-up examination of the patients revealed no improvement of end-diastolic volume, end-systolic volume, and left ventricular ejection fraction. However, CDC treatment reduced scar mass and improved viable heart mass as well as regional contractility and systolic wall thickening (Malliaras et al. 2014).

Between 2013 and 2018, hESC-derived cardiovascular progenitors were transplanted into ten patients who suffered from ischemic heart failure in a clinical trial performed in France (NCT02057900). The generated cells were embedded into fibrin gel and were injected while the patients surgically received a coronary artery bypass grafting and/or a mitral valve procedure. A follow-up after 1 year revealed no adverse events during the recoveries and no tumor formation. None of the patients suffered from arrhythmias, and the patients' symptoms improved with increased systolic motion. Silent alloimmunization (immune response to antigens of an allogeneic donor) occurred in three patients. The conclusion was that the derivation of cardiovascular progenitor cells from hESCs was feasible, and transplantation of these cells was safe for short and medium term (Menasche et al. 2018).

All in all, many clinical trials with stem cell-based therapies for heart diseases were performed in the last 20 years. Especially, BM-MSCs were used extensively in these trials but with no or only minimal success. Nevertheless, there are, alone in the USA, more than 60 direct-to-costumer businesses cardiomyocytes promising to cure patients suffering from heart diseases with stem cells. These treatments are not authorized by the US FDA and expose the patients to incalculable risks (Goff et al. 2017). Until the mechanisms related to heart development and metabolism are better understood and risk factors are eliminated, stem cell therapy for heart diseases is still only at the very beginning.

Diabetes

Diabetes is one of the major health problems of our time. In 2017 about 425 million adults worldwide were affected with increasing tendencies. Ninety percent of the patients suffer from type 2 diabetes (T2D) which usually develops later in

life. Type 1 diabetes (T1D) can already start at a young age and currently affects more than 1.1 million children (Diabetes Facts & Figures 2019). T1D is an autoimmune disease where immune cells attack and destroy pancreatic β cells. These cells are responsible for the production of insulin in response to dietary glucose. Insulin enables cells, especially those of the muscles, to take up glucose for energy generation. If insulin is missing, blood glucose levels increase dangerously high which acutely can cause a diabetic coma. In the long term, elevated glucose levels can harm almost every organ, leading to cardiovascular problems, retinopathy, nephropathy, and neuropathy. In T2D the pancreas still produces insulin, but the peripheral tissues become resistant toward it, which means that insulin loses its ability to channel glucose into the cells. Other insulin-associated functions, especially its promoting effect on fat storage, remain intact. Over time, the pancreas becomes exhausted in T2D and stops producing insulin similar to T1D. While T1D always needs treatment with insulin injections, only later stages of T2D depend on injections (Kharroubi and Darwish 2015).

From the 145 stem cell-based clinical studies for diabetes listed in August 2019 at https://clinicaltrials.gov/, about one third deal with T2D, while the majority has been designed for T1D. Several approaches for stem cell therapy of diabetes have been tested. They are based either on the concept of replacing nonfunctional β cells by stem cell-derived ones or on exploiting the capacities of MSCs or HSCs to act as immunomodulators and improve the regeneration of lost cells. In view of this, MSCs are able to suppress T-cell responses and inhibit differentiation of dendritic cells as well as B-cell proliferation. They can stimulate the production of anti-inflammatory cytokines and suppress that of pro-inflammatory cytokines, as well as of reactive oxygen species. In addition, they can improve insulin signal transduction, while HSCs stimulate angiogenesis in the damaged islets as well as regeneration of endothelial progenitor cells (Sneddon et al. 2018).

Clinical trials for T1D and T2D usually employ MSCs from different sources or BM-derived mononuclear cells/HSCs either alone or in combination by intravenous injection or injection into the pancreatic artery. All clinical trials so far proved safety of the procedure and also showed limited therapeutic effects of the treatment. However, follow-up periods usually were no longer than 12 months, so we lack information on long-term success, and also no systematic study has been performed so far to determine which stem cell source and which mode of application are the best (Sneddon et al. 2018).

MSCs also have a capacity for in vitro differentiation into islet-like insulin-producing cells. These differentiated cells have been successfully used in clinical trials, although it is not yet clear whether they really replaced damaged β cells or if the positive effects were again due to immunomodulatory effects of MSCs (Cho et al. 2018).

In order to replace non-functioning β cells, novel techniques based on ESCs have been developed. To this end, ESCs are in vitro differentiated into insulin-producing β-like cells. Unfortunately, these cells are less efficient in insulin synthesis than primary β cells, which is probably due to the missing 3D structure of the organ, which provides a dedicated niche for β cells. Therefore, differentiation protocols are currently being developed which take also the 3D structure into account (Sneddon et al. 2018).

Besides cell maturation/function, cell therapy of the pancreas faces two main challenges which even hamper treatment with adult pancreatic cells. First, the transplanted cells home-in inefficiently into the pancreas, and second, they are frequently attacked by the host's immune system. To overcome both issues, several companies have developed encapsulation devices made of, e.g., alginate, polytetrafluoroethylene, silicone, or polycaprolactone. These devices can be transplanted either intraperitoneally or subcutaneously. Ideally, they provide sufficient blood and oxygen supply which ensures proper function of the encapsulated cells and allow the secreted insulin to exit the device. In addition, they protect the transplanted cells from the host's immune system and in the worst case prevent the escape of tumorigenic cells (Sneddon et al. 2018). One clinical phase I/II trial has been conducted by Sernova

Corp. employing isolated islet cells in a pouch (NCT01652911). The method proved to be safe. Six weeks after transplantation, the pouches were explanted, and histological analysis demonstrated that islets retained their macrostructure and were connected to new blood vessels. However, the end goal of long-term insulin independence of the patients could not be achieved. Another phase I/II clinical trial which is running until 2021 is currently conducted by ViaCyte to test safety and efficacy for a device loaded with ESC-derived insulin-producing β-like cells (NCT02239354). Preliminary results indicate that the devices are well-tolerated and that insulin-producing cells persist up to 2 years after transplantation (Pullen 2018).

Another stem cell-based technique tested for the treatment of diabetes is the so-called stem cell educator therapy. In this setting, mononuclear cells are separated from the patient's blood in a closed-loop cell separator. They are then educated to gain immune tolerance by briefly co-culturing them with adherent UCB stem cells. The educated mononuclear cells are reinfused into the patient (Zhao et al. 2013). A clinical phase I/II trial, performed by Tianhe Stem Cell Biotechnologies, demonstrated safety as well as efficacy of this treatment in T2D patients over up to 56 weeks (NCT01415726). Patients not only tolerated the treatment well but also had improved values for glycosylated hemoglobin (HbA1C), c-peptide, and homeostasis model assessment (HOMA) of insulin resistance. The company is now performing a phase II trial including T1D patients as well (NCT03390231) which is supposed to be completed by the end of 2020.

Liver Diseases

The liver is the central metabolic organ of the human body. It is involved in nutrient as well as in drug metabolism, and it produces a plethora of proteins which are secreted and have essential functions elsewhere in the body. The liver is a very complex organ. Eighty percent of its cells are hepatocytes which are responsible for the metabolic homeostasis. They store, release, and synthesize glucose and lipids according to the body's energy needs. They are involved in protein breakdown and detoxify ammonium ions in the urea cycle. The most important function of hepatocytes from a pharmacological point of view is the so-called first- and second-phase metabolism where drugs either get activated or detoxified by cytochrome P450 enzymes and afterward by conjugation with distinct cofactors (Stanger 2015).

There are mainly two indications for stem cell therapy in the liver. First, stem cells can be used to mitigate congenital metabolic defects. Second they are employed to restore liver tissue in the fibrotic/cirrhotic liver.

Several metabolic diseases are known where genetic defects of key enzymes either reduce the output of a specific pathway or impair this pathway completely. Usually, these are rare diseases, affecting less than 1 in 1500–2500 people, and they manifest early in childhood. Well-known examples are Crigler-Najjar syndrome or the various urea cycle defects. In Crigler-Najjar syndrome, patients lack the enzyme uridine diphosphate glucuronosyltransferase which is responsible for conjugating bilirubin – a degradation product of hemoglobin – and thus enabling its excretion. If unconjugated bilirubin accumulates in the body, it causes severe icterus which even affects the brain (kernicterus) leading to encephalopathy and thus triggering neurological damage and long-term cognitive defects. The current standard of care is a lifelong phototherapy for 10–12 h a day in order to convert unconjugated bilirubin with blue-light irradiation into a water-soluble form that can be excreted via the bile. Urea cycle defects result from mutations in any of the cycle's enzymes. This cycle is needed to detoxify ammonium ions which are generated during protein degradation and can, like unconjugated bilirubin, cause encephalopathy (Sokal 2014).

Two characteristics make metabolic diseases especially amenable for stem cell therapy. First, they are caused by an isolated mutation of one specific gene, and second, significant metabolic improvement can be obtained by replacing only 5% of total hepatocytes with healthy ones, while as few as 10% of healthy cells are expected to normalize liver function (Sokal 2014). From 2012

to 2014, a clinical phase I/II trial took place to evaluate the safety of a stem cell treatment in 6 pediatric Crigler-Najjar syndrome and 14 urea cycle disorder patients (NCT01765283) (Smets et al. 2019). Patients obtained transhepatic infusions with Heterologous Human Adult Liver-Derived Progenitor Cells (HepaStem), a product by the Belgian company Promethera. These stem cells are obtained by isolation and short-term cultivation of parenchymal liver cells which display a mesenchymal-like phenotype and preferentially differentiate into hepatocytes. They were administered in an immunosuppressive setting, and the safety could be proven in a 12-month follow-up period. In addition, a certain level of efficiency could also be shown. A proof-of-concept study in a 3-year-old girl suffering from a urea cycle disorder not only showed preliminary efficiency of the transplanted cells but also indicated that these cells were proliferating in the liver which means that maybe less cells are needed than initially expected (Sokal et al. 2014).

Currently, there are 20 clinical trials listed which employ stem cells to treat liver cirrhosis, while an additional 40 trials cover both cirrhosis and fibrosis (https://clinicaltrials.gov/, 5.8.2019). Most of these trials are conducted with autologous or allogeneic MSCs from various sources. All chronic liver diseases associated with inflammation (viral hepatitis, primary biliary cirrhosis, alcoholic and non-alcoholic fatty liver disease) can potentially progress via a fibrotic stage to the end stage of liver cirrhosis. In this condition, functional liver tissue is replaced by scars made of fibrotic material. The liver gets rigid and cannot perform its functional tasks properly anymore. In this phase, orthotopic liver transplantation is the only option to save the patient's life. However, there is a shortage of suitable donor organs, and researchers are working on cell therapies in order to temporarily improve the liver's function, enabling the patient to survive longer time periods before transplantation with the final goal to replace liver transplantation altogether. The first completed trials which all employed some kind of MSC obtained, e.g., from autologous fat tissue or BM indicate an overall safety of the procedure with limited success (Hu et al. 2019).

Kidney Diseases

The kidneys are a bean-shaped pair of organs, which are essential for removal of toxins and waste products from the blood, maintenance of fluid homoeostasis, secretion of hormones, and water reabsorption from urine. An adult kidney is composed of around one million nephrons, the functional units of the kidney.

Conditions such as diabetes and high blood pressure as well as genetic disorders are the main causes of chronic kidney disease (CKD). CKD is characterized by a gradual loss of renal functions which peaks in kidney failure at the worst outcome. In addition, acute kidney injury (AKI), which is defined as abrupt occurring kidney damage or failure, can cause renal diseases. Until now, kidney failure is only treatable with dialysis and finally organ transplantation. Transplantation is the only treatment which can recover kidney function, but it is accompanied with lifelong immunosuppression. Moreover, the organ demand is much higher than the available donor organs.

An alternative to organ transplantation and dialysis could be stem cell-based therapy. Researchers work on growing stem cell-derived organs in vitro, which may be used for transplantations and might cope with the high demand for donor organs. Another option may be injection of isolated stem cells into the affected region. It is proposed that these cells then differentiate into the required cell type within the organ.

One hereditary kidney disease is polycystic kidney disease (PKD). Characteristically, fluid-filled cysts form in the tubules of the kidney which increase the organ size and can spread to other organs. Damage in the kidney may result in end-stage renal disease and kidney failure. The autosomal dominant form of this disease (ADPKD) is caused by a mutation of the genes *polycystin* (*PKD*) *1* or *2*. The respective proteins are located in primary cilia and involved in calcium-dependent signalling pathways. Their dysfunction influences cell proliferation and structure as well as fluid secretion (Igarashi and Somlo 2007). A clinical trial, performed in 2014, with a focus on ADPKD patients depended on the use of MSCs. It was hypothesized that an introduction of

BM-MSCs can improve kidney function while acting anti-apoptotic, anti-fibrotic, and anti-inflammatory (NCT02166489). Six patients were chosen for this trial. 2×10^6 cells/kg autologous BM-MSCs were infused intravenously. The patients were observed afterward during a time span of 12 months. Overall, the procedure was safe for the patients, but the effects of the MSCs could not be evaluated due to the trial design (Makhlough et al. 2017).

Since the kidney contains a great variety of different cell types, this organ is especially challenging to model in vitro. Successful treatment of kidney diseases can only be obtained, when the affected or damaged cell types are recognized. Upon better understanding about the affected cell types in the different kinds of kidney diseases, more treatments and therapies will be established in the future.

Diseases of the Central Nervous System (CNS)

The central nervous system (CNS), composed of the brain, spinal cord, and retina, is one of the most complex and less understood systems in the human body. CNS disorders and injuries, among them neurodegenerative diseases, retinal degeneration, and spinal cord injury, not only cause devastating consequences for the patients and family members but are also a major economic burden.

Neurological disorders are complex and usually irreversible, partially due to the limited potential of endogenous regeneration of the CNS. With the lack of effective therapeutic approaches, stem cell-based therapy holds promising potential to treat CNS disorders.

Retinal Degradation

Retinal disorders were the first CNS conditions to be targeted with cell therapy. The retina comprises several neuronal layers, among them the photoreceptors (PRs) which convert the light inputs into signals that are then transmitted to the brain via the optic nerve. The PRs are in contact with a monolayer of retinal pigment epithelium (RPE), essential to maintain PR homeostasis. Degeneration of the PR due to their malfunction or due to degeneration of the RPE can lead to visual decline that ultimately ends up in blindness.

The leading causes of retina degeneration and blindness in Western countries are age-related macular degeneration (AMD) and retinitis pigmentosa. AMD can be subclassified into non-vascular AMD (dry form) and neovascular AMD (wet form). Another variant of this disease, which is not age-associated, is Stargardt macular degeneration. It is the most prevalent inherited macular dystrophy affecting both children and adults. While AMD is caused by the degeneration of the RPE, retinitis pigmentosa is caused by the degeneration of the PR and/or RPE. Retinitis pigmentosa is a group of heterogeneous inherited conditions with a prevalence of 1:4000. Since the PR and the RPE cannot be endogenously regenerated, stem cell therapy is a promising therapeutic alternative for treating these retinal degeneration conditions (Ben M'Barek and Monville 2019).

RPE and PR cells can be obtained from fetal and adult tissue, specifically from cadavers (allogeneic RPE) or from patients (autologous RPE) through a nasal surgery. The use of fetal tissue implies ethical concerns and showed variability related to the developmental state of the fetus, leading to the abandonment of this source. Adult RPE cells showed a huge limitation due to the low amount of cells that can be obtained, as well as the complications inherent to the surgery in case of autologous RPE isolation. With the development of efficient protocols to differentiate ESCs and iPSCs into specific retinal subtypes like RPE cells, retinal cell replacement is a promising target of CNS cell therapy, with multiple clinical trials ongoing (Goldman 2016).

The first clinical trials using ESC-derived RPEs to treat AMD and Stargardt disease were performed in 2012 (NCT01344993 and NCT013450060). The results of these studies provided evidence that ESC-derived RPEs were safe to treat these diseases. As a consequence, many other clinical trials emerged placing retinal degeneration at the forefront of stem cell-based therapies. In 2014, the first clinical trial using autologous iPSC-derived RPEs was conducted at the RIKEN institute in Japan

(UMIN000011929). However, it was aborted 1 year later due to safety concerns. There are currently three ongoing clinical trials using iPSC-derived RPEs, with more expected to be approved. The results from these trials will help to optimize the best formulation strategy to combat retinal degeneration (Ben M'Barek and Monville 2019; Schwartz et al. 2015).

Neurodegenerative Diseases

Neurodegenerative diseases are a heterogeneous group that can have a hereditary or a spontaneous origin, which all culminate in the loss of neurons or glia cells in the CNS. Stem cell-based therapy is emerging as a promising candidate to treat some of these diseases including Parkinson's disease (PD), Alzheimer's disease, Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), among others. Cell replacement therapy was already performed as a proof of concept in some of these conditions.

PD is the second most common degenerative disease of the CNS, affecting 1% of the population above 60 years (Tysnes and Storstein 2017). PD is characterized by the degeneration of the dopaminergic neurons (DAn) in the *substantia nigra pars compacta* (SNpc), causing the characteristic motor factors of PD like rigidity, bradykinesia, tremors, and postural instability, along with other non-motor symptoms. Another hallmark of PD is the presence of aggregated α-synuclein in the Lewy bodies and Lewy neurites in the substantia nigra. The loss of DAn is initially restricted to the SNpc, but it will progressively spread to other regions of the brain during the development of the disease. The exact etiology of PD remains unknown, and nowadays it is considered to be a multifactorial disease resulting from a complex interplay between genetic factors and environmental cues. Currently, the available pharmacological regiments for PD only treat the symptoms of the disease and not the underlying neurodegeneration process (Antony et al. 2013; Kalia and Lang 2015). Therefore, new therapeutic strategies are needed, with stem cell-based therapy being the most promising approach to slow or even stop the progression of this devastating disease. The first cell therapy trials for PD were performed more than 30 years ago, using fetal ventral mesencephalic cells (fVM), a source of DAn. These first trials showed long-term efficacy and survival. However, it has been reported that in some cases the transplanted cells exhibited disease-related pathology and there was a huge variability in the clinical response probably due to the variability of the cells transplanted (Barker et al. 2015). In order to standardize the approach of cell therapy for PD, TRANSEURO (NCT01898390), a European consortium conducting phase I clinical trials to investigate the benefits of transplantation of allogeneic dopaminergic neuroblasts derived from fVM into PD patients, was created.

High expectations for future PD treatments lie on emerging therapies based on hPSC-derived DAns. The preferential use of iPSCs over ESCs has been fostered by the systematic banking of HLA-matched iPSCs, avoiding the need for immunosuppression regime in the recipient. The development of good manufacturing practice (GMP)-grade protocols to obtain authentic and functional midbrain DAn cells from hPSCs lead to the creation of the global network initiative GForce-PD. This network encompasses different consortia like the European NeuroStemCellRepair and the NYSTEM-PD based on ESC-derived DAn and the Japanese Center for iPS Cell Research and Application (CiRA) based on allogeneic transplantation from HLA-matched iPSC-derived DAn cells. As part of the GForce-PD network, the first clinical trial using iPSC-derived DAns to treat PD started in August 2018 at Kyoto University by a team headed by Takahashi as part of the Japanese Haplobank CiRA (UMIN000033564). Furthermore, some clinical trials outside the GForce-PD have already started. The Chinese Academy of Sciences is using HLA-matched ESCs (NCT03119636), while an Australian Clinical trial (NCT02452723) sponsored by Cyto Therapeutics Pty Limited is based on human parthenogenetic stem cell-derived neural cells (Barker et al. 2019; Wenker and Pitossi 2019).

With more iPSC-derived cell therapies to treat PD already planned, the follow-up of these trials will likely be beneficial, not only for PD but also to develop new approaches for other neurodegenerative diseases.

ALS is a neurodegenerative disorder characterized by progressive degeneration of the upper and

lower motor neurons. ALS has a devastating impact on the patient, and although the pathogenic basis is still unclear, stem cell-based therapy seems to have potential. The motor neuronal degeneration observed in ALS may arise from defective trophic support from astrocytes and microglia, whose replacement appears technically more feasible. Furthermore, the delivery of growth factors or extracellular vesicles might exert a neuroprotective effect.

Injection of human NSCs (NSI-566) originating from the spinal cord of an 8-week-old aborted fetus was performed in a phase I (NCT01348451) and phase II (NCT01730716) clinical trial. The results from these trials showed improved functional status scores of these patients, although survival did not differ from control groups. A phase 3 clinical trial is already planned (Goutman et al. 2018; Abati et al. 2019). Another trial using fetal tissue-derived NSCs in combination with gene therapy was approved in 2016. The human NSCs are engineered to produce Glial cell-derived neurotrophic factor (GDNF) (CNS10-NPC-GDNF) and injected in the spinal cord (NCT02943850). It is expected that these cells differentiate into neuroprotective astrocytes, increasing the survival of motor neurons in ALS patients. With more preclinical data suggesting that the malfunctioning of astrocytes is involved in ALS pathogenesis, a new clinical trial using ESC-derived astrocytes (NCT03482050) was approved in 2018. The results from these trials will give us insights about the safety and potential of PSCs in the treatment of ALS.

Beside these treatments, which are in the trial state, the Korean company Corestem has obtained approval for their product Neuronata-R in South Korea in 2014 after successful phase I and II clinical trials (NCT01363401). The treatment is based on repeated intrathecal injections of autologous BM-MSCs (Oh et al. 2015). In 2018 it also got permission as an orphan drug from the US FDA and EMA followed this classification in 2019.

Spinal Cord Injury (SCI)

SCI is one of the most devastating neurological injuries, including motor, sensory, and autonomic paralysis, with a prevalence (2016) of 27.04 million cases (GBD 2016 et al. 2019). SCI results from dislocation or fracture of the spine in the cervical of thoracic region and can be divided into two different phases, the acute and the chronic phases. In the acute phase, neurons and glia cells can undergo fast necrosis seconds to minutes after the injury (primary injury). The secondary injuries, depending on the extent of the primary injury, comprise the most critical phase. A cascade of destructive events occurs, including aberrant molecular signalling, inflammation, vascular changes, and secondary cellular dysfunctions. Weeks to months after the secondary injuries, the chronic phase takes place with scar formation as a consequence of the reactive gliolysis, formation of cysts, and gray matter demyelination (Cofano et al. 2019). Treatments for acute SCI are limited to the stabilization of the injured spine followed by rehabilitation in the chronic phase. Although some early medical treatments accounted for some significance recovery, there is no effective treatment.

The development of cell transplantation therapies to replace the damaged cells and to promote neural protection and regeneration of the injured spinal cord has shown great potential. A variety of different cell types have already been transplanted into injured spinal cord, in rodents and human, like olfactory ensheathing cells, oligodendrocyte precursor cells (OPCs), Schwann cells, NSCs derived from fetal tissue, and MSCs. In particular, from the 18 clinical trials conducted, 16 were based on MSCs. The big disadvantage of using MSCs for transplantation is their restricted differentiation potential. To overcome this, the first clinical trial using ESC-derived OPCs was approved in 2010 (NCT01217008). This study was terminated based on financial issues, and a second clinical trial was approved in 2015 (NCT02302157). This is the only PSCs-based clinical trial currently active, and the results are expected to be published soon (Csobonyeiova et al. 2019; Pereira et al. 2019). Recently the beneficial effects of iPSC-derived NSCs transplantation for tissue recovery were shown. After transplantation these cells can differentiate into neurons and oligodendrocytes, myelinate the host axons, and secrete neurotrophic factors to

prevent the secondary damage. In February 2019, the first clinical trial using allogenic iPSC-derived NSCs was approved. The study will be performed on four patients at the Keio University School of Medicine by a research team headed by Prof. Hideyuki Okano (Nagoshi et al. 2019; Tsuji et al. 2019)

Future Perspectives in Stem Cell-Based Therapy of the CNS

Although some neurodegenerative diseases share pathological mechanisms such as protein aggregation, misfolded proteins, and loss of local neurons, the underlying molecular mechanisms that lead to the CNS's progressive degeneration are unknown. There are still considerable problems in treating these diseases, and so far they do not make good targets for stem cell therapy, especially because the neurodegeneration is too widespread, diffuse, and currently irreversible. Alzheimer's disease, the most common and devastating form of dementia, is so far not a target for stem cell-based replacement therapy due to its high complexity, heterogeneity, and unclear pathogenesis. Other diseases are still not a target of stem cell-based clinical trials, but preclinical studies have shown the potential of such therapy. Huntington's disease(HD) is a fatal autosomal-dominant disease characterized by motor dysfunction, cognitive impairment, and psychiatric disturbances due to neurodegeneration. HD is caused by a mutation of the huntingtin (HTT) gene, and stem-cell based therapy is a potential strategy to restore neuronal function, replacing the lost neurons, and provide neurotrophic support. Several clinical trials using human fetal tissue to treat HD have been performed since 1990. The results from these very heterogeneous trials indicated that fetal transplantation in HD is safe but the clinical improvement did not last more than 4–6 years after transplantation. Finding an alternative to the fetal striatal graft transplantation is being a focus of research, with a set of preclinical studies based on PSCs ongoing based on PSCs. The results from these studies using PSC-based therapy are largely inconsistent. Thus, more preclinical work is necessary to confirm if PSCs have a therapeutic potential to treat HD.

Another disease where stem cell-based therapy is a potential target is ischemic stroke, an acute cerebrovascular disease caused by a decreased or interrupted blood supply in the brain. Ischemic stroke destroys a heterogeneous population of brain cells, making cell replacement therapy a difficult approach since neurons and glia cells that should be replaced are affected. Therefore, most clinical trials shifted to use MSCs in order to suppress the post-ischemic inflammatory response. The results from these clinical trials showed an improvement of the neurological symptoms in the patients. Although MSCs are an attractive and promising therapy, it is not a cell replacement therapy. The UK started a clinical trial already in phase II (NCT03629275) where an immortalized human NSC line (CTX0E03) is being used. Although (HLA-matched) iPSCs are still considered as a potential source for stem cell therapy to treat stroke, where the results from preclinical studies have been encouraging, additional studies and improvement in the strategies for cell replacement therapy are required.

Summary

A tremendous number of stem cell-based clinical trials (also for organs which we could not mention in this book chapter) are currently underway. Many of these focus on MSCs which can be applied in a wide variety of settings, to support regeneration, modulate the immune response, and maybe even restore damaged tissue (Fig. 4). As they do not have to be applied in an HLA-matched manner, allogeneic transplantation is facilitated. Cells differentiated from PSCs are currently emerging as treatment options for various diseases. Here, the hurdles for transplantation are higher because the possibility of immune rejection requires HLA matching and the tumorigenic potential of PSCs needs to be controlled.

While HSC and keratinocyte transplantation have been in the clinic for a long time, only selected MSC-based treatments have been approved in a limited number of countries. It

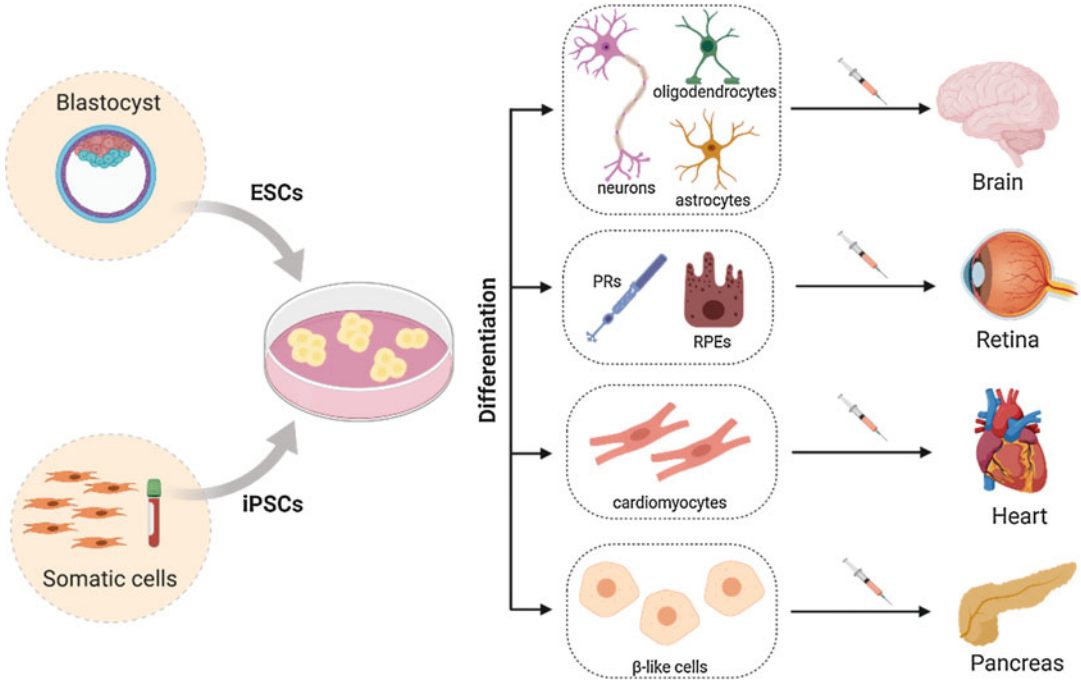


Fig. 4 Overview of the ongoing PSC-based clinical trials. PSCs isolated from the blastocyst (ESCs) or reprogrammed from somatic cells (iPSCs) are differentiated in vitro to generate neurons, oligodendrocytes, and astrocytes to treat central nervous system disorders, PRs and RPEs to treat retinal degenerative diseases, cardiomyocytes to treat heart diseases, and β-like cells to treat type 1 diabetes. After differentiation the cells are transplanted into the brain, retina, heart, or pancreas from the donor. Figure made in ©BioRender – biorender.com

will probably take much longer until PSC-based therapies will be available to a broad spectrum of patients. These novel treatments will need a close surveillance until we have fully understood their risks and benefits.

References and Further Reading

Abati E, Bresolin N, Comi GP, Corti S (2019) Preconditioning and cellular engineering to increase the survival of transplanted neural stem cells for motor neuron disease therapy. Mol Neurobiol 56(5):3356–3367. https://doi.org/10.1007/s12035-018-1305-4

Abu-Dawud R, Graffmann N, Ferber S, Wruck W, Adjaye J (2018) Pluripotent stem cells: induction and self-renewal. Philos Trans R Soc Lond Ser B Biol Sci 373(1750). https://doi.org/10.1098/rstb.2017.0213

alliancerm.org. https://alliancerm.org/available-products/. Accessed 18 Sept 2019

Al-Saif AM (2019) Gene therapy of hematological disorders: current challenges. Gene Ther. https://doi.org/10.1038/s41434-019-0093-4

Altman J, Das GD (1965) Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. J Comp Neurol 124:319–335. https://doi.org/10.1002/cne.901240303

Antony PMA, Diederich NJ, Krüger R, Balling R (2013) The hallmarks of Parkinson's disease. FEBS J 280:5981–5993. https://doi.org/10.1111/febs.12335

Ascheim DD, Gelijns AC, Goldstein D, Moye LA, Smedira N, Lee S, Klodell CT, Szady A, Parides MK, Jeffries NO, Skerrett D, Taylor DA, Rame JE, Milano C, Rogers JG, Lynch J, Dewey T, Eichhorn E, Sun B, Feldman D, Simari R, O'Gara PT, Taddei-Peters WC, Miller MA, Naka Y, Bagiella E, Rose EA, Woo YJ (2014) Mesenchymal Precursor Cells as Adjunctive Therapy in Recipients of Contemporary Left Ventricular Assist Devices. Circulation 129 (22):2287–2296. https://doi.org/10.1161/circulationaha.113.007412

Ballen KK, Verter F, Kurtzberg J (2015) Umbilical cord blood donation: public or private? Bone Marrow Transplant 50(10):1271–1278. https://doi.org/10.1038/bmt.2015.124

Barker RA, Drouin-Ouellet J, Parmar M (2015) Cell-based therapies for Parkinson disease – past insights and future potential. Nat Rev Neurol 11:492–503. https://doi.org/10.1038/nrneurol.2015.123

Barker RA, Farrell K, Valle Guzman N, He X, Lazic SE, Moore S, Morris R, Tyers P, Wijeyekoon R, Daft D, Hewitt S, Dayal V, Foltynie T, Kefalopoulou Z, Mahlknecht P, Lao-Kaim NP, Piccini P, Bjartmarz H, Björklund A, Lindvall O, Nelander-Wahlestedt J, Parmar M, Paul G, Widner H, Church A, Dunnett S, Peall K, Rosser A, Marc Gurruchaga J, Palfi S, Piroth T, Winkler C (2019) Designing stem-cell-based dopamine cell replacement trials for Parkinson's disease. Nat Med 2:1045–1053. https://doi.org/10.1038/s41591-019-0507-2

Barrandon Y, Green H (1987) Three clonal types of keratinocyte with different capacities for multiplication. Proc Natl Acad Sci U S A 84(8):2302–2306. https://doi.org/10.1073/pnas.84.8.2302

Bauer JW, Koller J, Murauer EM, De Rosa L, Enzo E, Carulli S, Bondanza S, Recchia A, Muss W, Diem A, Mayr E, Schlager P, Gratz IK, Pellegrini G, De Luca M (2017) Closure of a large chronic wound through transplantation of gene-corrected epidermal stem cells. J Invest Dermatol 137(3):778–781. https://doi.org/10.1016/j.jid.2016.10.038

Ben M'Barek K, Monville C (2019) Cell therapy for retinal dystrophies: from cell suspension formulation to complex retinal tissue bioengineering. Stem Cells Int 2019:1–14. https://doi.org/10.1155/2019/4568979

Caplan AI (1991) Mesenchymal stem cells. J Orthop Res 9(5):641–650. https://doi.org/10.1002/jor.1100090504

Caplan H, Olson SD, Kumar A, George M, Prabhakara KS, Wenzel P, Bedi S, Toledano-Furman NE, Triolo F, Kamhieh-Milz J, Moll G, Cox CS Jr (2019) Mesenchymal stromal cell therapeutic delivery: translational challenges to clinical application. Front Immunol 10:1645. https://doi.org/10.3389/fimmu.2019.01645

Chagastelles PC, Nardi NB (2011) Biology of stem cells: an overview. Kidney Int Suppl 1(3):63–67. https://doi.org/10.1038/kisup.2011.15

Champlin R (2003) Selection of autologous or allogeneic transplantation. In: Holland-Frei cancer medicine, vol 6. BC Decker, Hamilton

Cho J, D'Antuono M, Glicksman M, Wang J, Jonklaas J (2018) A review of clinical trials: mesenchymal stem cell transplant therapy in type 1 and type 2 diabetes mellitus. Am J Stem Cells 7(4):82–93

Cofano F, Boido M, Monticelli M, Zenga F, Ducati A, Vercelli A, Garbossa D (2019) Mesenchymal stem cells for spinal cord injury: current options, limitations, and future of cell therapy. Int J Mol Sci 20:2698. https://doi.org/10.3390/ijms20112698

Comella K, Parlo M, Daly R, Dominessy K (2018) First-in-man intravenous implantation of stromal vascular fraction in psoriasis: a case study. Int Med Case Rep J 11:59–64. https://doi.org/10.2147/IMCRJ.S163612

Connor B (2018) Concise review: the use of stem cells for understanding and treating Huntington's disease. Stem Cells 36:146–160. https://doi.org/10.1002/stem.2747

Csobonyeiova M, Polak S, Zamborsky R, Danisovic L (2019) Recent Progress in the regeneration of spinal cord injuries by induced pluripotent stem cells. Int J Mol Sci 20:3838. https://doi.org/10.3390/ijms20153838

Di WL, Lwin SM, Petrova A, Bernadis C, Syed F, Farzaneh F, Moulding D, Martinez AE, Sebire NJ, Rampling D, Virasami A, Zamiri M, Wang W, Hara H, Kadiyirire T, Abdul-Wahab A, Martinez-Queipo M, Harper JI, McGrath JA, Thrasher AJ, Mellerio JE, Qasim W (2019) Generation and clinical application of gene-modified autologous epidermal sheets in netherton syndrome: lessons learned from a phase 1 trial. Hum Gene Ther. https://doi.org/10.1089/hum.2019.049

Diabetes Facts & Figures (2019) International Diabetes Federation. https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html. Accessed 5 Aust 2019

Dominici M, Le Blanc K, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, Deans R, Keating A, Prockop D, Horwitz E (2006) Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. Cytotherapy 8(4):315–317. https://doi.org/10.1080/14653240600855905

Dowse R, McLornan DP (2017) Principles of haemopoietic stem cell transplantation. Medicine 45(5):322–326. https://doi.org/10.1016/j.mpmed.2017.02.009

Elstner A, Damaschun A, Kurtz A, Stacey G, Aran B, Veiga A, Borstlap J (2009) The changing landscape of European and international regulation on embryonic stem cell research. Stem Cell Res 2(2):101–107. https://doi.org/10.1016/j.scr.2008.10.003

Fitzsimmons REB, Mazurek MS, Soos A, Simmons CA (2018) Mesenchymal stromal/stem cells in regenerative medicine and tissue engineering. Stem Cells Int 2018:8031718. https://doi.org/10.1155/2018/8031718

Friedenstein AJ, Chailakhjan RK, Lalykina KS (1970) The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. Cell Tissue Kinet 3(4)

Frobel J, Hemeda H, Lenz M, Abagnale G, Joussen S, Denecke B, Saric T, Zenke M, Wagner W (2014) Epigenetic rejuvenation of mesenchymal stromal cells derived from induced pluripotent stem cells. Stem Cell Rep 3(3):414–422. https://doi.org/10.1016/j.stemcr.2014.07.003

Galipeau J, Sensebe L (2018) Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. Cell Stem Cell 22(6):824–833. https://doi.org/10.1016/j.stem.2018.05.004

GBD 2016, Traumatic Brain Injury and Spinal Cord Injury Collaborators SL, Theadom A, Ellenbogen RG, Bannick MS, Montjoy-Venning W, Lucchesi LR, Abbasi N, Abdulkader R, Abraha HN, Adsuar JC, Afarideh M, Agrawal S, Ahmadi A, Ahmed MB, Aichour AN, Aichour I, Aichour MTE, Akinyemi RO, Akseer N, Alahdab F, Alebel A, Alghnam SA, Ali BA, Alsharif U, Altirkawi K, Andrei CL, Anjomshoa M, Ansari H, Ansha MG, Antonio CAT, Appiah SCY, Ariani F, Asefa NG, Asgedom SW, Atique S, Awasthi A, Quintanilla BPA, Ayuk TB,

Azzopardi PS, Badali H, Badawi A, Balalla S, Banstola A, Barker-Collo SL, Bärnighausen TW, Bedi N, Behzadifar M, Behzadifar M, Bekele BB, Belachew AB, Belay YA, Bennett DA, Bensenor IM, Berhane A, Beuran M, Bhalla A, Bhaumik S, Bhutta ZA, Biadgo B, Biffino M, Bijani A, Bililign N, Birungi C, Boufous S, Brazinova A, Brown AW, Car M, Cárdenas R, Carrero JJ, Carvalho F, Castañeda-Orjuela CA, Catalá-López F, Chaiah Y, Champs AP, Chang J-C, Choi J-YJ, Christopher DJ, Cooper C, Crowe CS, Dandona L, Dandona R, Daryani A, Davitoiu DV, Degefa MG, Demoz GT, Deribe K, Djalalinia S, Do HP, Doku DT, Drake TM, Dubey M, Dubljanin E, El-Khatib Z, Ofori-Asenso R, Eskandarieh S, Esteghamati A, Esteghamati S, Faro A, Farzadfar F, Farzaei MH, Fereshtehnejad S-M, Fernandes E, Feyissa GT, Filip I, Fischer F, Fukumoto T, Ganji M, Gankpe FG, Gebre AK, Gebrehiwot TT, Gezae KE, Gopalkrishna G, Goulart AC, Haagsma JA, Haj-Mirzaian A, Haj-Mirzaian A, Hamadeh RR, Hamidi S, Haro JM, Hassankhani H, Hassen HY, Havmoeller R, Hawley C, Hay SI, Hegazy MI, Hendrie D, Henok A, Hibstu DT, Hoffman HJ, Hole MK, Rad EH, Hosseini SM, Hostiuc S, Hu G, Hussen MA, Ilesanmi OS, Irvani SSN, Jakovljevic M, Jayaraman S, Jha RP, Jonas JB, Jones KM, Shushtari ZJ, Jozwiak JJ, Jürisson M, Kabir A, Kahsay A, Kahssay M, Kalani R, Karch A, Kasaeian A, Kassa GM, Kassa TD, Kassa ZY, Kengne AP, Khader YS, Khafaie MA, Khalid N, Khalil I, Khan EA, Khan MS, Khang Y-H, Khazaie H, Khoja AT, Khubchandani J, Kiadaliri AA, Kim D, Kim Y-E, Kisa A, Koyanagi A, Krohn KJ, Defo BK, Bicer BK, Kumar GA, Kumar M, Lalloo R, Lami FH, Lansingh VC, Laryea DO, Latifi A, Leshargie CT, Levi M, Li S, Liben ML, Lotufo PA, Lunevicius R, Mahotra NB, Majdan M, Majeed A, Malekzadeh R, Manda A-L, Mansournia MA, Massenburg BB, Mate KKV, Mehndiratta MM, Mehta V, Meles H, Melese A, Memiah PTN, Mendoza W, Mengistu G, Meretoja A, Meretoja TJ, Mestrovic T, Miazgowski T, Miller TR, Mini G, Mirica A, Mirrakhimov EM, Moazen B, Mohammadi M, Molokhia M, Monasta L, Mondello S, Moosazadeh M, Moradi G, Moradi M, Moradi-Lakeh M, Moradinazar M, Morrison SD, Moschos MM, Mousavi SM, Murthy S, Musa KI, Mustafa G, Naghavi M, Naik G, Najafi F, Nangia V, Nascimento BR, Negoi I, Nguyen TH, Nichols E, Ningrum DNA, Nirayo YL, Nyasulu PS, Ogbo FA, Oh I-H, Okoro A, Olagunju AT, Olagunju TO, Olivares PR, Otstavnov SS, Owolabi MO, Mahesh PA, Pakhale S, Pandey AR, Pesudovs K, Pinilla-Monsalve GD, Polinder S, Poustchi H, Prakash S, Qorbani M, Radfar A, Rafay A, Rafiei A, Rahimi-Movaghar A, Rahimi-Movaghar V, Rahman M, Rahman MA, Rai RK, Rajati F, Ram U, Rawaf DL, Rawaf S, Reiner RC, Reis C, Renzaho AMN, Resnikoff S, Rezaei S, Rezaeian S, Roever L, Ronfani L, Roshandel G, Roy N, Ruhago GM, Saddik B, Safari H, Safiri S, Sahraian MA, Salamati P, Saldanha RF, Samy AM, Sanabria J, Santos JV, Milicevic MMMS, Sartorius B, Satpathy M, Savuon K, Schneider IJC, Schwebel DC, Sepanlou SG, Shabaninejad H, Shaikh MAA, Shams-Beyranvand M, Sharif M, Sharif-Alhoseini M, Islam SMS, She J, Sheikh A, Shen J, Sheth KN, Shibuya K, Shiferaw MS, Shigematsu M, Shiri R, Shiue I, Shoman H, Siabani S, Siddiqi TJ, Silva JP, Silveira DGA, Sinha DN, Smith M, Filho AMS, Sobhani S, Soofi M, Soriano JB, Soyiri IN, Stein DJ, Stokes MA, Sufiyan MB, Sunguya BF, Sunshine JE, Sykes BL, Szoeke CEI, Tabarés-Seisdedos R, Ao BJT, Tehrani-Banihashemi A, Tekle MG, Temsah M-H, Temsah O, Topor-Madry R, Tortajada-Girbés M, Tran BX, Tran KB, Car LT, Ukwaja KN, Ullah I, Usman MS, Uthman OA, Valdez PR, Vasankari TJ, Venketasubramanian N, Violante FS, Wagnew FWS, Waheed Y, Wang Y-P, Weldegwergs KG, Werdecker A, Wijeratne T, Winkler AS, Wyper GMA, Yano Y, Yaseri M, Yasin YJ, Ye P, Yimer EM, Yip P, Yisma E, Yonemoto N, Yoon S-J, Yost MG, Younis MZ, Yousefifard M, Yu C, Zaidi Z, Zaman SB, Zamani M, Zenebe ZM, Zodpey S, Feigin VL, Vos T, Murray CJL (2019) Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the global burden of disease study 2016. Lancet Neurol 18:56–87. https://doi.org/10.1016/S1474-4422(18)30415-0

Goff ZD, Kichura AB, Chibnall JT, Hauptman PJ (2017) A survey of unregulated direct-to-consumer treatment centers providing stem cells for patients with heart failure. JAMA Intern Med 177(9). https://doi.org/10.1001/jamainternmed.2017.2988

Goldman SA (2016) Stem and progenitor cell-based therapy of the central nervous system: hopes, hype, and wishful thinking. Cell Stem Cell. https://doi.org/10.1016/j.stem.2016.01.012

Gonzalez R, Hamblin MH, Lee J-P (2016) Neural stem cell transplantation and CNS diseases. CNS Neurol Disord Drug Targets 15:881–886

Gotherstrom C, Westgren M, Shaw SW, Astrom E, Biswas A, Byers PH, Mattar CN, Graham GE, Taslimi J, Ewald U, Fisk NM, Yeoh AE, Lin JL, Cheng PJ, Choolani M, Le Blanc K, Chan JK (2014) Pre- and postnatal transplantation of fetal mesenchymal stem cells in osteogenesis imperfecta: a two-center experience. Stem Cells Transl Med 3(2):255–264. https://doi.org/10.5966/sctm.2013-0090

Goutman SA, Brown MB, Glass JD, Boulis NM, Johe K, Hazel T, Cudkowicz M, Atassi N, Borges L, Patil PG, Sakowski SA, Feldman EL (2018) Long-term phase 1/2 intraspinal stem cell transplantation outcomes in ALS. Ann Clin Transl Neurol 5:730–740. https://doi.org/10.1002/acn3.567

Guenou H, Nissan X, Larcher F, Feteira J, Lemaitre G, Saidani M, Del Rio M, Barrault CC, Bernard FX, Peschanski M, Baldeschi C, Waksman G (2009) Human embryonic stem-cell derivatives for full reconstruction of the pluristratified epidermis: a preclinical study. Lancet 374(9703):1745–1753. https://doi.org/10.1016/S0140-6736(09)61496-3

Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, Salgado M, Martinez-Picado J, Nijhuis M, Wensing AMJ, Lee H, Grant P, Nastouli E, Lambert J, Pace M, Salasc F, Monit C, Innes AJ, Muir L, Waters L, Frater J, Lever AML, Edwards SG, Gabriel IH, Olavarria E (2019) HIV-1 remission following CCR5Delta32/Delta32 haematopoietic stem-cell transplantation. Nature 568(7751):244–248. https://doi.org/10.1038/s41586-019-1027-4

Gyurkocza B, Storb R, Storer BE, Chauncey TR, Lange T, Shizuru JA, Langston AA, Pulsipher MA, Bredeson CN, Maziarz RT, Bruno B, Petersen FB, Maris MB, Agura E, Yeager A, Bethge W, Sahebi F, Appelbaum FR, Maloney DG, Sandmaier BM (2010) Non-myeloablative allogeneic hematopoietic cell transplantation in patients with acute myeloid leukemia. J Clin Oncol 28(17):2859–2867. https://doi.org/10.1200/JCO.2009.27.1460

Hainzl S, Peking P, Kocher T, Murauer EM, Larcher F, Del Rio M, Duarte B, Steiner M, Klausegger A, Bauer JW, Reichelt J, Koller U (2017) COL7A1 editing via CRISPR/Cas9 in recessive dystrophic epidermolysis bullosa. Mol Ther 25(11):2573–2584. https://doi.org/10.1016/j.ymthe.2017.07.005

Henig I, Zuckerman T (2014) Hematopoietic stem cell transplantation-50 years of evolution and future perspectives. Rambam Maimonides Med J 5(4):e0028. https://doi.org/10.5041/RMMJ.10162

Hernon CA, Dawson RA, Freedlander E, Short R, Haddow DB, Brotherston M, MacNeil S (2006) Clinical experience using cultured epithelial autografts leads to an alternative methodology for transferring skin cells from the laboratory to the patient. Regen Med 1(6):809–821. https://doi.org/10.2217/17460751.1.6.809

Hirsch T, Rothoeft T, Teig N, Bauer JW, Pellegrini G, De Rosa L, Scaglione D, Reichelt J, Klausegger A, Kneisz D, Romano O, Secone Seconetti A, Contin R, Enzo E, Jurman I, Carulli S, Jacobsen F, Luecke T, Lehnhardt M, Fischer M, Kueckelhaus M, Quaglino D, Morgante M, Bicciato S, Bondanza S, De Luca M (2017) Regeneration of the entire human epidermis using transgenic stem cells. Nature 551(7680):327–332. https://doi.org/10.1038/nature24487

Horwitz EM, Prockop DJ, Fitzpatrick LA, Koo WW, Gordon PL, Neel M, Sussman M, Orchard P, Marx JC, Pyeritz RE, Brenner MK (1999) Transplantability and therapeutic effects of bone marrow-derived mesenchymal cells in children with osteogenesis imperfecta. Nat Med 5(3):309–313. https://doi.org/10.1038/6529

Hu C, Zhao L, Li L (2019) Current understanding of adipose-derived mesenchymal stem cell-based therapies in liver diseases. Stem Cell Res Ther 10(1):199. https://doi.org/10.1186/s13287-019-1310-1

Igarashi P, Somlo S (2007) Polycystic kidney disease. J Am Soc Nephrol 18(5):1371–1373. https://doi.org/10.1681/ASN.2007030299

Iijima H, Isho T, Kuroki H, Takahashi M, Aoyama T (2018) Effectiveness of mesenchymal stem cells for treating patients with knee osteoarthritis: a meta-analysis toward the establishment of effective regenerative rehabilitation. NPJ Regen Med 3:15. https://doi.org/10.1038/s41536-018-0041-8

Inoue H, Nagata N, Kurokawa H, Yamanaka S (2014) iPS cells: a game changer for future medicine. EMBO J 33(5):409–417. https://doi.org/10.1002/embj.201387098

Iriguchi S, Kaneko S (2019) Toward the development of true "off-the-shelf" synthetic T-cell immunotherapy. Cancer Sci 110(1):16–22. https://doi.org/10.1111/cas.13892

Izmiryan A, Ganier C, Bovolenta M, Schmitt A, Mavilio F, Hovnanian A (2018) Ex vivo COL7A1 correction for recessive dystrophic epidermolysis bullosa using CRISPR/Cas9 and homology-directed repair. Mol Ther Nucleic Acids 12:554–567. https://doi.org/10.1016/j.omtn.2018.06.008

JCR Pharmaceuticals Co. L (2017) Annual Report. https://ssl4.eir-parts.net/doc/4552/ir_material_for_fiscal_ym4/63251/00.pdf

Jessop H, Farge D, Saccardi R, Alexander T, Rovira M, Sharrack B, Greco R, Wulffraat N, Moore J, Kazmi M, Badoglio M, Adams G, Verhoeven B, Murray J, Snowden JA (2019) General information for patients and carers considering haematopoietic stem cell transplantation (HSCT) for severe autoimmune diseases (ADs): a position statement from the EBMT Autoimmune Diseases Working Party (ADWP), the EBMT Nurses Group, the EBMT Patient, Family and Donor Committee and the Joint Accreditation Committee of ISCT and EBMT (JACIE). Bone Marrow Transplant 54(7):933–942. https://doi.org/10.1038/s41409-019-0430-7

Jungbluth P, Spitzhorn LS, Grassmann J, Tanner S, Latz D, Rahman MS, Bohndorf M, Wruck W, Sager M, Grotheer V, Kropil P, Hakimi M, Windolf J, Schneppendahl J, Adjaye J (2019) Human iPSC-derived iMSCs improve bone regeneration in mini-pigs. Bone Res 7 (1):32. https://doi.org/10.1038/s41413-019-0069-4

Kalia LV, Lang AE (2015) Parkinson's disease. Lancet (London, UK) 386:896–912. https://doi.org/10.1016/S0140-6736(14)61393-3

Karimkhani C, Dellavalle RP, Coffeng LE, Flohr C, Hay RJ, Langan SM, Nsoesie EO, Ferrari AJ, Erskine HE, Silverberg JI, Vos T, Naghavi M (2017) Global skin disease morbidity and mortality: an update from the global burden of disease study 2013. JAMA Dermatol 153(5):406–412. https://doi.org/10.1001/jamadermatol.2016.5538

Kharroubi AT, Darwish HM (2015) Diabetes mellitus: the epidemic of the century. World J Diabetes 6(6):850–867. https://doi.org/10.4239/wjd.v6.i6.850

Kim JH, Kurtz A, Yuan BZ, Zeng F, Lomax G, Loring JF, Crook J, Ju JH, Clarke L, Inamdar MS, Pera M, Firpo MT, Sheldon M, Rahman N, O'Shea O, Pranke P, Zhou Q, Isasi R, Rungsiwiwut R, Kawamata S, Oh S, Ludwig T, Masui T, Novak TJ, Takahashi T, Fujibuchi W, Koo SK, Stacey GN (2017) Report of the international stem cell banking initiative workshop activity: current hurdles and progress in seed-stock banking of human pluripotent stem cells. Stem Cells Transl Med 6(11):1956–1962. https://doi.org/10.1002/sctm.17-0144

Klein J, Sato A (2000) The HLA system. N Engl J Med 434 (10):702–709. https://doi.org/10.1056/NEJM200009073431006

Kobayashi K, Stamm C, Nakauchi A (2019) First iPSC-Derived CAR T-Cell Therapy Created by Kyoto University CiRA and Takeda Collaboration Enters Process Development Toward Clinical Testing. https://www.takeda.com/newsroom/newsreleases/2019/first-ipsc-derived-car-t-cell-therapy-created-by-kyoto-university-cira-and%2D%2Dtakeda-collaboration-enters-process-development-toward-clinical-testing/. Accessed 6 Sept 2019

Kurtzberg J, Prockop S, Teira P, Bittencourt H, Lewis V, Chan KW, Horn B, Yu L, Talano JA, Nemecek E, Mills CR, Chaudhury S (2014) Allogeneic human mesenchymal stem cell therapy (remestemcel-L, Prochymal) as a rescue agent for severe refractory acute graft-versus-host disease in pediatric patients. Biol Blood Marrow Transplant 20(2):229–235. https://doi.org/10.1016/j.bbmt.2013.11.001

Lazarus HM, Haynesworth SE, Gerson SL, Rosenthal NS, Caplan AI (1995) Ex vivo expansion and subsequent infusion of human bone marrow-derived stromal progenitor cells (mesenchymal progenitor cells): implications for therapeutic use. Bone Marrow Transplant 16(4)

Little MT, Storb R (2002) History of haematopoietic stem-cell transplantation. Nat Rev Cancer 2(3):231–238. https://doi.org/10.1038/nrc748

Makhlough A, Shekarchian S, Moghadasali R, Einollahi B, Hosseini SE, Jaroughi N, Bolurieh T, Baharvand H, Aghdami N (2017) Safety and tolerability of autologous bone marrow mesenchymal stromal cells in ADPKD patients. Stem Cell Res Ther 8(1):116. https://doi.org/10.1186/s13287-017-0557-7

Malliaras K, Makkar RR, Smith RR, Cheng K, Wu E, Bonow RO, Marbán L, Mendizabal A, Cingolani E, Johnston PV, Gerstenblith G, Schuleri KH, Lardo AC, Marbán E (2014) Intracoronary cardiosphere-derived cells after myocardial infarction. J Am Coll Cardiol 63(2):110–122. https://doi.org/10.1016/j.jacc.2013.08.724

Marks PW, Witten CM, Califf RM (2017) Clarifying stem-cell Therapy's benefits and risks. N Engl J Med 376 (11):1007–1009. https://doi.org/10.1056/NEJMp1613723

Mavilio F, Pellegrini G, Ferrari S, Di Nunzio F, Di Iorio E, Recchia A, Maruggi G, Ferrari G, Provasi E, Bonini C, Capurro S, Conti A, Magnoni C, Giannetti A, De Luca M (2006) Correction of junctional epidermolysis bullosa by transplantation of genetically modified epidermal stem cells. Nat Med 12(12):1397–1402. https://doi.org/10.1038/nm1504

Maximow AA (1924) Relation of blood cells to connective tissues and endothelium. Physiol Rev 4(4):533–563. https://doi.org/10.1152/physrev.1924.4.4.533

Menasche P, Vanneaux V, Hagege A, Bel A, Cholley B, Parouchev A, Cacciapuoti I, Al-Daccak R, Benhamouda N, Blons H, Agbulut O, Tosca L, Trouvin JH, Fabreguettes JR, Bellamy V, Charron D, Tartour E, Tachdjian G, Desnos M, Larghero J (2018) Transplantation of human embryonic stem cell-derived cardiovascular progenitors for severe ischemic left ventricular dysfunction. J Am Coll Cardiol 71(4):429–438. https://doi.org/10.1016/j.jacc.2017.11.047

Michler RE (2018) The current status of stem cell therapy in ischemic heart disease. J Card Surg 33(9):520–531. https://doi.org/10.1111/jocs.13789

Ming G, Song H (2011) Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron. https://doi.org/10.1016/j.neuron.2011.05.001

Nagoshi N, Tsuji O, Nakamura M, Okano H (2019) Cell therapy for spinal cord injury using induced pluripotent stem cells. Regen Ther 11:75–80. https://doi.org/10.1016/J.RETH.2019.05.006

O'Connor NE, Mulliken JB (1981) Grafting of burns with cultured epithelium prepared from autologous epidermal cells. Lancet 1(8211):75–78

Oh KW, Moon C, Kim HY, Oh SI, Park J, Lee JH, Chang IY, Kim KS, Kim SH (2015) Phase I trial of repeated intrathecal autologous bone marrow-derived mesenchymal stromal cells in amyotrophic lateral sclerosis. Stem Cells Transl Med 4(6):590–597. https://doi.org/10.5966/sctm.2014-0212

Panch SR, Szymanski J, Savani BN, Stroncek DF (2017) Sources of hematopoietic stem and progenitor cells and methods to optimize yields for clinical cell therapy. Biol Blood Marrow Transplant 23(8):1241–1249. https://doi.org/10.1016/j.bbmt.2017.05.003

Panés J, García-Olmo D, Van Assche G, Colombel JF, Reinisch W, Baumgart DC, Dignass A, Nachury M, Ferrante M, Kazemi-Shirazi L, Grimaud JC, de la Portilla F, Goldin E, Richard MP, Leselbaum A, Danese S (2016) Expanded allogeneic adipose-derived mesenchymal stem cells (Cx601) for complex perianal fistulas in Crohn's disease: a phase 3 randomised, double-blind controlled trial. Lancet 388(10051):1281–1290. https://doi.org/10.1016/s0140-6736(16)31203-x

Passweg JR, Baldomero H, Basak GW, Chabannon C, Corbacioglu S, Duarte R, Kuball J, Lankester A, Montoto S, de Latour RP, Snowden JA, Styczynski J, Yakoub-Agha I, Arat M, Mohty M, Kroger N, European Society for B, Marrow T (2019) The EBMT activity survey report 2017: a focus on allogeneic HCT for nonmalignant indications and on the use of non-HCT cell therapies. Bone Marrow Transplant. https://doi.org/10.1038/s41409-019-0465-9

Pellegrini G, Rama P, Mavilio F, De Luca M (2009) Epithelial stem cells in corneal regeneration and epidermal gene therapy. J Pathol 217(2):217–228. https://doi.org/10.1002/path.2441

Pereira IM, Marote A, Salgado AJ, Silva NA (2019) Filling the gap: neural stem cells as a promising therapy for spinal cord injury. Pharmaceuticals (Basel, Switzerland) 12. https://doi.org/10.3390/ph12020065

Petrof G, Abdul-Wahab A, McGrath JA (2014) Cell therapy in dermatology. Cold Spring Harb Perspect Med 4(6). https://doi.org/10.1101/cshperspect.a015156

Petrof G, Lwin SM, Martinez-Queipo M, Abdul-Wahab A, Tso S, Mellerio JE, Slaper-Cortenbach I, Boelens JJ, Tolar J, Veys P, Ofuya M, Peacock JL, Martinez AE, McGrath JA (2015) Potential of systemic allogeneic mesenchymal stromal cell therapy for children with recessive dystrophic epidermolysis bullosa. J Invest

Dermatol 135(9):2319–2321. https://doi.org/10.1038/jid.2015.158

Pullen LC (2018) Stem cell-derived pancreatic progenitor cells have now been transplanted into patients: report from IPITA 2018. Am J Transplant 18(7):1581–1582. https://doi.org/10.1111/ajt.14954

Rahman MS, Spitzhorn L-S, Wruck W, Hagenbeck C, Balan P, Graffmann N, Bohndorf M, Ncube A, Guillot PV, Fehm T, Adjaye J (2018) The presence of human mesenchymal stem cells of renal origin in amniotic fluid increases with gestational time. Stem Cell Res Ther 9(1). https://doi.org/10.1186/s13287-018-0864-7

Rama P, Matuska S, Paganoni G, Spinelli A, De Luca M, Pellegrini G (2010) Limbal stem-cell therapy and long-term corneal regeneration. N Engl J Med 363(2): 147–155. https://doi.org/10.1056/NEJMoa0905955

Rheinwald JG, Green H (1975) Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. Cell 6(3):331–343. https://doi.org/10.1016/s0092-8674(75)80001-8

Sakai D, Schol J, Foldager CB, Sato M, Watanabe M (2017) Regenerative technologies to bed side: evolving the regulatory framework. J Orthop Translat 9:1–7. https://doi.org/10.1016/j.jot.2017.02.001

Schachinger V, Erbs S, Elsasser A, Haberbosch W, Hambrecht R, Holschermann H, Yu J, Corti R, Mathey DG, Hamm CW, Suselbeck T, Werner N, Haase J, Neuzner J, Germing A, Mark B, Assmus B, Tonn T, Dimmeler S, Zeiher AM (2006) Improved clinical outcome after intracoronary administration of bone-marrow-derived progenitor cells in acute myocardial infarction: final 1-year results of the REPAIR-AMI trial. Eur Heart J 27(23):2775–2783. https://doi.org/10.1093/eurheartj/ehl388

Schwartz SD, Regillo CD, Lam BL, Eliott D, Rosenfeld PJ, Gregori NZ, Hubschman J-P, Davis JL, Heilwell G, Spirn M, Maguire J, Gay R, Bateman J, Ostrick RM, Morris D, Vincent M, Anglade E, Del Priore LV, Lanza R (2015) Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. Lancet 385:509–516. https://doi.org/10.1016/S0140-6736(14)61376-3

Seaberg RM, van der Kooy D (2003) Stem and progenitor cells: the premature desertion of rigorous definitions. Trends Neurosci 26:125–131. https://doi.org/10.1016/S0166-2236(03)00031-6

Sebastiano V, Zhen HH, Haddad B, Bashkirova E, Melo SP, Wang P, Leung TL, Siprashvili Z, Tichy A, Li J, Ameen M, Hawkins J, Lee S, Li L, Schwertschkow A, Bauer G, Lisowski L, Kay MA, Kim SK, Lane AT, Wernig M, Oro AE (2014) Human COL7A1-corrected induced pluripotent stem cells for the treatment of recessive dystrophic epidermolysis bullosa. Sci Transl Med 6(264):264ra163. https://doi.org/10.1126/scitranslmed.3009540

Segers VFM, Lee RT (2008) Stem-cell therapy for cardiac disease. Nature 451(7181):937–942. https://doi.org/10.1038/nature06800

Sheyn D, Ben-David S, Shapiro G, De Mel S, Bez M, Ornelas L, Sahabian A, Sareen D, Da X, Pelled G, Tawackoli W, Liu Z, Gazit D, Gazit Z (2016) Human induced pluripotent stem cells differentiate into functional mesenchymal stem cells and repair bone defects. Stem Cells Transl Med 5(11):1447–1460. https://doi.org/10.5966/sctm.2015-0311

Sipp D, Caulfield T, Kaye J, Barfoot J, Blackburn C, Chan S, De Luca M, Kent A, McCabe C, Munsie M, Sleeboom-Faulkner M, Sugarman J, van Zimmeren E, Zarzeczny A, Rasko JEJ (2017) Marketing of unproven stem cell-based interventions: a call to action. Sci Transl Med 9(397). https://doi.org/10.1126/scitranslmed.aag0426

Smets F, Dobbelaere D, McKiernan P, Dionisi-Vici C, Broue P, Jacquemin E, Lopes AI, Goncalves I, Mandel H, Pawlowska J, Kaminska D, Shteyer E, Torre G, Shapiro R, Eyskens F, Clapuyt P, Gissen P, Pariente D, Grunewald S, Yudkoff M, Binda MM, Najimi M, Belmonte N, Vos B, Thonnard J, Sokal E (2019) Phase I/II trial of liver derived mesenchymal stem cells in pediatric liver based metabolic disorders: a prospective, open label, multicenter, partially randomized, safety study of one cycle of heterologous human adult liver-derived progenitor cells (HepaStem(R)) in urea cycle disorders and Crigler-Najjar Syndrome patients. Transplantation. https://doi.org/10.1097/TP.0000000000002605

Sneddon JB, Tang Q, Stock P, Bluestone JA, Roy S, Desai T, Hebrok M (2018) Stem cell therapies for treating diabetes: progress and remaining challenges. Cell Stem Cell 22(6):810–823. https://doi.org/10.1016/j.stem.2018.05.016

Sokal EM (2014) Treating inborn errors of liver metabolism with stem cells: current clinical development. J Inherit Metab Dis 37(4):535–539. https://doi.org/10.1007/s10545-014-9691-x

Sokal EM, Stephenne X, Ottolenghi C, Jazouli N, Clapuyt P, Lacaille F, Najimi M, de Lonlay P, Smets F (2014) Liver engraftment and repopulation by in vitro expanded adult derived human liver stem cells in a child with ornithine carbamoyltransferase deficiency. JIMD Rep 13:65–72. https://doi.org/10.1007/8904_2013_257

Spitzhorn L-S, Rahman MS, Schwindt L, Ho H-T, Wruck W, Bohndorf M, Wehrmeyer S, Ncube A, Beyer I, Hagenbeck C, Balan P, Fehm T, Adjaye J (2017) Isolation and molecular characterization of amniotic fluid-derived mesenchymal stem cells obtained from caesarean sections. Stem Cells Int 2017:1–15. https://doi.org/10.1155/2017/5932706

Spitzhorn LS, Kordes C, Megges M, Sawitza I, Gotze S, Reichert D, Schulze-Matz P, Graffmann N, Bohndorf M, Wruck W, Kohler JP, Herebian D, Mayatepek E, Oreffo RO, Haussinger D, Adjaye J (2018) Transplanted human pluripotent stem cell-derived mesenchymal stem cells support liver regeneration in Gunn rats. Stem Cells Dev. https://doi.org/10.1089/scd.2018.0010

Spitzhorn LS, Megges M, Wruck W, Rahman MS, Otte J, Degistirici O, Meisel R, Sorg RV, Oreffo ROC, Adjaye J (2019) Human iPSC-derived MSCs (iMSCs) from

aged individuals acquire a rejuvenation signature. Stem Cell Res Ther 10(1):100. https://doi.org/10.1186/s13287-019-1209-x

Stanger BZ (2015) Cellular homeostasis and repair in the mammalian liver. Annu Rev Physiol 77:179–200. https://doi.org/10.1146/annurev-physiol-021113-170255

Steinhoff G, Nesteruk J, Wolfien M, Kundt G, Börgermann J, David R, Garbade J, Große J, Haverich A, Hennig H, Kaminski A, Lotz J, Mohr F-W, Müller P, Oostendorp R, Ruch U, Sarikouch S, Skorska A, Stamm C, Tiedemann G, Wagner FM, Wolkenhauer O (2017) Cardiac function improvement and bone marrow response –. EBioMedicine 22:208–224. https://doi.org/10.1016/j.ebiom.2017.07.022

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131(5):861–872. https://doi.org/10.1016/j.cell.2007.11.019

Tang Y, Yu P, Cheng L (2017) Current progress in the derivation and therapeutic application of neural stem cells. Nat Publ Group 8:3108. https://doi.org/10.1038/cddis.2017.504

Thomas ED, Lochte HL Jr, Lu WC, Ferrebee JW (1957) Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. N Engl J Med 257 (11):491–496. https://doi.org/10.1056/NEJM195709122571102

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM (1998) Embryonic stem cell lines derived from human blastocysts. Science 282(5391):1145–1147

Tolar J, McGrath JA, Xia L, Riddle MJ, Lees CJ, Eide C, Keene DR, Liu L, Osborn MJ, Lund TC, Blazar BR, Wagner JE (2014) Patient-specific naturally gene-reverted induced pluripotent stem cells in recessive dystrophic epidermolysis bullosa. J Invest Dermatol 134 (5):1246–1254. https://doi.org/10.1038/jid.2013.523

Toma C, Pittenger MF, Cahill KS, Byrne BJ, Kessler PD (2002) Human mesenchymal stem cells differentiate to a cardiomyocyte phenotype in the adult murine heart. Circulation 105(1):93–98. https://doi.org/10.1161/hc0102.101442

Tsuji O, Sugai K, Yamaguchi R, Tashiro S, Nagoshi N, Kohyama J, Iida T, Ohkubo T, Itakura G, Isoda M, Shinozaki M, Fujiyoshi K, Kanemura Y, Yamanaka S, Nakamura M, Okano H (2019) Concise review: laying the groundwork for a first-in-human study of an induced pluripotent stem cell-based intervention for spinal cord injury. Stem Cells 37:6–13. https://doi.org/10.1002/stem.2926

Tysnes O-B, Storstein A (2017) Epidemiology of Parkinson's disease. J Neural Transm 124:901–905. https://doi.org/10.1007/s00702-017-1686-y

van Rood JJ (1966) Leucocyte groups and histocompatibility. Vox Sang 11(3):276–292. https://doi.org/10.1111/j.1423-0410.1966.tb04603.x

Wenker SD, Pitossi FJ (2019) Cell therapy for Parkinson's disease is coming of age: current challenges and future prospects with a focus on immunomodulation. Gene Ther. https://doi.org/10.1038/s41434-019-0077-4

Wenzel D, Bayerl J, Nystrom A, Bruckner-Tuderman L, Meixner A, Penninger JM (2014) Genetically corrected iPSCs as cell therapy for recessive dystrophic epidermolysis bullosa. Sci Transl Med 6(264):264ra165. https://doi.org/10.1126/scitranslmed.3010083

Wu M, Zhang R, Zou Q, Chen Y, Zhou M, Li X, Ran R, Chen Q (2018) Comparison of the biological characteristics of mesenchymal stem cells derived from the human placenta and umbilical cord. Sci Rep 8(1):5014. https://doi.org/10.1038/s41598-018-23396-1

Yang YK (2018) Aging of mesenchymal stem cells: implication in regenerative medicine. Regen Ther 9:120–122. https://doi.org/10.1016/j.reth.2018.09.002

Zhao Y, Jiang Z, Zhao T, Ye M, Hu C, Zhou H, Yin Z, Chen Y, Zhang Y, Wang S, Shen J, Thaker H, Jain S, Li Y, Diao Y, Chen Y, Sun X, Fisk MB, Li H (2013) Targeting insulin resistance in type 2 diabetes via immune modulation of cord blood-derived multipotent stem cells (CB-SCs) in stem cell educator therapy: phase I/II clinical trial. BMC Med 11:160. https://doi.org/10.1186/1741-7015-11-160

Part II

Clinical Pharmacokinetics

Dose Finding in Single Dose Studies by Allometric Scaling

35

Zheng Lu, Rüdiger Kaspera, Yoichi Naritomi, and Tianli Wang

Contents

Abstract

This chapter reviews common methodologies and applications of scaling for clearance (CL), oral bioavailability (F), volume of distribution (V_d), and half-life ($t_{1/2}$) with respect to its procedures, evaluation, performance, and modifications for the dose finding in single dose studies. Methods of allometric scaling have been well established in drug development to predict the dose for single dose studies for first-in-human trials from preclinical data, generally from one to three or more species such as mouse, rat, dog, or monkey. Allometric scaling is the study of body size to diverse

Z. Lu (✉) · T. Wang
Clinical Pharmacology and Exploratory Development, Astellas Pharma, Northbrook, IL, USA
e-mail: zheng.lu@astellas.com; tlwang2011@gmail.com

R. Kaspera
Clinical Pharmacology and Exploratory Development, Astellas Pharma Europe B.V., Leiden, The Netherlands
e-mail: rudiger.kaspera@astellas.com

Y. Naritomi
Analysis and Pharmacokinetics Research Labs., Astellas Pharma Inc., Ibaraki, Japan
e-mail: yoichi.naritomi@astellas.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_3

biological characteristics, like clearance. The clearance and other PK parameters are proportional to the body weight among different species. Dose finding in single dose studies by allometric scaling can be either by dose-by-factor approach or pharmacokinetically-guided approach. Besides traditional simple allometry and allometric scaling based on the rule of exponents, many newly proposed methods based on the allometric scaling with different correction factors have been shown either to improve the accuracy of predications or to explore the possibilities of predictions for protein therapeutics. Although the allometric scaling approaches have been widely used to predict human PK parameters of small molecules which is critical to next dose finding step for single ascending dose study, there is a need to have more mechanistic methods based on the allometric scaling, especially for protein therapeutics or some formulations, like liposomal, to improve the predictabilities of first-in-human dose in single ascending dose study.

Purpose and Rationale

In first-in-human (FIH) studies, the single ascending dose (SAD) study, cohorts of six to eight subjects are given a single dose of the new chemical entity (NCE) and monitored for safety, tolerability, and pharmacokinetics (PK) to identify the maximum tolerated dose (MTD) and/or dose-limiting toxicity. Because it is the FIH study, the starting dose in a SAD study must be low enough to be safe but not too low as lead to an excessive numbers of dose levels to reach the MTD. Methods of allometric scaling have been well established in drug development to predict the dose for single dose studies for first-in-human trials from preclinical data. This chapter reviews common methodologies and applications of scaling for clearance (CL), volume of distribution (V_d), and half-life ($t_{1/2}$) with respect to its procedures, evaluation, performance, and modifications for the dose finding in single dose studies.

Procedure

Since drug efficacy and toxicity are usually associated with drug exposure, the projection of human pharmacokinetic (PK) for an early assessment of efficacious doses and dosing regimens is important during clinical development, especially prior to FIH studies (Deng et al. 2011). The allometric scaling approaches have been widely used to predict human PK parameters of small molecules, generally from three or more species such as mouse, rat, dog, or monkey.

Interspecies Scaling of CL

Systemic CL or so-called total CL does not identify the mechanism of the elimination process and considers drug elimination from the entire body. Allometric scaling is applicable to compounds mainly eliminated by cytochrome P450 (P450)-mediated metabolism and to compounds that undergo non-P450 metabolism or are eliminated mainly by urinary excretion as the unchanged drug, and protein binding is inconsequential (Boxenbaum 1984).

There have been a number of comprehensive reviews on the accuracy and precision of various methods to estimate human systemic CL. The most comprehensive report was by Lombardo et al. (2013a) who presented an analysis of 400 drugs (the largest known database in existence) administered intravenously to rats, dogs, monkey, and humans. They used 37 different methods to estimate human CL and calculated the median bias and geometric mean fold-relative error for each method, of which the most important ones are described here and their performance discussed in the following section.

Simple Allometry and Allometric Scaling Based on Rule of Exponents

Simple allometry scaling is based on a power function $CL = a \times Y^b$, generally using three or more species PK data, where Y may be the body weight (BW) or body surface area (SA) and a and b are the coefficient and exponent, respectively. For compounds with high hepatic extraction

ratios, body surface area is expected to be more predictive (Zou et al. 2012a).

The more commonly used equation is:

$$CL_{human} = CL_{animal} \times (BW_{human}/BW_{animal})^{0.75}$$

which is called single species scaling (SSS) (Hosea et al. 2009). This method uses a fixed scaling exponent with one species PK data. Caldwell et al. (2004) proposed that human PK predictions can be obtained using the simple allometric scaling from rats with a fixed exponent: CL(human) approximately = 40 × CL(rat) (L/hr).

Rule of exponents (ROE) can also be applied to human CL prediction as below:

$$\begin{aligned}
&\text{When } 0.55 \le b \le 0.70, \ CL = a \times BW^b \\
&\text{When } 0.71 \le b \le 1.0, CL_{human} = \\
&\quad a \times (MLP_{animal} \times CL_{animal})^b / MLP_{human} \\
&\text{When } b > 1.0, \ CL_{human} = a \times \\
&\quad (BrW_{animal} \times CL_{animal})^b / BrW_{human} \\
&\text{When } b > 1.3, \ CL \text{ may be overpredicted} \\
&\text{When } b < 0.55, \ CL \text{ may be underpredicted}
\end{aligned}$$

where MLP represents maximum lifespan potential and BrW indicates brain weight (Mahmood and Brian 1996).

Allometric Scaling of Hepatically Eliminated Drugs

The interspecies scaling of biliary and renal CL is not discussed within the chapter and the reader is referred to respective publications.

The Liver Blood Flow Method

For hepatically eliminated drugs, a common approach is to extrapolate human CL by the hepatic blood flow (LBF) ratio between humans and animals (Zou et al. 2012a).

$$CL_{human} = CL_{animal} \times (LBF_{human}/LBF_{animal})$$

It has been found that the mouse and monkey LBF methods were more accurate approach for human CL prediction than rat and dog LBF methods (Stoner et al. 2004; Ward and Smith 2004).

Normalization by In Vitro CL

The in vitro results from hepatic metabolism studies using microsomes, hepatocytes, and liver slices can be extrapolated in order to incorporate intrinsic CL. Combining the in vitro intrinsic with the in vivo CL of animals, Lave et al. (1997) predicted the human CL using allometric scaling. The in vivo CL of each species was normalized by multiplying the ratio of CL in human hepatocytes or microsomes versus CL in animal hepatocytes or microsomes. The method seemed to have not been widely used and may not show superiority over other allometric scalings.

Allometric Scaling of Unbound Drug CL

Plasma protein binding of many drugs varies considerably among animal species and only unbound drug can be eliminated. Therefore, protein binding has been considered to potentially influence the distribution and elimination of drugs (Zou et al. 2012a).

$$\text{Unbound CL} = CL/f_u$$

$$\text{Unbound CL} = a \times (BW)^b$$

where f_u is the unbound fraction in plasma.

Although the f_u in rats is observed to be representative of the average f_u in animals, correction for protein binding in each animal species would be more favorable than just considering only rats and humans (Tang and Mayersohn 2005). Tang and Mayersohn (2005) proposed the unbound fraction-corrected intercept method (FCIM) using the ratio of unbound fraction in plasma (f_u) between rats and human (Rf_u), based on 61 sets of CL values in animal species:

$$CL\,(mL/min) = 33.35 \times (a/Rf_u)^{0.77}$$

where a is a coefficient obtained from allometric scaling. It has been shown that, practically, unbound CL cannot be predicted more accurate than total CL, except for some drugs. Typically, protein binding corrections will not be made unless the ratio of f_u between rats and humans is tenfold or more.

Computational (In Silico) Approaches

Wajima et al. (2003) proposed an approach to predict human oral CL (mL/min/kg) by using experimental data for oral CL of the rat and dog, molecular weight (MW), clogP, and the number of hydrogen bond acceptors (Ha):

$$\begin{aligned}\log(CL) = &-0.5927 + 0.7386\log(CLrat) \\ &+ 0.5040\ \log(CLdog) + 0.06014\ c\ logP \\ &- 0.1862\ \log(CLdog) \times c\ logP \\ &+ 0.02893\ MW \times c\ logP + 0.02551\log(CLrat) \\ &\times \log(CLrat) \times c\ logP - 0.03029\ \log(CLrat) \\ &\times \log(CLrat) \times Ha - 0.03051\log(CLrat) \times MW \\ &\times c\ logP + 0.08461\log(CLdog) \\ &\times \log(CLdog) \times \log(CLdog) \\ &- 0.2510\ \log(CLdog) - 0.2510\ \log(CLdog) \\ &\times \log(CLdog) \times MW + 0.04607\log(CLdog) \\ &\times c\log P \times c\ logP - 0.003596\ c\ logP \\ &\times c\ logP \times Ha + 0.0005963\ c\log P \\ &\times Ha \times Ha\end{aligned}$$

Unlike other allometric scaling approaches, the *Wajima* method incorporates molecular structure parameters and is not dependent on the BW. The authors obtained clearance for 68 drugs either eliminated from renal excretion as unchanged drugs or extensively metabolized. The method gave a very good prediction of CL for the drugs studies.

Interspecies Scaling of Oral Bioavailability

The described above allometric scaling approaches predict human systemic CL. For orally administered drugs, it is vital to predict human oral bioavailability (F). F can be predicted based on preclinical in vivo estimates. One method is to use the average of all preclinical species. For example, if mouse, rat, and dog are 40%, 50%, and 60%, the human estimate is 50%. This method, however, is a rule of thumb and should be used carefully.

Since F can be expressed as:

$$F = F_a \times F_g \times F_h$$

where F_a denotes the fraction of the compound absorbed, F_g denotes intestinal availability, and F_h denotes hepatic availability, ideally for predicting F, it is necessary to predict all three parameters: F_a, F_g, and F_h. Assuming that total CL is equal to hepatic CL, it is possible to predict F_h using allometric scaling. On the other hand, to predict F_g and F_h, in vitro-in vivo extrapolation methods (IVIVE) using hepatic microsomes, hepatocytes, and intestinal microsomes have been actively investigated. These approaches are out of the scope of this chapter, and the reader is referred to respective publications.

Interspecies Scaling of V_d

V_d is commonly extrapolated pharmacokinetic parameter from animals. V_d of the central compartment (V_c) is most important in establishing the safety or toxicity for FIH studies by providing initial estimate of the plasma concentration following intravenous administration, and can be predicted with more accuracy than V_d at steady state (V_{ss}) or V_d by area (V_β). For majority of drugs, the exponents of the allometry for V_d revolve around 1.0. It has been suggested that if exponents of the allometry are > 1.1, then the V_d may be dramatically overestimated (Mahmood 2005).

Interspecies Scaling of $t_{1/2}$

Unlike CL and V_d, the $t_{1/2}$ is not directly related to the physiological body function, and thus, the correlation between BW and $t_{1/2}$ across species is poor (Mahmood 2005). Caldwell et al. (2004) proposed a simple allometric scaling for human $t_{1/2}$ approximately $= 4 \times t_{1/2}$ (rat) (hr), using 145 drugs. Another approach is to indirectly predict $t_{1/2}$ from CL and V_c. Alternatively, Mahmood (1998) suggested use of the allometry of mean residence time (MRT) versus BW to predict first MRT. $t_{1/2}$ can be then predicted by dividing the predicted MRT by 1.44. It is suggested that different approaches should be used to provide a range of predicted human $t_{1/2}$ before scientific judgment is used to select an appropriate estimate (Mahmood 2005).

Interspecies Scaling for Protein Therapeutics

It has been commonly believed that the nonhuman primate, usually the cynomolgus monkey, is the most relevant species for conducting preclinical PK studies for therapeutic monoclonal antibodies (mAbs). Human CL of mAbs can be reasonably projected based on monkey CL alone, by simple allometry with a fixed exponent of 0.85 for soluble antigen targets or 0.90 for membrane-bound targets. The dosage range for PK parameter determination was assumed to be linear (Deng et al. 2011).

Evaluation

Two tragic stories, of Tusko (West et al. 1962) which occurred 55 years ago, and of TGN1412 which occurred in 2006, are used as classical examples. Both fatal cases were due to overdoses in first trials.

Tusko, a 14-year-old Indian male elephant, died after intramuscularly dosing of 0.1 mg/kg which was a total dose of 297 mg lysergic acid diethylamide (LSD) on its body weight of 2970 kg. The trial tried to mimic a temporary form of madness in a zoo elephant. The dose was selected based on the observation that the rage in cats was produced with intravenous dose of 0.15 mg/kg LSD. Later, based on the calculations using the allometric approach, the actual dose of LSD to Tusko should have been much less and was in the range of 3–56 mg from different allometric approaches.

TGN1412 was intended to be used to treat leukemia and autoimmune disease such as rheumatoid arthritis. It is an agonistic monoclonal antibody which can bypass the requirement for T cell antigen receptor signaling and activates human T cells by only stimulating co-stimulatory receptor CD28 in the immune system. In the FIH clinical trial, it was dosed at 0.1 mg/kg to six healthy volunteers (Expert scientific group 2006). The dose selection was based on the no observed adverse effect level (NOAEL) which is considered as 50 mg/kg from the repeated dose toxicity study in cynomolgous monkeys (as described in the draft USA Food and Drug Administration (FDA) guideline "Estimating the Safe Starting Dose in Clinical Trials for Therapeutics in Adult Healthy Volunteers," 2002) with an additional safety factor of 160. However, extra precautions were not taken when antibodies are used to stimulate rather than neutralize components of the immune system. More than 90% of the CD28 receptors were bound by TGN1412 with proposed FIH dose of 0.1 mg/kg based on later calculations. Without any knowledge on the behavior of this compound in humans, the receptor occupancy of more than 90% was too high and induced massive production of cytokines and uncontrolled inflammatory responses which were observed in all six healthy volunteers in this trial. In conclusion, the preclinical development studies that were performed with TGN1412 did not predict a safe dose for use in humans, even though current regulatory requirement were met. Although the above two stories represented the failed evaluations of allometric scaling for dose finding in single dose studies, the importance of allometric scaling for the selection of "first time dose" in a species appears to be of immense significance. In addition, an understanding of the pharmacokinetic-pharmacodynamic relationship contributes to a much improved judgment.

The most widely used method for FIH dose estimation are "dose-by-factor" approach which is based on the NOAELs in multiple species and the "pharmacokinetically guided" approach. Both of these approaches rely on allometric scaling either of the dose itself or of drug clearance. For the NOAEL-based approach, the following case is used: The NOAEL in the 4-week rat toxicity study was 10 mg/kg/day and 3 mg/kg/day in the 4-week dog toxicity study, the human equivalent doses (HED) using body surface area conversion factor (BSA-CF), 0.16 for rats and 0.54 for dogs, were calculated as 1.6 mg/kg/day, then the maximum recommended starting dose (MRSD) in the FIH clinical trial is estimated as 9.7 mg/man by applying the default safety factor of 10 and based on a 60 kg body weight for a man. Basically, the dose by factor approach applies an exponent for body surface area (0.67), which account for differences in metabolic rate, to convert doses between

animals and humans. Thus, HED is determined by the below equation:

$$\mathrm{HED(mg/kg)} = \mathrm{Animal\ NOAEL(mg/kg)} \times (\mathrm{BW_{animal}(kg)/BW_{human}(kg)})^{0.33}$$

However, the dose by factor approach based on NOAEL does not take into account of systemic exposure (AUC) and the safety factor applied in the calculation of MRSD is very empirical. In the pharmacokinetically guided approach, systemic exposure instead of dose is extrapolated from animal to human, and difference in potency, free fraction in plasma and bioavailability between animals and humans should be also taken into account for the extrapolation. FIH dose from pharmacokinetically guided approach is calculated by the below equation:

$$D = AUC \times CL/F$$

in which AUC is extrapolated human AUC based on the animal AUC corresponding to NOAEL or lowest animal AUC if a NOAEL and its corresponding AUC are available from more than one animal species. Or by the equation:

$$D = C_{ss} \times \tau \times CL/F$$

in which C_{ss} is extrapolated human steady-state plasma concentration based on the animal C_{ss} and τ is the dosing interval. With extrapolated AUC or C_{ss}, the key elements left to project a dose in humans to produce a target AUC or C_{ss} are CL and absolute oral F based on above two equations. Besides predicted from in vitro data, the human oral F is predicted in some practice based on in vivo estimates using the average of all preclinical species. This method, however, is a rule of thumb and should be used with caution. The methods of allometric scaling of CL and F from animals to humans have been extensively discussed in the other sections of this chapter.

In another practice of interspecies allometric scaling to predict human PK parameters and FIH dose, oral plasma PK of ST-246 (Amantana et al. 2013) smallpox therapeutic was evaluated in mice, rabbits, monkeys, and dogs. Simple allometry relating animal oral plasma CL (CL/F) to animal BW was used to determine human CL/F. Using a 70 kg body weight, the human CL/F was predicted as 254 L/h from the approach of simple allometry (point estimate). Based on the ROE, the CL/F was predicted by using the MLP correction, since the scaling exponent was approximately 1.0. The point estimate of human CL/F was predicted as 51.4 L/h from the approach of MLP-corrected allometry. In order to establish good safety margin in a FIH study, a relatively lower CL/F was considered in this practice to determine a safe dose. With a pharmacokinetic-guided approach, the starting oral dose of 485 mg is the product of the lowest observed systemic exposure value (AUC) among the species utilized in this study which is 9.43 h $*$ μg/mL in dog and the scaled human CL/F which is 51.4 L/h based on the approach of MLP-corrected allometry. The trial was conducted from the low and median dose levels of 400 mg and 600mg to 800mg and the observed CL from these three levels of dosing are in close proximity to the predicted human CL/F from the approach of MLP-corrected allometry. Hence, this evaluation shows that allometric scaling of animal PK is useful in dose selection for FIH trials.

The similar approach was used to predict a FIH dose of 7-O-succinyl macrolactin A (SMA) (Keumhan et al. 2017), based on allometric scaling of PK data from mice, rats, and dogs. The human CL of SMA was first predicted by both simple allometric scaling and MLP-corrected allometric scaling of estimated CL from mice, rats, and dogs. The first-in-man dose of SMA was calculated by multiplying the efficacious exposure (AUC) with the predicted human CL from MLP-corrected allometry, which predicted a lower value of human CL of SMA.

Interspecies allometric scaling, including simple allometric scaling and allometric scaling with correction factors of MLP or BrW, has been increasingly applied in recent years to predict human PK properties of mAbs from preclinical data. However, PK allometric scaling across species with above allometric scaling fails in some cases of nonlinear PK and qualitative and quantitative difference in disposition pathways which

are typical for mAbs. Because the PK profiles of mAbs can be affected by their antigen through a target-mediated drug disposition, allometric scaling with correction factor of antigen concentration (AC) was evaluated for human CL estimation of four types of mAbs, including bevacizumab, etanercept, infliximab, and adalimumab (Wang et al. 2016b). In this evaluation, the plasma concentration of vascular endothelial growth factor (VEGF), which is the antigen of bevacizumab, was detected by enzyme-linked immunosorbent assay (ELISA) kits. The concentrations of tumor necrosis factor α which is the antigen of the rest three mAbs were obtained from published studies. The mean CLs of the mAbs in rabbit and dog were divided by the AC of the species and the product plotted as a function of BW on a log-log scale as in the below equation:

$$\mathrm{CL}/\mathrm{AC} = \mathrm{a} \times \mathrm{BW}^{\mathrm{b}}$$

The predicted human CL of 4.05 mL/day/kg of bevacizumab was close to the observed human CL of 5.73 mL/day/kg based on AC-corrected allometry and allometric scaling having the best prediction of human CL of etanercept and infliximab in comparison with other approaches including simple allometric scaling. Scaling with correction factors of MLP or BrW has equivalent good prediction of human CL of adalimumab with simple allometric scaling and scaling with a correction factor of BrW. These results indicated that AC has reasonably corrected the additional PK differences among the species besides the BW for mAbs. Although further evaluations AC-corrected allometry need to be conducted in the multiple species scaling of mAbs that showed nonlinear PK profiles, it may provide us a new perspective to estimate human PK parameters of mAbs from preclinical data to better find the dose of FIH trial of mAbs.

The allometric scaling for pegylated liposomal and nanoparticle anticancer drugs was first evaluated with the PK of CKD-602(S-CKD602), doxorubicin (Doxil®), and cisplatin (SPI-077) which were all available from mice, rats, dogs, and phase I clinical studies (Caron et al. 2011). Because proposed CL pathways for nanoparticles and liposomes are the monocytes and macrophages of the mononuclear phagocytel system (MPS), liver weight, spleen weight, monocyte count, spleen blood flow, and liver blood flow, and potential factors associated with the MPS were evaluated to determine if parameters other than BW can best allometrically scale the disposition of these anticancer agents. The variable with the strongest relationship to liposomal clearance across all agents was total monocyte count. A − 20.4%, 186%, and −78.2% difference existed between predicted and actual CL for S-CKD602, Doxil®, and SPI-077, respectively, when adding monocyte count into the allometric equation. This evaluation provided the preliminary evidence that factors associated with the MPS, such as monocyte count, may improve the prediction of CL in humans of drugs with liposomal formulations.

Critical Assessment of the Method

Human Systemic CL

The performance of prediction of human systemic clearance by allometric scaling was investigated for nearly 400 compounds (in addition with respect to charge class) by Lombardo et al. (2013a), and followed the assessment by the PhRMA initiative (Ring et al. 2011). In the first, the lowest mean-fold error, as well as frequency within twofold from predicted to observed CL, was observed for the following methods: (a) SSS using monkey, directly, or (b) including a correction of differences in liver blood flow, (c) FCIM, and (d) multiple linear regression (MLR*)* rat-dog. MLR is based on a logarithmic scaling of two species (log(CL human) = 0.4 × log (CL rat) + 0.4 × log(CL dog) – 0.4, Lombardo et al. 2013a).

The FCIM performed very well with a GMFE (geometric mean-fold error) of 1.9 with 62% of compounds <twofold. Across charge classes, SSS including monkey, and FCIM performed best with about 60% of compounds showing a GMFE <2. In general, the correction for the fraction of drug unbound in plasma tends to worsen the accuracy of all methods. The latter publication (Ring et al. 2011) also demonstrated that the SSS dog and MLR (rat-dog), as well as FCIM, perform

with a GMFE < 2 for about 2/3 of the compounds. The *Wajima* method performed well with 60% of the compounds <twofold and 87% < threefold, whereas IVIVE was the least well-performing method. In addition, the *Wajima* method was assessed more thoroughly for PK profile estimation (Lombardo et al. 2016), finding a well predictability for i.v., but the prediction of p.o. PK parameters including absorption constant and F requires more estimation and validation.

V_{ss}

The performance of the prediction of V_{ss} by allometric scaling was investigated by another publication from Lombardo et al. (2013b). The review recommends the use of multiple and best performing methods: SSS using dog (with a GMFE of <2, with 2/3 of the compounds below a twofold prediction), *Øie-Tozer* method (rat, dog, monkey, GMFE 1.5, and 79% under twofold), and *Wajima* method (if monkey data are not available, and GMFE 1.9 and 70% under twofold). Predictive performance decreased when protein binding correction was used and should only be used when human and animals several magnitudes apart. These methodologies are thought to be the most practicable and considered together on their convergence using a coefficient of variation. However, the approach may depend on compound chemistry. The *Øie-Tozer* method only performs best for classical small molecules (Stepensky 2011) and can generate aberrant values for $f_{u,T}$ (fraction unbound in tissue) for polar compounds of low volume, and be sensitive to the ratio of extravascular to intravascular binding (Waters and Lombardo 2010). In addition, the approach to average across the *best* methods improved the prediction of V_d (as well as CL). Using an average of the *Øie-Tozer*, the assumption that V_d (monkey equals V_d (human), and the *Wajima* and MLR improved the prediction (GMFE 1.5 and 88% below twofold) (Lombardo et al. 2016). This data set of 54 marketed drugs also yielded a good superposition of concentration time profiles using the *Wajima* superposition method, finally leaving only 18% of the compounds with a relatively poor prediction. Superiority of the *Øie-Tozer* method over allometry was also shown by Zou et al. (2012b).

$t_{1/2}$

Earlier investigations showed that CL, V_d, and elimination $t_{1/2}$, predicted by using pharmacokinetic constants, were comparable with values predicted by simple allometry (Mahmood 1999). As $t_{1/2}$ is a hybrid of CL and V_d, its prediction has been more thoroughly assessed within methods of human concentration-time profile predictions. Sinha et al. (2011) reported that 65% of 29 chosen compounds had a predicted <twofold error in the prediction of $t_{1/2}$, when C_{max} was estimated using simple allometry, ROE, correction using protein binding, and a direct approach using compartmental modeling. Vz (V_d during terminal phase)/F was estimated using simple allometry and allometry with protein binding corrections. The *Wajima* superposition method (Lombardo et al. 2016), which derived a $t_{1/2}$ from a concentration-time profile prediction using scaled V_d and CL, showed that 69% of 54 compounds were within a twofold error, while 83% were within a threefold error. The PHRMA C_{ss}-MRT method predicted $t_{1/2}$ (p.o.) with lower confidence, 42–49% of 106 to 247 compounds were predicted <twofold, and 60–72% < threefold (Vuppugalla et al. 2011). Additional methods such as physiologically based pharmacokinetic (PBPK) modeling and time invariant methods (Dedrick plot) are not within scope of this chapter and the reader is referred to the PHRMA publication and others (Poulin et al. 2011). However, PBPK-approaches seem to become more common in industry for PK profile predictions. For this purpose, the knowledge of essential physiochemical and preclinical PK parameter is essential (i.e., MW, logP, pKa, fraction unbound, blood-plasma ratio, and a clearance prediction) (e.g., Poulin et al. 2011).

Allometric Scaling of Protein Therapeutics

As indicated earlier, a comparative investigation was published by Deng et al. (2011). Thirteen monoclonal antibodies were tested for which a

better correlation was obtained between the observed human CL and the estimated human CL based on cynomolgus monkey PK data and an allometric scaling exponent of 0.85 for CL than other scaling approaches. Human concentration-time profiles were also reasonably predicted from the cynomolgus monkey data using species-invariant time method with a fixed exponent of 0.85 for CL and 1.0 for V_d. The predictive error was less than 50% for CL and below 100% for V_d. In a preceding investigation, six antibodies showed also a reasonable prediction of CL, V_d, but $t_{1/2}$ requires at least three animal species, when four methods were tested: (a) simple allometry, (b) maximum lifespan potential, (c) product of BrW and CL, and (d) fixed exponent (Mahmood 2009). For only one antibody, EGF/r3, the prediction was significantly higher. The prediction of the $t_{1/2}$ was reported with a %error of 43–107, and the use of fixed exponents can have large variability. Using only monkey data for 18 monoclonal antibodies (Ling et al. 2009), pharmacokinetic parameters may be best predicted with a time-invariant method and simple allometry for soluble or membrane-bound receptors. A suggested optimal exponent of 0.85–0.9 gave a percent error of up to 116% (maximum). For the prediction of PK-profiles in particular for nonlinear PK, species-invariant time methods (Dedrick plot), physiologically based PK models, or target-mediated drug-disposition models need to be considered, including PK/PD relationships or FcRn-binding (Wang et al. 2016a).

An additional example of four types of mAbs, including bevacizumab, etanercept, infliximab, and adalimumab, was discussed above (Wang et al. 2016b).

Overall, it is advised that a range of methodologies should be used and scientifically judged for dose selection and safety. Both studies show that clearance of selected biomacromolecules follows well-defined, size-related physiologic relationships, and preclinical PK studies provide reasonable estimates of human disposition under the premises of linear pharmacokinetics.

Coagulation factors have also been tested with a similar outcome. Using simple allometry with exponent to predict human pharmacokinetics for five factors, a three species scaling prediction was advantageous over a two species scaling with good correlation and an error of 11–34% for CL, 13–65% for V_{ss}, but large for $t_{1/2}$ (1–647%). A fixed exponent and single species scaling should be avoided (Mahmood 2009).

Modifications of the Method

Recently, chimeric mice with humanized liver have been developed and used for in vivo PK studies. The livers of the mice have been repopulated with human hepatocytes and express human drug-metabolizing enzymes. Therefore, chimeric mice with humanized liver are considered model animals for mimicking human drug metabolism and pharmacokinetics. Sanoh et al. (2015) reported that SSS using chimeric mice with humanized liver data showed excellent predictability of human CL and V_{ss} for various drugs metabolized by P450 and/or non-P450 enzymes. Further, they predicted human intravenous plasma-concentration curves using the complex Dedrick plot with scaling exponents for CL and V_{ss}, which enabled good prediction. Miyamoto et al. (2017) compared the predictability of SSS using data from chimeric mice with humanized liver, monkeys, and rats for 30 compounds metabolized by various drug-metabolizing enzymes. Data from chimeric mice with humanized liver produced the most accurate prediction of human CL among the species (approximately 80% within threefold range). Data from chimeric mice with humanized liver and monkeys produced comparably predictions of human V_{ss} with approximately 80% of compounds falling within threefold range. In contrast, data from rats produced low predictability. These results suggest that chimeric mice with humanized liver are useful for predicting human pharmacokinetics.

References and Further Reading

Amantana A, Chen Y, Tyavanagimatt SR, Jones KF, Jordan R, Chinsangaram J, Bolken TC, Leeds JM, Hruby DE (2013) Pharmacokinetics and interspecies allometric scaling of ST-246, an oral antiviral therapeutic for treatment of orthopoxvirus infection. PLoS One 8(4):e61514

Boxenbaum H (1984) Interspecies pharmacokinetic scaling and the evolutionary-comparative paradigm. Drug Metab Rev 15(5–6):1071–1121

Caldwell GW, Masucci JA, Yan Z, Hageman W (2004) Allometric scaling of pharmacokinetic parameters in drug discovery: can human CL, Vss and t1/2 be predicted from in-vivo rat data? Eur J Drug Metab Pharmacokinet 29(2):133–143

Caron WP, Clewell H, Dedrick R, Ramanathan RK, Davis WL, Yu N, Tonda M, Schellens JH, Beijnen JH, Zamboni WC (2011) Allometric scaling of pegylated liposomal anticancer drugs. J Pharmacokinet Pharmacodyn 38(5):653–669

Deng R, Iyer S, Theil FP, Mortensen DL, Fielder PJ, Prabhu S (2011) Projecting human pharmacokinetics of therapeutic antibodies from nonclinical data: what have we learned? MAbs 3(1):61–66

Expert Scientific Group (2006) Expert scientific group on phase one clinical trials final report. Stationery Office, Norwich

Estimating the safe starting dose in clinical trials for therapeutics in adult healthy volunteers (2002) U.-S. Department of Health and Human Services Food and Drug Administration, Centre for Drug Evaluation and Research (CDER), Centre for Biologics Evaluation and Research (CBER)

Hosea NA, Collard WT, Cole S, Maurer TS, Fang RX, Jones H, Kakar SM, Nakai Y, Smith BJ, Webster R, Beaumont K (2009) Prediction of human pharmacokinetics from preclinical information: comparative accuracy of quantitative prediction approaches. J Clin Pharmacol 49(5):513–533

Keumhan N, Wonku K (2017) Calculation of a first-in-man dose of 7-O-Succinyl Macrolactin A based on allometric scaling of data from mice, rats and dogs. Biomol Ther, pp 1–11

Lave T, Dupin S, Schmitt C, Chou RC, Jaeck D, Coassolo P (1997) Integration of in vitro data into allometric scaling to predict hepatic metabolic clearance in man: application to 10 extensively metabolized drugs. J Pharm Sci 86(5):584–590

Ling J, Zhou H, Jiao Q, Davis HM (2009) Interspecies scaling of therapeutic monoclonal antibodies: initial look. J Clin Pharmacol 49(12):1382–1402

Lombardo F, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, Harriman SP, Berellini G, Liric Rajlic I, Obach RS (2013a) Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 2: clearance. J Clin Pharmacol 53(2):178–191

Lombardo F, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, Harriman SP, Berellini G, Rajlic IL, Obach RS (2013b) Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 1: volume of distribution at steady state. J Clin Pharmacol 53(2):167–177

Lombardo F, Berellini G, Labonte LR, Liang G, Kim S (2016) Systematic evaluation of Wajima superposition (steady-state concentration to mean residence time) in the estimation of human intravenous pharmacokinetic profile. J Pharm Sci 105(3):1277–1287

Mahmood I (1998) Interspecies scaling: predicting volumes, mean residence time and elimination half-life. Some suggestions. J Pharm Pharmacol 50(5):493–499

Mahmood I (1999) Prediction of clearance, volume of distribution and half-life by allometric scaling and by use of plasma concentrations predicted from pharmacokinetic constants: a comparative study. J Pharm Pharmacol 51(8):905–910

Mahmood I (2005) Interspecies pharmacokinetic scaling: principles and application of allometric scaling. Pine House Publishers, Rockville

Mahmood I (2009) Pharmacokinetic allometric scaling of coagulation factors and tissue-type plasminogen activators. Haemophilia 15(5):1109–1117

Mahmood I, Balian JD (1996) Interspecies scaling: predicting clearance of drugs in humans. Three different approaches. Xenobiotica 26(9):887–895

Miyamoto M, Iwasaki S, Chisaki I, Nakagawa S, Amano N, Hirabayashi H (2017) Comparison of predictability for human pharmacokinetics parameters among monkeys, rats, and chimeric mice with humanised liver. Xenobiotica 47(12):1052–1063

Poulin P1, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, Ring BJ, Adkison KK, Ku MS, He H, Vuppugalla R, Marathe P, Fischer V, Dutta S, Sinha VK, Björnsson T, Lavé T, Yates JW (2011) PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci 100(10):4127–4157

Ring BJ, Chien JY, Adkison KK, Jones HM, Rowland M, Jones RD, Yates JW, MS K, Gibson CR, He H, Vuppugalla R, Marathe P, Fischer V, Dutta S, Sinha VK, Björnsson T, Lavé T, Poulin P (2011) PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 3: comparative assessement of prediction methods of human clearance. J Pharm Sci 100 (10):4090–4110

Sanoh S, Naritomi Y, Fujimoto M, Sato K, Kawamura A, Horiguchi A, Sugihara K, Kotake Y, Ohshita H, Tateno C, Horie T, Kitamura S, Ohta S (2015) Predictability of plasma concentration-time curves in humans using single-species allometric scaling of chimeric mice with humanized liver. Xenobiotica 45(7): 605–614

Sinha VK, Vaarties K, De Buck SS, Fenu LA, Nijsen M, Gilissen RA, Sanderson W, Van Uytsel K, Hoeben E, Van Peer A, Mackie CE, Smit JW (2011) Towards a better prediction of peak concentration, volume of distribution and half-life after oral drug administration in man, using allometry. Clin Pharmacokinet 50(5): 307–318

Stepensky D (2011) The Øie-Tozer model of drug distribution and its suitability for drugs with different pharmacokinetic behavior. Expert Opin Drug Metab Toxicol 7(10):1233–1243

Stoner CL, Cleton A, Johnson K, DM O, Hallak H, Brodfuehrer J, Surendran N, Han HK (2004) Integrated oral bioavailability projection using in vitro screening data as a selection tool in drug discovery. Int J Pharm 269(1):241–249

Tang H, Mayersohn M (2005) A novel model for prediction of human drug clearance by allometric scaling. Drug Metab Dispos 33(9):1297–1303

Vuppugalla R, Marathe P, He H, Jones RD, Yates JW, Jones HM, Gibson CR, Chien JY, Ring BJ, Adkison KK, MS K, Fischer V, Dutta S, Sinha VK, Björnsson T, Lavé T, Poulin P (2011) PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 4: prediction of plasma concentration-time profiles in human from in vivo preclinical data by using the Wajima approach. J Pharm Sci 100(10):4111–4126

Wajima T, Fukumura K, Yano Y, Oguma T (2003) Prediction of human pharmacokinetics from animal data and molecular structural parameters using multivariate regression analysis: oral clearance. J Pharm Sci 92(12):2427–2440

Wang J, Iyer S, Fielder PJ, Davis JD, Deng R (2016a) Projecting human pharmacokinetics of monoclonal antibodies from nonclinical data: comparative evaluation of prediction approaches in early drug development. Biopharm Drug Dispos 37(2):51–65

Wang L, Qiang W, Cheng Z (2016b) Allometric scaling of therapeutic monoclonal antibodies using antigen concentration as a correction factor: application to the human clearance prediction. J Pharm Sci 105(3): 1335–1340

Ward KW, Smith BR (2004) A comprehensive quantitative and qualitative evaluation of extrapolation of intravenous pharmacokinetic parameters from rat, dog, and monkey to humans. I. Clearance. Drug Metab Dispos 32(6):603–611

Waters NJ, Lombardo F (2010) Use of the Øie-Tozer model in understanding mechanisms and determinants of drug distribution. Drug Metab Dispos 38(7):1159–1165

West LJ, Pierce CM, Thomas WD (1962) Lysergic acid diethylamide: its effects on a male Asiatic elephant. Science 138(3545):1100–1103

Zou P, Yu Y, Zheng N, Yang Y, Paholak HJ, LX Y, Sun D (2012a) Applications of human pharmacokinetic prediction in first-in-human dose estimation. AAPS J 14(2):262–281

Zou P, Zheng N, Yang Y, LX Y, Sun D (2012b) Prediction of volume of distribution at steady state in humans: comparison of different approaches. Expert Opin Drug Metab Toxicol 8(7):855–872

Pharmacokinetic Aspects of Multiple Dose Studies

36

Steven G. Woolfrey and James Gilmour Morrison

Contents

Abstract

The development of a new chemical entity (NCE) in man is highly regulated and normally starts with a single ascending dose study (SAD) followed by a multiple ascending dose study (MAD), both in young healthy volunteers using the route and dose regime proposed in the final product. The aim of a MAD study is to establish the safety, maximum tolerated dose, and pharmacokinetics of the NCE on

S. G. Woolfrey (✉)
Independent Pharmaceutical Consultant – Clinical Pharmacology and Pharmacokinetics, Newcastle Upon Tyne, UK
e-mail: Steven.Woolfrey@outlook.com

J. G. Morrison
GW Pharmaceuticals, Cambridge, UK

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_4

repeated dosing, usually to steady state. This chapter describes the pharmacokinetic aspects of MAD studies and some of the design modifications that can be included in the study design. The MAD may be one of the few opportunities for assessing the pharmacokinetics of the compound in detail following multiple dosing, albeit not usually in patients. Previous data from the SAD should be used to optimize the design of the MAD including blood sampling regime, lower limit of sensitivity for the bioanalytical assay, and pharmacokinetic analysis. In addition to the pharmacokinetics and safety, the design should consider the need to characterize drug-related material in urine, carry out a preliminary assessment of the metabolic route, and characterize the potential for CYP3A4 induction. Suggestions are also provided for an approach when the impact of prandial status is unknown and more frequent dosing other than once daily is required.

Purpose and Rationale

The clinical development of a new chemical entity (NCE) is highly regulated (ICH–E8 1997), although the approach varies with the type of therapeutic class. Oncology agents are toxic at all dose levels and are normally only tested in oncology patients. Other therapeutic agents may be tested in normal volunteers, and it is these products that are the object of this chapter.

The clinical development of a non-oncology NCE normally begins with a single ascending dose (SAD) study in healthy young male volunteers, aged between 18 and 45 years. As virtually all drugs are administered repeatedly, it is important that the safety and pharmacokinetics are also investigated following multiple dosing. A SAD study is therefore normally followed by a multiple ascending dose (MAD) study, also in healthy young male volunteers. In this, the investigator aims to establish the safety, maximum tolerated dose at steady state, and pharmacokinetics of the NCE before proceeding into patients. This may be one of the few opportunities for assessing the pharmacokinetics of the compound in detail following multiple dosing, albeit not usually in patients. The route and, ideally, the dose regime should be those proposed for the final product at registration. The choice of subjects is usually young healthy male volunteers, but regulatory authorities now recognize the importance of studying females at an early stage of development (Liu and Mager 2016) and encourage their inclusion (ICH 2005).

The aim of this chapter is to describe the process for setting up and assessing the pharmacokinetic components of MAD Phase I studies in man. It is assumed that a SAD study will have already been carried out and data are available from that study. For the purpose of this chapter, it is generally assumed that young healthy male volunteers are used for the assessment, measuring plasma levels of the NCE given by the oral route; however, various design modifications are also considered.

A commonly used approach for the design of MAD studies in healthy volunteers is to administer multiple doses of the NCE to steady state, with detailed pharmacokinetic assessment on Day 1 and at steady state. In order to characterize the approach and time to steady state, pre-dose samples are also taken throughout the dosing period. Typically, this type of study is carried out in small groups of separate subjects per dose group, without consideration of the statistical power.

Any clinical study should always try to maximize data collection in order to justify the use of human volunteers, minimize development time, and justify the overall cost of the study. Consequently, it is usual practice to expand a MAD study to provide a preliminary investigation of metabolism and urinary pharmacokinetics where appropriate and contribute to an overall Phase I assessment of the importance of metabolic polymorphisms. Other subgroups are also possible depending on the properties of the NCE.

Procedure

The protocol and design of a typical MAD study in healthy male volunteers have a fairly standard layout, and the main protocol headings are

shown below. The number of dose groups in a MAD study will vary depending on the nature of the compound and the pharmacokinetics known at the time of the study. The use of three or four ascending dose groups is common practice for a compound with a reasonable safety profile, where predictable pharmacokinetic are expected.

A MAD study is normally carried out immediately after a SAD study. The primary purpose of a MAD study is to characterize the safety, as well as the pharmacokinetics after multiple dosing. The sections below only cover the pharmacokinetic aspects of the MAD study, reflecting the objectives of this work.

Protocol Outline

Objectives

As MAD studies involve an assessment of safety, as well as the pharmacokinetics of the NCE, the protocol has to reflect both objectives. It is normal practice to have safety as the primary objective and the pharmacokinetic characterization as a secondary objective. Investigation of other aspects of the NCE, e.g., preliminary gender effect, may also be stated here as tertiary objectives.

A typical protocol entry for the objectives of a MAD study is given below:

- Primary – To assess the safety and tolerability after repeated oral doses of ABC1234 to steady state
- Secondary – To assess the pharmacokinetic parameters after repeated oral doses of ABC1234 to steady state

Study Design

A variety of designs are possible for a MAD study, but each group needs to involve a small number of placebo subjects for comparison of the safety data. Some plasma samples may be used from the placebo subjects to confirm that the placebo subjects were not dosed with active compound, but otherwise they play no further part in the pharmacokinetic assessment. For example, plasma samples from the placebo subjects might be taken at the anticipated T_{max} and just before the next dose for bioanalysis on each main pharmacokinetic sampling day.

Although statistical analysis is carried out on the safety and pharmacokinetic data, the study is not powered to achieve statistical significance. The overall objective is to use the smallest number of subjects to characterize the multiple dose safety and pharmacokinetics of the NCE.

As the design of a MAD study involves both active and placebo-treated subjects, they are normally run as double-blind designs, involving the generation and use of a subject randomization sheet with the usual precautions and rules for accidental and required unblinding, respectively.

The design preferred by the author is a double-blind design of three or four treatment groups of 12 subjects, nine treated with active, and three with placebo. Subjects are randomly assigned to a group and each group treated repeated with an increasing dose of drug in the selected prandial state to steady state. The likely time to steady state is assessed from the results of the SAD study. The safety and pharmacokinetic data are assessed at each dose level before progressing to the next higher dose.

Inclusion Criteria

Overall inclusion criteria applied in a SAD study should be applied to the MAD study. As with virtually all Phase I studies, the subjects are normally men between 18 and 45 years in good health. Permitted concomitant medications allowed in the SAD study should normally be allowed in the MAD, but for both study types, the list should be very minimal. Any concomitant drugs that could affect the safety and pharmacokinetics of the NCE are normally excluded.

Treatments and Doses

The type of formulation at the time of the SAD and MAD studies is normally experimental hard gelatin capsules. It is recommended that the same formulation is used for the SAD and MAD studies to eliminate any differences due to formulation. If other formulations need to be examined, these should be included as separate subgroups.

The dosing process should be identical for the MAD and SAD studies and between MAD doses. Each daily treatment should be given as a combination of capsules with a standard volume of water (typically 200 ml) in the morning for a once-daily regime. A discussion on the use of fed and fasted prandial state is given in section "Critical Assessment of the Method."

Pharmacokinetic Data

In order to characterize the plasma pharmacokinetics of NCE, frequent blood samples for the measurement of total drug taken over a dosing interval on Day 1 and at the end of the study over the last dosing interval are required. About 12 to 15 samples are normally sufficient to characterize the pharmacokinetic of most NCE. Beyond the end of the last dosing interval, additional samples are taken in order to characterize the terminal half-life at steady state, typically another two to five samples depending on the expected half-life. Blood samples are taken pre-dose (C_{trough}) on selected days during the dosing phase to measure the approach and time to steady state. In all cases, the authors recommend duplicate pharmacokinetic plasma samples, one transported to the site of bioanalysis and the other retained at the clinical site, in case of problems in transport or bioanalysis.

A pre-dose blood sample is normally collected to investigate allelic variants of drug metabolism enzymes and drug transporters as part of the development process as described in section "Genotyping Data."

If the NCE is cleared in urine, then urine should be collected for pharmacokinetic analysis as described in section "Need to Measure Drug in Urine."

Bioanalytical work needs to be carried out to GLP and GCP standards using fully validated assays. Stability studies are needed to support the storage regime for plasma (normally frozen at −20 °C or −70 °C) and collection process from blood.

Current development regulation requires that the pharmacokinetic analysis is carried out using a model-independent approach (Rowland and Tozer 2011). A description of the software, pharmacokinetic parameters, statistical techniques, and quality assurance procedures needs to be included in the protocol. Statistical analysis may be described in detail in the protocol or, more usually, in a separate statistical analysis plan.

The level of detail provided in the protocol for the pharmacokinetic analysis should specify the parameters, their formula, and procedures for assessment, for example, half-life estimates and extrapolations. This is particularly important if the pharmacokinetic analysis is carried out by an independent organization.

Evaluation

The pharmacokinetic analysis of a NCE is carried out using actual blood sampling times on individual plasma concentration-time profiles. Data can only be excluded if there is a fully documented reason indicating a problem with the bioanalysis, dosing, or subject compliance, which must be fully explained and documented in the final study report.

The following pharmacokinetic parameters form the backbone of all MAD studies determined for each individual plasma concentration-time profile:

Day 1 – C_{max}, t_{max}, $AUC_{last,}$ and AUC_{τ}
Day 14 – C_{max}, t_{max}, AUC_{τ} and $t_{1/2}$
C_{trough} plasma concentrations on Days 2–14

Other parameters may be needed as outlined in section "Example."

Assessment of dose proportionality is an important objective of the MAD study. One approach is to assess dose proportionality for C_{max} and AUC_{τ} on Day 1 and at steady state separately, using the empirical power model (Parameter $= \alpha \times \text{dose}^{\beta}$), along with an "estimation" interpretation, according to the recommendations of Gough et al. (1995). Within-subject and total standard deviations for $\log(C_{max})$ and $\log(AUC_{\tau})$ are estimated.

Accumulation of the NCE is also assessed in the MAD study using the Day 1 and steady-state data for C_{max} and AUC_{τ} (R_{Cmax} and R_{ac}).

Accumulation is examined statistically for log (ratio of day steady state/Day 1) C_{max} and AUC $_{\tau}$, with a linear fixed effects model. Accumulation ratios are assessed for each dose level separately, as well as pooled across dose levels within the fixed effects model framework.

The occurrence of steady state is typically assessed by fitting C_{trough} values to a nonlinear mixed effects model in order to predict the time to achieve 90% of the steady-state trough concentration, taking into account any dose differences.

Half-life determined at steady state is compared across the doses using a linear mixed effect model.

Summary statistics are calculated for all parameters and the data then listed on an individual basis and with the summary statistics. The data may be plotted to show both the individual and mean values so that trends can be easily identified and investigated. Reporting is carried out according to ICH guidelines (ICH-E3 1995).

Critical Assessment of the Method

Use of Previous Data to Design Study

In order to optimize the sampling regime and choose the doses for the MAD study, it is recommended that data from a clinical SAD study are used to predict the pharmacokinetics on multiple dosing, including accumulation, time to steady state, and exposure at steady state. Such modeling and simulation will help provide confidence that the likely steady-state exposures will be within safe limits, but are adequate to allow a pharmacological effect in Phase II. It will also provide confidence that the bioanalytical assay has sufficient sensitivity to characterize the pharmacokinetics on all sampling days.

Preliminary Pharmacokinetic Analysis

It is common practice to carry out an interim pharmacokinetic analysis for each dose of the MAD study to check exposure levels and the pharmacokinetics prior to progression to the next dose. As the bioanalysis can take a couple of days to turn around, many investigators use a "N-1" approach to assess each dose: Typically this preliminary analysis uses nominal blood sampling times and bioanalytical data that has not been subjected to a full quality control check, on the basis that it will be repeated in the final analysis using actual sampling time data and final analytical data. The preliminary pharmacokinetics of the previous dose is examined at the same time as the safety data from the current dose in order to make a decision on whether it is feasible to proceed to the next dose in the dose escalation. In order to carry out the interim analysis in a timely manner, close coordination is required between sample transfer to the site of bioanalysis, the bioanalysis itself, and the pharmacokinetic analysis.

Approach to a Double-Blinded Study

Multiple dose Phase I studies are normally carried out in a double-blinded manner in order to aid the interpretation of the safety data. The bioanalyst and pharmacokineticist, however, effectively "unblind" the study in their interim analyses. It should be remembered that often the bioanalysis cannot wait until the completion and "lock" of the safety data, particularly if preliminary pharmacokinetic data have to be released prior to the next dose increase. Pragmatic procedures, therefore, need to be written into the protocol to protect the safety assessment from accidental "unblinding," which, for example, can be relevant if adverse events are recorded for individual subjects. Pragmatic approaches to interim pharmacokinetic data include removing subject identifiers for individual data and only presenting summary data, e.g., mean and SD.

Dosing in Fed or Fasted State

The effect of the prandial state on the pharmacokinetics of the NCE may be unknown at the point of carrying out a MAD study. If it has been examined prior to the MAD study, then the prandial state that significantly maximizes the exposure

(typically greater than twofold) is recommended, as this will probably be recommended for the final product. If data are unavailable or previous data have shown no major influence of food, then dosing in the MAD study should be carried out after an overnight fast, with a light breakfast allowed 2 h after dosing. The same prandial state for dosing should be used throughout the MAD.

Blood Sampling

The blood sampling regime is fundamental to the successful outcome of any pharmacokinetic assessment. Great care must, therefore, be taken to ensure that an adequate number of samples, appropriately placed during the dosing regime, are built into the design that yield levels above the lower limit of detection of the assay. The data from the SAD study and predictions of the pharmacokinetics on multiple doing should be used to help decide on the blood sampling regime and the target assay sensitivity.

Practical considerations also need to be considered with respect to the actual timing of blood samples: Sampling during the night when the subjects are normally asleep is possible but should obviously be kept to a minimum and avoided if at all possible.

Genotyping Data

The pharmacokinetics of some drugs can be affected by genetic polymorphisms of drug metabolism enzymes, which may lead to an apparent high level of pharmacokinetic variability and, in extreme cases, safety concerns. By the time an NCE enters clinical development, some knowledge is available on the likely impact of the major genetic polymorphisms from in vitro data. As a result, subjects should be screened for the major or compound-specific metabolic genetic polymorphisms, in order to assess the effect of these on the pharmacokinetics of the NCE at the end of the Phase I program. At this point data will be available from a limited number of subjects showing a number of allelic variants of metabolism enzymes.

Modification of the Method

Additional Pharmacokinetic Parameters After Oral Dosing

There are many pharmacokinetic parameters that can be calculated to characterize the pharmacokinetics of a compound after multiple oral dosing, and the most common used for summarizing the data are those given in section "Evaluation."

Three optional parameters that are often calculated, include the apparent volume of distribution, clearance, and mean residence time, all estimated at steady state in MAD studies. It must be remembered that the apparent volume of distribution and clearance is influenced by the bioavailability of the compound, which may well be unknown early in clinical development.

There are a variety of apparent volumes of distribution terms that may be calculated. It is generally regarded, however, that the volume of distribution at steady state (V_{ss}/F) is the most robust (ICH-E3 1995), although the apparent volume term for the terminal phase (V_z/F) is the one usually calculated.

Other parameters that are sometimes stated include C_{min}, $C_{average}$ (at steady state), t_{last}, and $t_{1/2\ effective}$, the latter being the half-life estimated from the time to achieve steady state.

In some cases, there may be a lag in the absorption of a compound, possibly due to the dispersion of an oral dose form. Such a delay is usually observed in the SAD study, and it may be important to include t_{lag} in the MAD study. If a lag time is not observed after a single dose, it is probably unnecessary to measure this parameter in the MAD study, unless there is a significant change in dosage formulation.

Impact of Dosing More Frequently than Once Daily

In some cases, a NCE is administered more frequently than once daily, in order to achieve adequate exposure during the dosing interval. The principles that have been developed for multiple dosing assuming once-a-day dosing also apply to

the situation when dosing is more frequent. In essence, the pharmacokinetic analysis is carried per dosing interval.

With once-daily dosing, the impact of diurnal variation on exposure is usually not apparent. With more frequent dosing (e.g., twice daily), however, diurnal variation may have a significant effect on exposure and, ultimately, the therapeutic effect. This can be assessed by characterizing the pharmacokinetics at steady state under two dose periods, these being carefully chosen to maximize the likely differences (e.g., one during the day and one at night). When setting up such comparisons, the number of blood samples that can reasonably be taken may become problematical either because of their number and volume of blood or because of the need to sample intensively during the night. Care needs to be taken to make sure that the exposure data will be comparable between the two dosing periods and other Phase I studies, e.g., SAD.

When developing a drug, the final therapeutic dose is highly influenced by the exposure ratio between the proposed human dose and the no observed adverse effect level (NOAEL) or no effect level (NOEL) in the most sensitive animal toxicology species. The animal data are often, but not always, generated by once-daily dosing and so a strategy needs to be established beforehand on how the animal/human exposure comparisons will be determined.

Need to Measure Drug in Urine

The need to sample urine for the assessment of the NCE will depend whether the compound is eliminated via the kidneys to any great extent. This is often assessed in the single ascending dose study; however, it may at this very early stage be based on preliminary non-validated assays. It is recommended that if >5% of the dose is renally excreted as parent drug, then the quantification of drug in urine via a fully validated assay needs to be carefully considered, as part of the MAD study.

There are a variety of ways to characterize the urinary pharmacokinetics of a NCE. It is recommended, however, that a urine collection is made over a dosing interval at steady state from which the following parameters can be calculated using standard pharmacokinetic principles (Rowland and Tozer 2011):

- Amount excreted under a dosing interval ($Ae_{0\text{-}T}$)
- Fraction of the dose excreted per dosing interval ($Fe_{0\text{-}T}$)
- Urinary Clearance ($Cl_{R0\text{-}T}$)

Care needs to be taken to ensure a complete collection of urine (as the volume of urine/dosing period is critical to the assessment) and that the urinary collection and storage process maintain the integrity of the compound. It is usual to collect the urine over various collection periods within a dosing interval from which aliquots for bioanalysis can be frozen and sent for analysis. As for plasma, duplicate samples for bioanalysis are recommended, one being maintained at the site as a backup.

Characterization of Metabolites

The development of any NCE requires the characterization of its route of metabolism and excretion in man (ICH E8). In the past, this was done in a small group of human subjects using a radiolabeled NCE, usually with ^{14}C, the study often carried out early in the Phase I development program. With the advancement of mass spectroscopy, it is now possible for many NCE to carry out a significant part of the metabolic characterization using unlabeled NCE, as demonstrated by Xiao et al. (2016). The MAD study provides a better opportunity than the SAD for this work in that the drug will be at steady state.

A potential option in MAD studies is to take large plasma and urine samples at various times over a dosing interval at steady state and examine these for the presence of drug-related material via mass spectroscopy. Although this work may not completely eliminate the eventual need for a human radiolabeled metabolic study, it should allow such a study to be pushed back to late in the development process, when the success of the NCE is more certain.

It is also worth considering that current regulatory guidelines (FDA 2016; ICH-M3 (R2)) may require the quantification of major drug-related metabolites in both animal toxicology and human clinical studies. A knowledge of the human metabolism at a relatively early stage in the clinical development may therefore be critical to avoid having to repeat studies, purely because of a lack of exposure data.

Assessment of CYP3A4 Induction

Characterizing the potential of a NCE to induce CYP3A4 is an important part of the development process of any drug. In vitro studies can give a preliminary indication of the likelihood of induction potential but are not definitive. A commonly used approach has been to look at the change in the urinary ratio of 6β-hydroxy cortisol/cortisol before dosing and at steady state in the MAD study, although the results can be variable (Galteau and Shamsa 2003).

More recently, however, the use of plasma 4β-hydroxy cholesterol has proved a more robust technique (Diczfalusy et al. 2009) that has largely replaced the use of 6β-hydroxy cortisol/cortisol urinary ratio. For plasma 4β-hydroxy cholesterol, plasma samples are taken before dosing and at selected days during the study at pre-dose and a ratio of Day X/Day 0 then used to assess CYP3A4 status (Dutreix et al. 2013, 2014). For example, samples for 4β-hydroxy cholesterol could be taken at baseline, Day 3, Day 6, and Day 9s to 14 for a 2-week MAD study.

Example

To illustrate the type of data that can be obtained using a typical MAD, the results of an actual study are described below for NCE, ABC1234.

In this example, the pharmacokinetics of ABC1234 had already been studied in a SAD study in healthy male volunteers. Based on predictions of steady state, the available safety data and likely pharmacological levels, it had been decided to carry out a MAD study at four dose levels up to a maximum daily dose of 60 mg. Predictions indicated that steady state would be achieved before Day 14. The assay for the measurement of drug was adequately sensitive with a limit of quantification (LOQ) of 0.2 ng/mL. The study was carried out in a double-blinded manner in young healthy male volunteers, with the objective of characterizing the safety and pharmacokinetics of the compound after once-daily dosing for 14 days.

The compound was administered in hard gelatin capsules once daily for 14 days to groups of 12 subjects, 3 of which were matched placebos. A blood sampling regime, based on that from the SAD study, was used on Days 1 and 14. Pre-dose samples (C_{trough}) were taken from Days 2 to 14. After the last dose on Day 14, blood sampling was continued up to 72 h post last dose in order to characterize the terminal half-life at steady state. Previous human studies had indicated urinary excretion was $<5\%$ of the dose; thus, urine was not collected for the assessment of urinary pharmacokinetics. A pre-dose blood sample was collected from all subjects in order to genotype for metabolic polymorphisms, although the results were not considered in the current analysis.

The pharmacokinetic profiles of ABC1234 on Days 1 and 14 are shown in Figs. 1 and 2, with the approach to steady state given in Fig. 3. The derived pharmacokinetic parameters are shown in Tables 1 and 2, which included V_z/F, Cl_{ss}/F, and MRT. No evidence of a lag time in absorption (t_{lag}) had been observed in the SAD study, and the dosage form was identical; hence, it was not measured here. All blood samples were collected within $\pm 15\%$ of the theoretical sampling time specified in the protocol.

ABC1234 appeared rapidly in plasma following oral administration of 10, 20, 40, and 60 mg of ABC1234 to healthy young male subjects. On Day 1, following a single oral administration, plasma concentrations were quantifiable up to the last sampling time (24 h post-dose) at all dose levels. On Day 14 following repeated daily administration, plasma concentrations were quantifiable up to the last sampling time (72 h post-dose) at the 20, 40, and 60 mg dose levels. ABC1234 was not quantifiable in any of the pre-

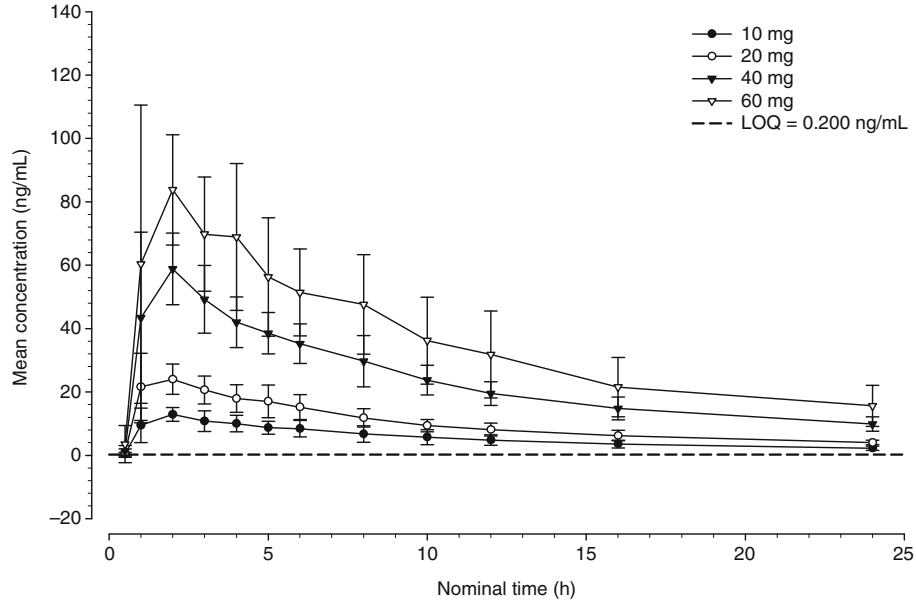


Fig. 1 Plasma concentrations (Mean ± SD) of ABC1234 after a single oral administration on Day 1

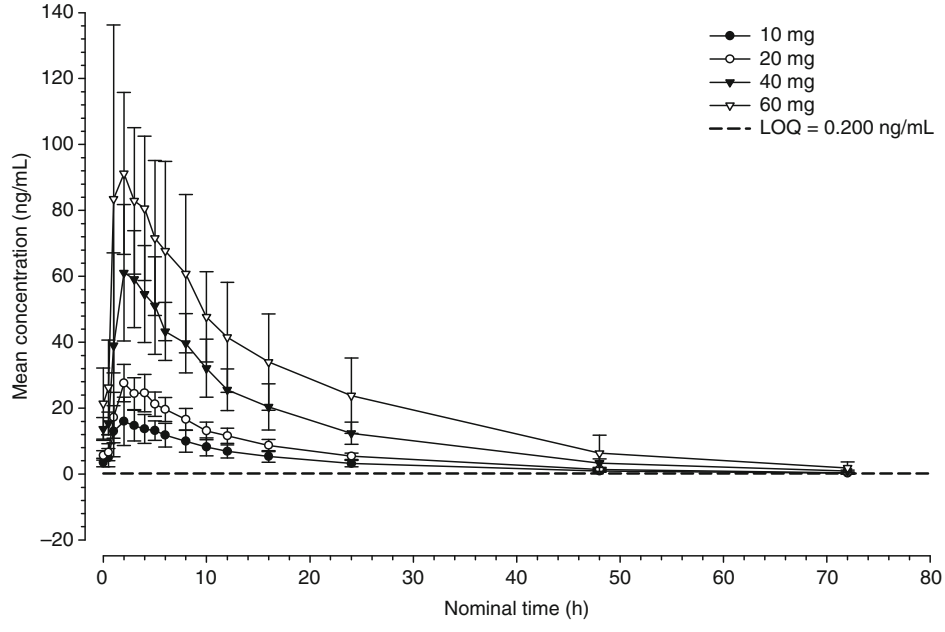


Fig. 2 Plasma concentrations (Mean ± SD) of ABC1234 at steady state on Day 14 after repeated daily oral administration for 14 days

dose Day 1 samples for subjects who received ABC1234 or in any of the samples analyzed from subjects who received placebo.

Mean C_{max} ranged from 13.7 ng/mL to 101 ng/mL on Day 1 and 18.1 ng/mL to 106 ng/mL on Day 14, over the dose range studied. The

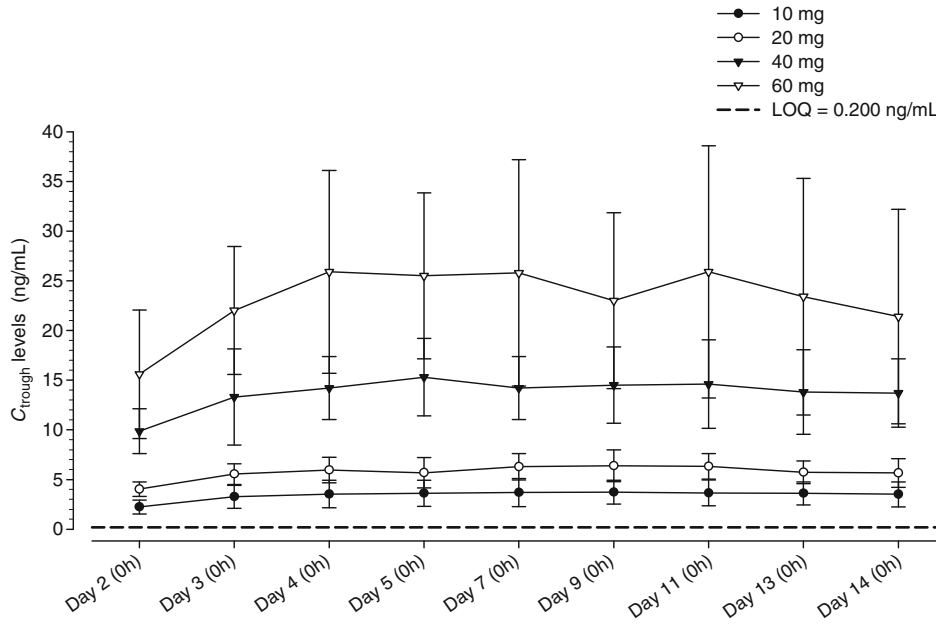


Fig. 3 C_{trough} plasma concentrations (Mean ± SD) of ABC1234 following daily administration for 14 days

Table 1 Pharmacokinetic parameters of ABC1234 after a single oral administration

PK parameter	10 mg	20 mg	40 mg	60 mg
C_{max} (ng/mL)	13.7 ± 2.76 (20) [13.5]	28.6 ± 5.09 (18) [28.1]	61.6 ± 13.6 (22) [60.2]	101 ± 26.3 (26) [97.7]
t_{max} (h)	2.00 (1.00, 3.00)	2.00 (1.00, 2.00)	2.00 (1.00, 2.00)	2.00 (1.00, 4.00)
AUC_{last} (ng h/mL)	129 ± 38.0 (29) [125]	234 ± 47.3 (20) [230]	558 ± 99.7 (18) [551]	840 ± 229 (27) [816]
$AUC_{0–24}$ (ng h/mL)	131 ± 38.3 (29) [126]	237 ± 47.7 (20) [232]	563 ± 100 (18) [556]	849 ± 233 (27) [824]

Tabulated values are Mean ± SD (CV%) [Geometric Mean] except for t_{max} where values are Median (Min, Max): Minor differences between AUC_{last} and $AUC_{0–24}$ reflect small variation in the actual sampling times around the 24 h point, but all were within ±15% of nominal

increase in C_{max} showed no major deviation from dose proportionality, although it should be noted that the lower 90% confidence interval for the β estimate of the log transformed power model was slightly greater than 1 on Day 1 (90% CI: 1.01,1.19). For a sixfold increase in dose, C_{max} increased by 7.20-fold on Day 1 (90% CI: 6.14, 8.45) and 6.23-fold on Day 14 (90%CI: 4.99, 7.79).

There was no accumulation of C_{max} over the dosing period (accumulation ratio pooled across doses was 1.08; 90% CI: 0.998, 1.16). Within-subject and total subject variability was low (19% and 26%, respectively).

Mean $AUC_{0–24}$ increased in a manner that was dose-proportional on both Days 1 and 14: For a sixfold increase in dose, $AUC_{0–24}$ increased by 6.80-fold on Day 1 (90% CI: 5.77, 8.02) and 6.19-fold on Day 14 (90% CI: 5.09, 7.52).

Consistent with the dose-proportional increase in C_{max} and $AUC_{0–24}$, V_z/F, Cl_{ss}/F, and MRT were independent of dose.

Over the dosing period, there was minimal accumulation of $AUC_{0–24}$ (accumulation ratio pooled across doses was 1.33; 90% CI: 1.29, 1.38). Within-subject and total subject variability was low, with values of 7.5% and 24%, respectively.

Table 2 Pharmacokinetic parameters of ABC1234 at steady state on Day 14 after daily oral administration for 14 days

PK parameter	10 mg	20 mg	40 mg	60 mg
C_{max} (ng/mL)	18.1 ± 7.41 (41) [16.9]	27.8 ± 5.45 (20) [27.3]	66.3 ± 18.8 (28) [64.1]	106 ± 35.6 (34) [102]
t_{max} (h)	2.00 (1.02, 5.02)	2.00 (2.00, 3.00)	3.00 (1.00, 4.00)	2.00 (1.00, 6.00)
$t_{1/2}$ (h)	14.1 ± 2.33 (17) [13.9]	12.7 ± 1.09 (9) [12.6]	12.6 ± 0.978 (8) [12.6]	12.6 ± 2.13 (17) [12.4]
$AUC_{0–24}$ (ng h/mL)	190 ± 58.9 (31) [183]	313 ± 60.5 (19) [308]	723 ± 177 (25) [705]	1150 ± 402 (35) [1100]
V_z/F (L)	824 ± 211 (26) [802]	882 ± 151 (17) [871]	783 ± 186 (24) [763]	726 ± 213 (29) [698]
Cl_{ss}/F (L/h)	41.5 ± 11.7 (28) [39.9]	48.6 ± 9.28 (19) [47.8]	43.2 ± 10.3 (24) [42.1]	41.2 ± 13.6 (33) [39.0]
MRT (h)	17.4 ± 2.10 (12) [17.3]	16.7 ± 1.05 (6) [16.6]	16.4 ± 1.40 (9) [16.4]	17.3 ± 2.78 (16) [17.1]

Tabulated values are Mean ± SD (CV%) [Geometric Mean] except for t_{max} where values are Median (Min, Max)

On repeated once-daily dosing, steady state was reached after the second or third dose at all dose levels (Fig. 3). Median time to steady state pooled across doses was 2.1 days (90th percentile 2.4 days).

Arithmetic mean $t_{1/2z}$ of ABC1234 on Day 14 following repeated daily administration was 12.6–14.1 h over the dose range studied. The difference in mean $t_{1/2z}$ between doses was not statistically significant ($p = 0.2333$), consistent with the dose-proportional nature of the pharmacokinetics.

In conclusion, ABC1234 appeared rapidly in plasma with a median t_{max} of 2–3 h and then declined with a terminal half-life of approximately 13 h. C_{max} and $AUC_{0–24}$ increased in a dose-proportional manner, reaching steady state by about Day 2. After repeated daily administration, there was minimal accumulation of $AUC_{0–24}$ (accumulation ratio: 1.33) over the dosing period. Within-subject and total subject variability for $AUC_{0–24}$ was low (7.5% and 24%, respectively).

References and Further Reading

Diczfalusy U, Kanebratt KP, Bredberg E, Andersson TB, Böttiger Y, Bertilsson L (2009) 4β-hydroxycholesterol as an endogenous marker for CYP3A4/5 activity. Stability and half-life of elimination after induction with rifampicin. Br J Clin Pharmacol 67(1):38–43

Dutreix C, Munarini F, Lorenzo S, Roesel J, Wang Y (2013) Investigation of CYP3A4-mediated drug-drug interactions on midostaurin in healthy volunteers. Cancer Chemother Pharmacol 72:1223–1234

Dutreix C, Lorenzo S, Wang Y (2014) Comparison of two endogenous biomarkers of CYP3A4 activity in a drug–drug interaction study between midostaurin and rifampicin. Eur J Clin Pharmacol 70(8):915–920

FDA (2016) Guidance for industry: safety testing of drug metabolites, revision 1

Galteau MM, Sharmsa F (2003) Urinary 6β-hydroxycortisol: a validated test for evaluating drug induction or drug inhibition mediated through CYP3A in humans and in animals. Eur J Clin Pharmacol 59(10):713–733

Gough K, Hutchinson M, Keene O, Byrom B, Ellis S, Lacey L, McKellar J (1995) Assessment of dose proportionality: report from the statisticians in the pharmaceutical industry/pharmacokinetics UK joint working party. Drug Inf J 29:1039–1048

ICH (2005) Gender considerations in the conduct of clinical trials

ICH M3 (R2) (2008) Nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals

ICH-E3 (1995) Structure and content of clinical study reports

ICH-E8 (1997) General considerations for clinical trials

Lui KA, Mager NA (2016) Women's involvement in clinical trials: historical perspective and future implications. Pharm Pract (Granada) 14(1):708. 1–9

Rowland M, Tozer TN (2011) Clinical pharmacokinetics and pharmacodynamics: concept and applications, 4th edn. Lippincott Williams & Wilkins, Philadelphia

Xiao W, Shen G, Zhuang X, Ran X, Zhu M, Li H (2016) Characterization of human metabolism and disposition of levo-tetrahydropalmatine: qualitative and quantitative determination of oxidative and conjugated metabolites. J Pharm Biomed Anal 128:371–381

Dose Linearity and Proportionality

37

Tanja Eisenblaetter, Lenore Teichert, Ronald Burnette, and Paul Hutson

Contents

T. Eisenblaetter (✉)
München, Germany
e-mail: t.eisenblaetter@web.de

L. Teichert
Biostatistics, Sanofi-Aventis Deutschland GmbH,
Frankfurt am Main, Germany
e-mail: Lenore.Teichert@sanofi.com

R. Burnette · P. Hutson
University of Wisconsin School of Pharmacy, Madison,
WI, USA
e-mail: ronald.burnette@wisc.edu; paul.hutson@wisc.edu

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_5

A desirable characteristic of a drug is linear pharmacokinetic properties to facilitate dose and dose regimen adjustment in patients. "Linear pharmacokinetics" implies that any concentration–time profiles normalized for dose and time are superimposable (Ludden 1991). Thus, one of the necessary conditions for linear pharmacokinetics is dose proportionality, and its assessment is a fundamental pharmacokinetic analysis conducted during the clinical development of a new drug candidate.

Dose Proportionality

If the concentration of a drug (usually in plasma) at any given time is proportional to the dose of the drug administered, then that drug is said to be dose proportional (Smith 2004). If the dose of such a drug is doubled (or tripled or halved), so is the concentration. Mathematically, dose proportionality at a given time point implies that for any dose equal or above 0

$$C \ \alpha \ \text{dose} \tag{1}$$

or replacing the proportionality with an equality

$$C = \alpha * \text{dose} \tag{2}$$

where C is the concentration at a given time point after dosing and α is some regression constant. A relationship between dose and C in case of dose proportionality is illustrated in Fig. 1.

When the concentration is normalized for dose, Eq. 2 passes into Eq. 3 for any dose above 0, illustrating that dose-normalized concentrations being constant are conditions equivalent to dose proportionality of these concentrations.

$$C/\text{dose} = \alpha \,(\text{for dose} > 0) \tag{3}$$

Area under the curve (AUC) and maximum concentration (C_{max}) are generally used to test for dose-normality instead of comparing raw plasma drug concentrations. However, there is no reason why other dose-dependent concentration measures (e.g., trough concentrations) cannot be used. A dose-proportional compound should exhibit dose proportionality for any dose-dependent concentration measure (e.g., minimum concentration, steady-state concentration, amount excreted via kidneys in a given time period).

Dose Linearity

Dose linearity is not to be confused with linear pharmacokinetics. Dose linearity is a weaker condition, even weaker than dose proportionality. It can be described by simple linear regression of the exposure measure C against dose

$$C = \alpha_0 + \alpha * \text{dose} \tag{4}$$

where α_0 is an intercept term and α is a regression constant. If the intercept term α_0 is 0, then Eq. 4 simplifies to Eq. 2 which is the definition of dose proportionality. Dose proportionality implies no

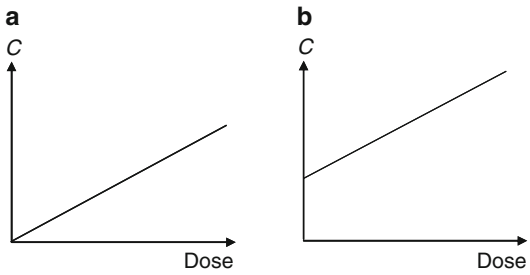


Fig. 1 Relationship between dose and concentration (C) in case of dose proportionality (**a**) and general dose linearity (**b**), both with $\alpha > 0$

drug exposure if the dose is 0. If the intercept term α_0 takes a nonzero value, the result is dose linearity, but without dose proportionality (Cawello et al. 1999). Although no dose is administered, the exposure can be larger than 0, which is typically observed for endogenous compounds. This is why a formal assessment of dose linearity and determination of the value of α is of minor practical importance of most drug candidates. The relationship between dose and C in case of dose proportionality and general dose linearity is illustrated in Fig. 1.

Mechanisms Leading to Lack of Dose Proportionality

Lack of dose proportionality (implying nonlinear pharmacokinetics) may be due to many mechanisms but is typically due to the saturation of some component in the system, such as metabolizing enzymes or transporters. Ludden (1991) classified nonlinearity into causes due to nonlinear absorption, nonlinear distribution, or nonlinear elimination.

Common causes of dose non-proportionality due to nonlinear absorption include saturation of carrier-mediated uptake, poor aqueous solubility or slow release from the formulation, and saturation of presystemic metabolism. Common mechanisms of nonlinear distribution include saturable protein binding, red blood cell binding, and tissue distribution. Lack of dose proportionality due to saturable elimination includes saturable elimination at metabolic enzymes, saturable renal elimination at transporters, and auto-induction.

Lack of dose proportionality may have implications with regard to safety and efficacy. For a drug that shows dose-dependent absorption, typically higher doses lead to less absorption and sub-proportional drug concentrations. In this case, efficacy becomes a concern. For a drug that shows saturable elimination, higher doses lead to higher than proportional concentrations and increased risk of adverse events. This becomes more of a concern when a drug has a narrow therapeutic window. Related to this is the issue of insufficient predictability of the concentration and response or toxicity from a change in dose.

Clinical Assessment of Dose Linearity/Proportionality

The assessment of dose linearity/proportionality typically starts with early exploratory single-dose clinical studies (see section "Dose Proportionality") providing PK data over a considerable dose range (Frick et al. 2006). These data support exposure–response relationships and subsequent dose selection in patients for the potential submission for new-drug approval (US FDA 2014). Dose linearity/proportionality may be assessed during drug development by more complex study designs (see section "Dose Linearity") to further support PK/PD relationships, ending with comprehensive confirmatory studies (see section "Mechanisms Leading to Lack of Dose Proportionality") to support drug labeling, dosage form modifications (EMEA CPMP 2014), or the use of several dosage strengths (EMEA CPMP 2010).

Statistical Assessment of Dose Linearity/Proportionality

Dose proportionality is a mathematically ideal concept, which physiologically will never be met in a strict sense. For instance, there is no biological setting where Eq. 2 can hold for an unlimited range of doses. In addition, even if the true expected concentrations for a drug suggested

ideal dose proportionality within a certain dose range, due to biological variability within and between subjects this could never be fully proven.

This leads to the assessment of dose proportionality being a statistical question, that is, to testing and estimation problems. Statistical analyses are intended to determine to what extent the data are compatible with the model of dose proportionality, to quantify deviations from the ideal, and to support the derivation of clinical implications.

The clinical question to be answered is whether the deviations of expected exposure from dose proportionality are of clinical relevance or not. Deviations are relevant if they are large enough to expose the patient to likely risk from concentration-dependent effects within the range of approved doses.

Descriptive Analyses

Descriptive analyses normally include the presentations of descriptive statistics for concentration-related parameters (typical statistics: number of non-missing observations, mean, standard deviation, minimum, median, maximum, geometric mean, coefficient of variation). This can be supplemented by the presentation of the corresponding descriptive statistics for dose-normalized parameters.

Graphical display could include scatterplots of PK parameters over dose for the raw or for the dose-normalized parameters. For the latter, dose proportionality is represented by a dose-independent (horizontal) level of the PK parameters (e.g., CL, V) plotted against dose. Parameters such as AUC that are expected to be dose-dependent are usually normalized to dose before plotting against the dose (e.g., Fig. 8). In addition, this presentation accounts for the fact that dose-normalization typically standardizes the variability.

Discrete Model

For any study design with a set of fixed doses, one approach to test for deviations from dose proportionality can be based on classical linear models with a fixed effect for dose for log-transformed dose-normalized parameters. If a significant dose effect is found, strict dose proportionality can be considered refuted. However, this does not necessarily imply that the deviations are of any clinical relevance.

Pairwise comparison of exposure allows estimation of the ratio of PK parameters AUC or C_{max} for two given doses. Given dose proportionality, the ratio is expected to be equal to the ratio of doses. Alternatively, ratios for dose-normalized parameters could be assessed and compared with the value of 1 which is to be expected under dose proportionality (e.g., Table 6).

Power Model

Quantifying deviations from dose proportionality for any dose demands the use of statistical models where parameters can be estimated together with measures of imprecision (usually confidence intervals) and where dose proportionality is characterized by certain values of the model parameters. Assuming log-normal distributions for exposure-related parameters like AUC and C_{max} suggests model deviations from dose proportionality that are multiplicative rather than additive. A statistically useful way to model this is known as the power model (Smith et al. 2000)

$$C = \alpha * \text{dose}^{\beta} \tag{5}$$

Ideal dose proportionality is met when $\beta = 1$: that is, when $C = \alpha$ * dose (Eq. 2).

Deviations from dose proportionality ($C = \alpha$ * dose) are modeled as factors depending on β, which can be seen in

$$C = \alpha^{*}\text{dose}^{\beta} = \alpha^{*}\text{dose}^{*}\text{dose}^{\beta-1} \tag{6}$$

For an r-fold change of a given dose an $r^{\beta-1}$-fold change of the exposure can be expected if the drug demonstrates dose proportionality.

$$C/\text{dose} = \alpha^{*}\text{dose}^{\beta-1} \tag{7}$$

Values of β above 1 represent higher than proportional drug exposure (super-proportional), whereas values of β below 1 represent lower

exposure (sub-proportional). Restated, the factor $r^{\beta-1}$ can be considered as the dose-normalized ratio between exposure following the r-fold change of a dose versus the initial dose. Different scenarios are given as examples in Table 1 below.

For instance, for a drug with a value of 1.25 for β, in case of doubling the dose (r = 2) the exposure would be more than doubled: it would be 19% higher than expected under dose proportionality. Figure 2 shows the deviation from dose-proportional exposure for different values of β.

It will be of clinical importance to assess whether the range for "r" is large enough to cover clinically relevant dose ranges.

One important beneficial statistical feature of the power model is that log transformation leads to a very simple linear model with a normal distribution for log(C):

$$\log(C) = \log(\alpha) + \beta^{*}\log(\text{dose}) + \text{Error} \quad (\text{for dose} > 0) \qquad (8)$$

Estimates of β together with its confidence limits can be derived using standard statistical methods for linear models (Gough et al. 1995; Smith et al. 2000). Based on that, it follows that for a given dose ratio a confidence interval for the "deviation factor" $r^{\beta-1}$ can be calculated. This leads to the analysis being linked to questions of bioequivalence. If, for instance, a classical range of 0.8–1.25 can be considered for a drug as an acceptable range for bioequivalence, the deviation from dose proportionality may be considered as irrelevant as long as the confidence interval for $r^{\beta-1}$ does not violate this interval of 0.8–1.25. A 90% confidence interval is commonly used (US FDA 2001).

Table 1 Factor for deviation from dose proportional exposure for different values of β and r

	β				
Dose factor r	0.5	0.75	1	1.25	1.5
0.1	3.16	1.78	1.00	0.56	0.32
0.5	1.41	1.19	1.00	0.84	0.71
1.0	1.00	1.00	1.00	1.00	1.00
2.0	0.71	0.84	1.00	1.19	1.41
10	0.32	0.56	1.00	1.78	3.16

Example 1: Exploratory Assessment of Dose Linearity/Proportionality – Single Dose Study Design

Purpose and Rationale

The primary objective of first-in-human studies is generally to assess the safety and tolerability of a drug in healthy volunteers. Pharmacokinetic and

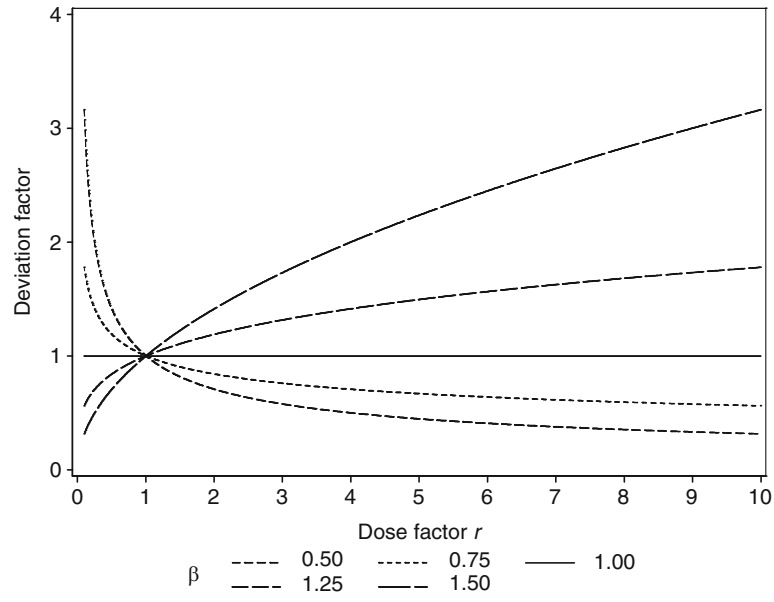


Fig. 2 Factor for deviation from dose proportional exposure for different values of beta (β)

pharmacodynamic data are often collected in these dose-escalation studies, so a secondary objective is thus the evaluation of pharmacokinetic parameters to describe the dose effect on drug absorption (t_{max}, t_{lag}), elimination ($t_{1/2z}$), and exposure (C_{max}, AUC) to assess the drug linearity/proportionality. The selection of the starting dose in these initial human trials is generally based on preclinical data from the most sensitive species in toxicology studies. The tested dose range should cover the potential therapeutic dose and should allow the determination of a maximum tolerated dose in humans and the safety margin.

The design of an exploratory assessment of dose linearity/proportionality during conduct of a first-in-human study for candidate drug "X001" is presented below. The description is limited to pharmacokinetic data although safety/tolerability and pharmacodynamic data were also obtained.

Primary objectives were to assess the clinical and laboratory safety/tolerability of X001, and secondary objectives were to assess the pharmacokinetics (PK) and pharmacodynamics (PD) following ascending single oral doses of X001 under fasted conditions. It was a single center, double-blind, placebo-controlled, randomized, escalating single oral dose study.

Study Design

Healthy young male subjects aged between 18 and 45 years with a body weight between 50 and 90 kg and a body mass index between 18 and 28 kg/m2 were recruited. These subjects were randomized to receive X001 oral odoses of 1, 2, 5, 10, 20, 40, 80, 160, 300 or 500 mg, or placebo. The placebo-controlled study design was chosen for purposes of safety and PD assessment, and was not mandatory for the PK objective. Six subjects in each dose step were randomized to X001 and two subjects to a matching placebo. The capsules were administered with 240 mL of non-carbonated water after an overnight fast and with a 4 h post-administration fasting period.

Eighty subjects were recruited, plus eight additional subjects for optional dose levels.

Starting with the lowest dose, each of the subsequent doses was administered only if the preceding dose was safe and well tolerated. The decision to proceed to the next higher dose (n + 1) was based on the full range of safety parameters of the last dose (n) and pharmacokinetic data of the previous dose (n − 1). Subjects entered the study unit the evening before the study of drug administration and were assessed for their baseline characteristics on the morning of the day of drug administration. Subjects remained in the study unit for 48 h after the dose was administered. Standard safety/tolerability criteria (e.g., adverse events, biochemistry, hematology, urinalysis, ECG, and vital signs) and additional pharmacodynamic parameters (e.g., postprandial blood glucose) were assessed in this study but are out of the scope of this chapter.

Blood samples to determine X001 were collected before dosing, and at 0.25 h, 0.5 h, 1 h, 1.5 h, 2 h, 3 h, 4 h, 6 h, 8 h, 12 h, 24 h, 36 h, and 48 h post-dose. X001 concentrations in plasma were assayed using liquid chromatography–tandem mass spectrometry (LC–MS/MS) with a validated lower limit of quantification (LLOQ) for X001 of 1 ng/mL.

The following pharmacokinetic parameters were calculated using standard non-compartmental techniques: maximum concentration (C_{max}), time to maximum concentration (t_{max}), area under the concentration–time curve from time of drug administration to last quantifiable concentration time point (AUC_{last}), area under the concentration–time curve from time of drug administration extrapolated to infinity (AUC), terminal elimination half-life ($t_{1/2z}$), and time to the first quantifiable concentration (t_{lag}). Plasma concentrations and pharmacokinetic parameters were listed by standard descriptive statistics (N, mean, SD, SE, min, median, max, CV%, and geometric mean).

Dose Proportionality

Dose proportionality of C_{max} and AUC was evaluated using the log-transformed power model with dose as the fixed effect:

$$\text{Log}\,(\text{PK parameter}) = \text{Log}\,(\alpha) + \beta^{*}\text{Log}\,(\text{dose}) + \text{Error} \quad (9)$$

This model was tested for goodness-of-fit using the plot of residuals. Since there was no evidence of goodness-of-fit (residuals randomly distributed around the origin), estimates for β with 90% confidence intervals (CI) were obtained by ordinary least squares regression. Estimates and 90% CI for PK parameter associated with an r-fold (r = 2 and r = highest dose/lowest dose) increase in dose were obtained by exponentiating r to the powers of the β estimate ($\widehat{\beta}$) and confidence limits:

$$r^{\widehat{\beta} \pm t_{0.95},_{df} SE\left(\widehat{\beta}\right)} \quad (10)$$

Dose Effect

The effect of dose upon $t_{1/2z}$ was assessed using a linear fixed effects model on log-transformed values:

$$\text{Log}\left(t_{1/2z}\right) = \text{dose} + \text{Error} \quad (11)$$

Point estimate and 90% CI for the geometric means of $t_{1/2z}$ were pooled across dose levels and separately for each dose level.

Results

Mean X001 plasma concentration time profiles following a single oral dose of 1 mg, 2 mg, 5 mg, 10 mg, 20 mg, 40 mg, 80 mg, 160 mg, 300 mg, and 500 mg X001 are presented in Fig. 3. X001 was rapidly absorbed, showing peak plasma concentrations approximately 1 h post-dose, irrespective of dose.

A summary of the descriptive statistics of main X001 PK parameters is given in Table 2. The relationship of individual AUC values versus dose, with linear regression and the 95% confidence range is illustrated in Fig. 4. The results of dose proportionality analysis are summarized in Table 3. X001 exposure increased with increasing doses. 90% CI of β-estimates for C_{max} and AUC were 0.97–1.03 and 1.00–1.06, respectively, thus including the unity and demonstrating dose proportionality. The mean terminal elimination half-life ($t_{1/2z}$) increased from 1 h to 2 h with increases in dose. Statistical analysis revealed that the dose had a significant effect ($p < 0.001$) on the terminal elimination half-life ($t_{1/2z}$).

Typically, this kind of early first-in-human study covers a very broad dose range, often leading to dose disproportionality over the entire dose range. At very low doses, PK parameters such as elimination half-life are often not reliable due to

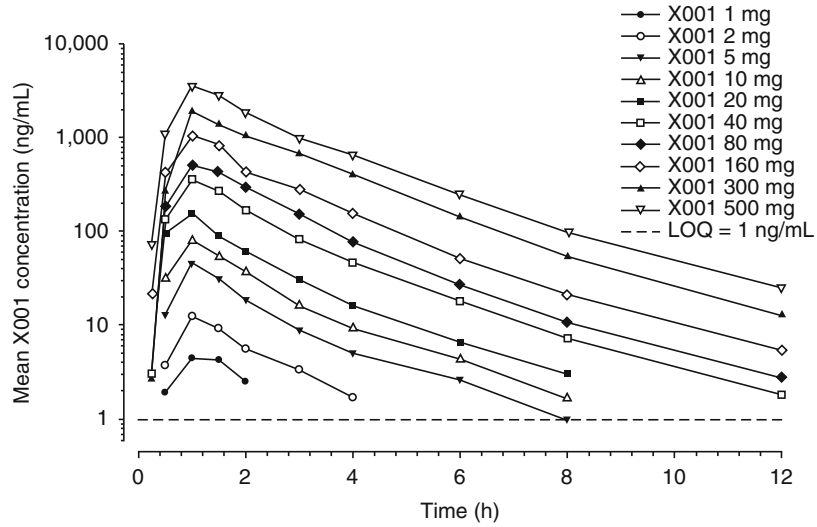


Fig. 3 Mean X001 plasma concentrations (semi-logarithmic scale)

Table 2 Key PK parameters and descriptive statistics of X001 by dose

	AUC (ng*h/mL)				C_{max} (ng/mL)				$t_{1/2z}$ (h)			
Dose (mg)	Mean	CV%	SD	Geometric mean	Mean	CV%	SD	Geometric mean	Mean	CV%	SD	Geometric mean
1	10.5	20	2.1	10.4	6.99	36	2.5	6.61	0.968	63	0.609	0.852
2	25.7	37	9.4	24.4	13.1	5	0.7	13.1	1.36	27	0.36	1.31
5	84.4	35	29.1	80.5	50.5	42	21.0	47.3	1.93	18	0.36	1.90
10	160	18	29	158	93.0	28	26.4	90.2	1.62	18	0.29	1.60
20	283	33	94	269	164	40	66	152	1.58	4	0.07	1.58
40	714	33	237	683	390	30	117	375	1.79	19	0.33	1.77
80	1,170	26	301	1,140	554	21	117	543	1.76	14	0.25	1.74
160	2,150	7	155	2,140	1,210	28	334	1,170	1.78	12	0.22	1.77
300	4,350	32	1,370	4,150	2,140	16	346	2,110	2.07	31	0.65	2.00
500	7,910	18	1,440	7,800	3,880	14	540	3,850	2.04	28	0.57	1.98

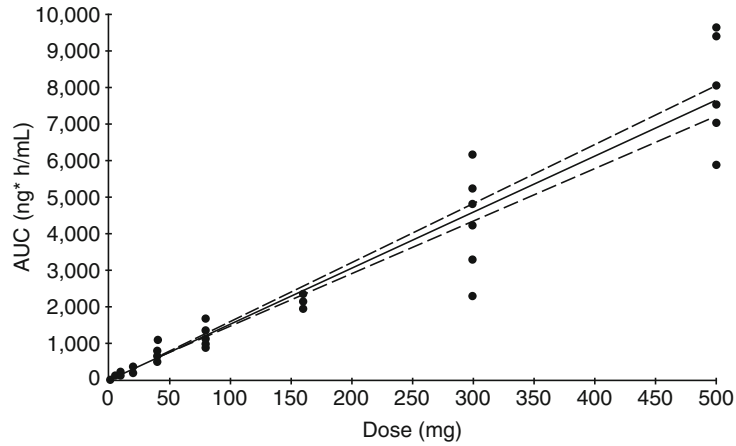


Fig. 4 Relationship of individual AUC values for X001 versus dose with linear regression (bold line) and the 95% confidence range (dashed line)

Table 3 Estimates with 90% CI for r-fold increase in dose

Parameter	Dose ratio	Ratio	
		Estimate	90% CI
C_{max} (ng/mL)	r = 2	2.00	1.96–2.05
	r = 500	506	414–619
	β-Estimate	1.00	0.97–1.03
AUC (ng∗h/mL)	r = 2	2.04	1.99–2.08
	r = 500	592	487–720
	β-Estimate	1.03	1.00–1.06

r = 500 = highest/lowest dose

the limits of analytical quantification. At high doses, saturation of absorption or elimination processes or limitations in drug release from the dosage form due to insolubility may decrease the exposure and prevent a dose proportional increase. In these cases a pivotal investigation of dose linearity/proportionality becomes necessary. It typically includes at least three different doses covering the anticipated therapeutic dose range proposed for regulatory approval. More details are provided in section "Mechanisms Leading to Lack of Dose Proportionality."

Discussion

As shown in Fig. 3 for the two lower doses of 1 mg and 2 mg, X001 could only be detected up to 2 h and 4 h, respectively, resulting in very limited data points in the elimination phase of the drug. Therefore, the quantification limit of the analytical method might have compromised the reliability of the PK parameters obtained for these doses (1 mg and 2 mg). For example, the low value of approximately 1 h for the calculated $t_{1/2z}$ for the 1 mg and 2 mg doses (Table 2) differs from the value of approximately 2 h for the doses from 5 mg to 500 mg. If the low doses are of further interest (e.g., as pharmacologically active dose), the lower limit of quantification of the analytical method could be improved in order to guarantee a reasonable PK assessment.

Evaluation of dose-proportionality is part of the first study in humans which is always a single-dose study but could also be applied to early multiple-dose studies. This kind of explorative data related to dose linearity/proportionality can be used to predict the exposure for further planned dose steps, including doses outside the dose range investigated so far. If a notable nonlinear effect is seen by this exploratory evaluation, a more elaborate study will need to be performed as described in section "Mechanisms Leading to Lack of Dose Proportionality." In some cases dose linearity/proportionality assessment using the described explorative method is not possible due to high interindividual variability in PK characteristics. In these cases an intra-individual crossover design is preferred as described in section "Dose Linearity."

Example 2: Exploratory Assessment of Dose Linearity/Proportionality – Crossover Study Design

Purpose and Rationale

Drug candidate "X002" had already been carefully explored before this study for pharmacodynamics, pharmacokinetics, and safety in a large number of subjects. However, no formal evaluation of the dose–exposure–response relationship had yet been conducted.

The primary objective of this study was to investigate the dose–exposure–response relationship of X002 after single subcutaneous injections of 0.075 units/kg body weight (U/kg), 0.15 U/kg, and 0.3 U/kg. The secondary objective was to assess the safety and tolerance of X002. The description is limited to pharmacokinetic data, although safety, tolerability, and pharmacodynamic data were also obtained.

The number of doses was restricted to three to allow for intra-subject comparisons in a crossover design.

Study Design

This was a single-center, single-blind, randomized, single dose, three-way crossover study comparing three single doses of X002 (0.075 U/kg, 0.15 U/kg, and 0.3 U/kg) injected subcutaneously. The design was a full crossover including all six possible treatment sequences.

All subjects were male patients between 18 and 55 years, a body mass index between 18 kg/m^2 and 30 kg/m^2, and within the context of the underlying disease had normal findings in the following assessments: medical history, physical examination, laboratory values, electrocardiogram (ECG), blood pressure, pulse rate, and core body temperature, unless the investigator considered any abnormality to be clinically irrelevant and not interfering with the safety of the subject and the scientific integrity of the study.

Subjects were randomized to one of the six sequences of three dose levels of X002. The study consisted of five trial periods – trial period 0 (screening visit), trial periods 1, 2, 3 (X002 treatment visits), and trial period 4 (follow-up visit). X002 was injected subcutaneously into the predefined body region on three different study days.

Serum concentrations were measured at time point 0 (prior to dosing) and after 10 min, 20 min, 30 min, 40 min, 50 min, 60 min, 70 min, 80 min, and 90 min as well as after 2 h, 2.5 h, 3 h, 4 h, 5 h, 6 h, 8 h, and 10 h after the injection of study medication. If the end of the period for pharmacodynamic assessment occurred earlier than the maximum of 10 h post-study medication injection, no further samples were taken. Serum concentrations of X002 were analyzed using a radioimmunoassay. The lower limit of quantification (LLOQ) was 5.0 μU/mL.

Area under the concentration–time curve for the time between 0 h and 2 h after dosing ($AUC_{(0–2h)}$, μU∗min/mL) was considered the primary PK parameter and was derived using standard non-compartmental methods as were $AUC_{(0–end)}$ (μU∗min/mL), and mean residence time (MRT, h).

Maximum concentration (C_{max}, μU/mL) and time to C_{max} (t_{max}, h) were derived using a compartmental method. No adjustments of the alpha levels were made for multiple analyses.

Serum concentrations and pharmacokinetic parameters were listed individually and summarized per dose by standard descriptive statistics (number of non-missing observations, geometric mean, mean, standard deviation, standard error of the mean, minimum, median, 25%- and 75%-quantiles, maximum, coefficient of variation (%)). Individual profiles per subject and median profiles were plotted by dose. Box and whisker plots of primary pharmacokinetic variables were generated per dose.

No formal sample size calculation was performed for this study. The sample size of the total 18 subjects, 3 subjects per sequence, was considered a standard approach for evaluation of dose–exposure–response relationship.

Assessment of the Dose–Exposure Relationship

AUC-values ($AUC_{(0–2\ h)}$ and $AUC_{(0–end)}$) and C_{max} were plotted over the dose per kg body weight

(U/kg) as well as over the total dose (U), for each subject. Corresponding plots were generated for the exposure parameters normalized to a dose of 0.15 U/kg or to 10 U, respectively.

Geometric means for AUC-values and C_{max} were plotted over dose with a regression line forced through the origin point. Geometric means for AUC-values and C_{max} normalized to a dose of 0.15 U/kg were also plotted over dose together with a regression line.

AUC values (through 2 h and through the end of sampling), C_{max}, and MRT were natural log-transformed and compared among the three dose levels using a linear ANOVA model with adjustment for dose, period, sequence, and subject within sequence (discrete model). 95% confidence intervals for pairwise dose differences were calculated and retransformed to derive the respective confidence limits for mean ratios of the pairwise treatment comparisons, that is, for 0.15 U/kg versus 0.075 U/kg, and for 0.3 U/kg versus 0.15 U/kg. Dose proportionality within the commonly accepted bioequivalence criteria (0.80–1.25) is confirmed for a doubling of the dose, when the confidence interval for a treatment ratio is within 1.60–2.50. T_{max} was analyzed by nonparametric analysis for pairwise comparisons with 95% nonparametric confidence intervals for the respective median difference in dose. A power model was applied to assess dose proportionality. This analysis was based on the actual doses of X002 administered.

Results

All 18 randomized subjects were treated, completed the study, and were evaluable for the PK component of the study. Median X002 serum concentration time profiles following the single subcutaneous doses of 0.075 U/kg, 0.15 U/kg, and 0.3 U/kg are presented in Fig. 5. X002 was rapidly absorbed, showing peak serum concentrations approximately 1 h post-dose.

All subjects showed a monotonically increasing dose–exposure relationship in $AUC_{(0\text{-}2\ h)}$, $AUC_{(0\text{-end})}$, and C_{max} for X002. T_{max} generally increased slightly with dose. Descriptive statistics for key PK parameters of X002 are given in Table 4. Boxplots for $AUC_{(0–2\ h)}$ of X002 are given in Fig. 6, showing the monotonic relationship and expected increased variability with higher doses.

Geometric means for $AUC_{(0–2\ h)}$ of X002 are shown in Fig. 7, suggesting dose proportionality.

Geometric means for $AUC_{(0–2\ h)}$ of X002 normalized on a dose of 0.15 U/kg are given in Fig. 8. The point estimates for treatment ratios together with 95% confidence intervals are presented in

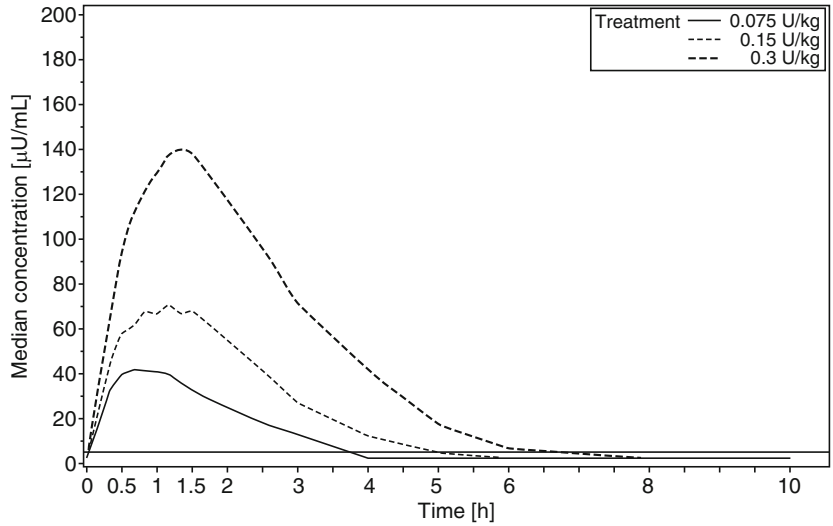


Fig. 5 Median X002 concentration (µU/mL) profiles over time after dosing

Table 4 Key PK parameters of X002 and descriptive statistics by dose

Variable (Unit)	Geometric mean (arithmetic mean ± SD)		
	0.075 U/kg ($N = 18$)	0.15 U/kg ($N = 18$)	0.3 U/kg ($N = 18$)
$AUC_{(0–2\ h)}$ (μU∗min/mL)	3,792 (3,855 ± 677)	6,676 (6,832 ± 1,461)	12,992 (13,237 ± 2,559)
$AUC_{(0–end)}$ (μU∗min/mL)	5,341 (5,372 ± 589)	11,196 (11,284 ± 1,456)	24,891 (25,076 ± 3,209)
C_{max} (μU/mL)	42 (43 ± 9)	72 (73 ± 16)	140 (142 ± 25)
MRT (min)	115 (122 ± 50)	121 (125 ± 34)	134 (136 ± 28)
T_{max} (min)[a]	47 (34−99)	57 (44−93)	72 (50−112)

[a]Median (minimum–maximum) reported

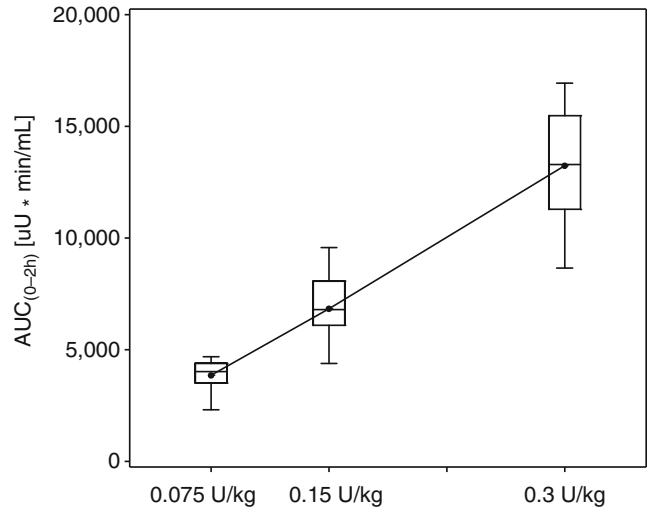


Fig. 6 Boxplots of $AUC_{(0–2\ h)}$ (μU∗min/mL) for X002 per dose

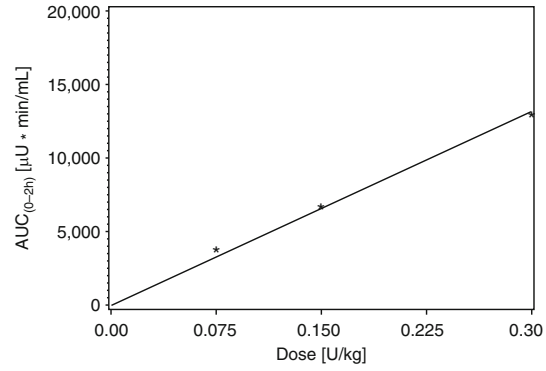


Fig. 7 Plots of geometric mean $AUC_{(0–2\ h)}$ (μU∗min/mL) for X002 per dose

Table 5. All 95% confidence intervals are fully contained within the range of 1.60–2.50, and as a consequence the 90% confidence intervals are within that range as well. Thus, exposure can be assumed to behave according to dose proportionality for a doubling of doses within the dose range investigated.

Due to the variation in body weight, the administered, weight-based doses varied between 4 U and 31 U, a 7.75-fold range. Individual plots of $AUC_{(0–2\ h)}$ over the total dose (U) are given in Fig. 9. The figure also represents the finding of dose monotony in this exposure parameter for each subject.

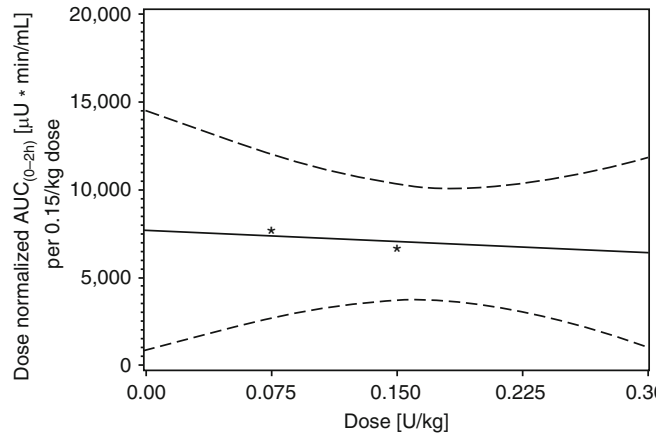


Fig. 8 Plots of geometric mean $AUC_{(0–2\ h)}$ (μU∗min/mL) for X002 – dose normalized on 0.15 U/kg

Table 5 Pairwise comparisons of key PK parameters for X002

	Point estimate (95% confidence interval)	
Variable	Ratio of 0.15–0.075 U/kg dose	Ratio of 0.3–0.15 U/kg dose
$AUC_{(0–2\ h)}$	1.8 (1.6–1.9)	1.9 (1.8–2.1)
$AUC_{(0–end)}$	2.1 (2.0–2.2)	2.2 (2.1–2.3)
C_{max}	1.7 (1.6–1.9)	2.0 (1.8–2.1)

Individual plots of $AUC_{(0–2\ h)}$ dose normalized for 10 U, over the total dose (U) are given in Fig. 10.

Each deviation from a horizontal line indicates a deviation from dose proportional results for an individual. However, deviations are generally small.

The results of the analysis of the power model for the dose-proportionality of increasing total doses of X002 are presented in Table 6. The 90% confidence intervals for the dose normalized ratios for a value of $r = 2$ are fully contained within the classical bioequivalence range, confirming the dose proportional behavior of exposure for a doubling of dose. A factor of 7.75 between the highest and the lowest total dose was measured instead of the expected factor of 4. According to the results from the power model, dose proportionality could not be shown for the full, 7.75-fold dose range. However, it was determined by extrapolation that up to a 3.7-fold (r) for $AUC_{(0–2\ h)}$, up to 5.0 r for $AUC_{(0–end)}$, and up to 3.0 r for C_{max} satisfied criteria for dose-normalized dose-proportionality. These dose ranges, for which dose proportionality can be assumed, were considered to cover clinical needs and expected range of actual dose adjustments of the drug.

Discussion

The study was designed as an explorative study, because a strictly confirmatory study was not deemed necessary at this stage of the development of X002. In another setting this might be required. For this case, a single primary analysis used to decide about dose proportionality will have to be specified upfront.

Based on the crossover design, the study design allowed to assess dose–exposure relationship and to investigate dose proportionality based on intra-subject comparisons, which is not possible in any setting with parallel groups. In addition, due to the dosing per kg of body weight and the variability of body weight between subjects, a broader range of doses could be observed.

Due to the crossover design with complete blocks, the number of doses investigated had to be limited to a small number. For situations where a broader dose range has to be investigated, crossover designs with incomplete blocks could be set up.

Different statistical approaches were used and delivered consistent findings. The study allowed to draw conclusions for a clinical relevant range of doses.

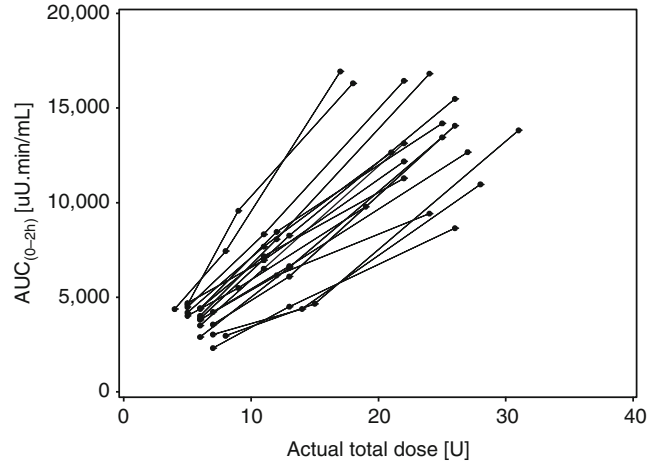


Fig. 9 Individual $AUC_{(0–2\ h)}$ (μU∗min/mL) over actual total dose (U) – per subject

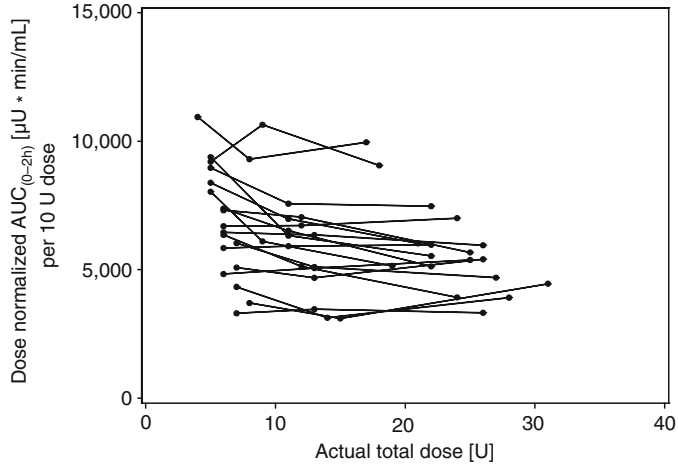


Fig. 10 Individual $AUC_{(0–2\ h)}$ (μU∗min/mL) over actual total dose (U) – per subject and dose normalized on 10 U

Table 6 Estimates with 90% CI for r-fold increase in dose

		Ratio		Dose-normalized ratio	
Parameter	Dose ratio	Estimate	90% CI	Estimate	90% CI
$AUC_{0–2\ h}$	r = 2	1.8440	1.7784, 1.9119	0.9220	0.8892, 0.9559
	r = 7.75	6.0964	5.4781, 6.7838	0.7866	0.7069, 0.8753
	$r^* = 3.7324$	3.1985	2.9859, 3.4260	0.8570	0.8000, 0.9179
	β-Estimate	0.8828	0.8306, 0.9350		
$AUC_{0–end}$	r = 2	2.1576	2.1145, 2.2016	1.0788	1.0572, 1.1008
	r = 7.75	9.6960	9.1353, 10.2917	1.2511	1.1788, 1.3280
	$r^* = 5.0073$	5.9723	5.6989, 6.2591	1.1927	1.1381, 1.2500
	β-Estimate	1.1094	1.0803, 1.1385		
C_{max}	r = 2	1.8164	1.7425, 1.8936	0.9082	0.8713, 0.9468
	r = 7.75	5.8315	5.1580, 6.5940	0.7524	0.6655, 0.8508
	$r^* = 3.0718$	2.6284	2.4574, 2.8115	0.8557	0.8000, 0.9153
	β-Estimate	0.8611	0.8012, 0.9211		

(r^*) = highest dose ratio compatible with dose proportionality by the equivalence approach

Example 3: Confirmatory Assessment of Dose Linearity/Proportionality – Single and Repeated Dose Crossover Design

Purpose and Rationale

The aim of a confirmatory dose proportionality study is to assess the PK of a drug at doses bracketing the anticipated therapeutic dose (i.e., usually at one dose below therapeutic dose and one dose above), using the most appropriate design (i.e., crossover, within-subject comparison) and an adequate number of subjects, in order to assess what are the variations in exposure when the dose needs to be adjusted, for example, in special populations or in case of concomitant medications.

The design of the confirmatory dose linearity/proportionality study during advanced clinical development for candidate drug "X003" is presented below. The study description is limited to the primary objective of pharmacokinetic data, although safety/tolerability data were also obtained as secondary objective.

In the case of candidate drug X003 the exploratory assessment of dose linearity/proportionality from the first-in-human study did not allow accurate characterization of the deviation from proportionality of X003 pharmacokinetics because of the parallel group design, low number of subjects, and relatively high variability of X003 PK. For a twofold increase in dose, there was a threefold increase in exposure with large 95% confidence interval of 1.9–4.1. In addition, no multiple dose PK data were available at doses below the anticipated therapeutic dose of 400 mg administered twice daily (BID).

Study Design

This was a single center, randomized, non-placebo-controlled, open-label, single and repeated BID oral dose, three treatment, three period, crossover study with a washout of 14 days between periods.

Healthy young white men aged between 18 and 35 years with a body weight between 50 kg and 90 kg and a body mass index between 18 kg/m^2 and 28 kg/m^2 were included. Eighteen subjects were accrued in order to have at least 12 subjects complete the study (US FDA 2001). Subjects were randomized and treated with 200 mg, 400 mg, and 800 mg of X003. The tablets were administered with 200 mL of non-carbonated water on Day 1 at 8:00 a.m. after the end of a standardized breakfast (single dose), on Days 2–13 at 8:00 a.m. and 8:00 p.m., and on Day 14 at 8:00 a.m. after the end of a standardized meal. Tablets containing 100 mg and 400 mg X003 were used to provide 200 mg (2 × 100 mg), 400 mg (1 × 400 mg), and 800 mg (2 × 400 mg) doses.

Subjects were hospitalized the evening before first drug administration (Day 1) to Day 2 (morning) and from Day 13 (evening) to Day 15 (morning) of each period. The subjects were discharged and then visited the study unit every morning and every evening from Days 2 to 13 for blood-sampling and/or study drug administration.

After the last dose was administered on the morning of Day 14, subjects returned for blood sampling on the evening of Day 15, and on the mornings of Days 16 through 18. The duration of study participation for each subject was in total 12–15 weeks: 3–21 days for subject selection, 14 days for period 1, 14 days for washout period 1, 14 days for period 2, 14 days for washout period 2, 14 days for period 3, and 10–12 days for follow-up period.

Blood samples to determine X003 were collected before dosing and at 0.5 h, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h, 8 h, 10 h, 12 h, 24 h, 36 h, 48 h, 72 h, and 96 h after dosing on Days 1 and 14, and before morning dosing for trough determinations on Days 6, 7, 8, 9, 10, and 12. X003 concentrations in plasma were determined using a validated liquid chromatography–tandem mass spectrometry (LC–MS/MS) method with a limit of quantification (LOQ) for X003 of 0.5 ng/mL. Standard safety/tolerability criteria (e.g., adverse events, biochemistry, hematology, urinalysis, ECG, and vital signs) were assessed in this study but are not the subject of this chapter.

Primary and secondary objectives were to assess the deviation from dose proportionality and safety/tolerability, respectively, of X003 after 200 mg, 400 mg, and 800 mg single and twice daily repeated oral doses of X003 for 10 days. The following pharmacokinetic parameters were assessed using non-compartmental analysis:

Day 1:
Maximum concentration (C_{max}), time to maximum concentration (t_{max}), area under the concentration time curve from time of drug administration to time 12 h ($AUC_{(0–12\ h)}$) and to last quantifiable concentration time point (AUC_{last}), area under the concentration–time curve from time of drug administration extrapolated to infinity (AUC), terminal elimination half-life ($t_{1/2z}$).
Day 14:
C_{max}, T_{max}, $AUC_{(0–12\ h)}$, $t_{1/2z}$.
Days 6, 7, 8, 9, 10, and 12: trough concentration (C_{trough}).

Plasma concentrations and pharmacokinetic parameters were listed by standard descriptive statistics (N, mean, SD, SE, min, median, max, CV%, geometric mean) for each dose level on Days 1 and 14. X003 C_{max} (Days 1 and 14), AUC (Day 1), $AUC_{(0–12\ h)}$ (Day 14), and $t_{1/2z}$ were log-transformed, and t_{max} was rank-transformed.

C_{max} and AUC at Day 1, and C_{max} and $AUC_{(0–12\ h)}$ at Day 14 were analyzed with a "random intercepts and random slopes" mixed model in SAS PROC MIXED separately for Days 1 and 14. The parameters were assumed to follow a multiplicative power model, which is equivalent to the log-transformed power model, and has the form:

$$\begin{aligned}\mathrm{Log}\,(\mathrm{parameter}) = {} & [\mathrm{Log}\,(\alpha)\mathrm{i} - \mathrm{Log}\,(\alpha)] \\ & + [\beta_i - \beta] \\ & \times \mathrm{Log}\,(\mathrm{dose}) \\ & + \mathrm{period} + \mathrm{Error} \end{aligned} \quad (12)$$

In the model, Log(α) and β were the estimates of intercept and slope, respectively, estimated by generalized least squares (GLS) with restricted maximum likelihood (REML) estimates of random effects.

Goodness-of-fit was assessed by visual inspection of residuals plots. If there was no evidence of lack-of-fit (residuals randomly distributed around the origin), estimates for β with 90% confidence intervals were computed within the mixed model framework. Estimates with 90% CI for PK parameter increases associated with an r-fold (r = 2 and r = 4 = highest dose/lowest dose) increase in dose were obtained by exponentiating r to the powers of the ($\widehat{\beta}$) and confidence limits (Eq. 10) and also, subsequently, converting these to a dose-normalized scale by dividing by r. Three cases were considered: r = 2 (doubling of dose), r = 4 (high/low dose), and the maximum dose ratio "r" compatible with dose proportionality by an equivalence approach (i.e., 90% confidence limits for the dose normalized increase in the PK parameter is within 0.80–1.25).

Dose Effect

Differences of $t_{1/2z}$ and T_{max} between doses were tested for significance with p-values from the linear fixed effects model with fixed terms for sequence, period, day, dose, and the dose-by-day interaction, and a random term for subjects within sequence.

If the dose-by-day interaction was not significant ($p \geq 0.05$), the interaction term was dropped from the model and the model was refit. The p-values for the dose and day effects were reported in the context of the reduced model.

If the dose-by-day interaction was significant ($p < 0.05$), the p-value for the dose effect was computed in a mixed effects model. Days 1 and 14 were fit separately with fixed terms for sequence, period and dose, and a random term for subjects within sequence. The p-value for the day effect was computed in a mixed effect model, fit separately for each dose with fixed terms for sequence and day and a random term for subjects within sequence. Each model described above was fit by GLS with REML estimates of random effects, using SAS PROC MIXED.

Accumulation Effects

To assess accumulation effects from Days 1 to 14, the PK parameters were analyzed with a mixed effects model with fixed terms for sequence, period, day, dose and the dose-by-day interaction, and a random term for subjects within sequence. The model was fit again by GLS with REML estimates of random effects using SAS PROC MIXED.

If the dose-by-day interaction was significant ($p < 0.05$), accumulation was assessed for each dose group separately within the mixed model framework. Otherwise, the term was dropped from the model and accumulation effects were assessed for all dose groups. For C_{max} and $AUC_{(0–12\ h)}$, the difference in means between Days 14 and 1, with 95% CI, was computed within the mixed model framework and converted to an accumulation ratio of adjusted geometric means by the antilog transformation.

Steady State

The occurrence of steady state for X003 was assessed separately at each dose level by fitting the trough values with a nonlinear mixed effects model using the SAS NLMIXED procedure. The BID trough values corresponding to days 6, 7, 8, 9, 10, and 12 were utilized for the steady-state assessment. Model goodness-of-fit was evaluated by graphical inspection of the fitted curves and the within-subject residuals, and by graphical inspection of histograms of the estimated subject-specific random parameters (e.g., evidence of outliers, or subpopulations by gender, age, metabolizer status). For each subject and each dose level, the day at which 90% of the estimated subject-specific steady-state trough concentration is reached was predicted from the model. The overall time-to-steady state was determined as the 50th percentile (for average steady state) and the 90th percentile (for individual steady state) of these individual predicted values at each dose level. The 95% CIs for the 50th and 90th percentiles were calculated by nonparametric methods.

Results

Plasma Concentrations and Pharmacokinetic Parameters

Mean (SD) X003 plasma concentration versus time curves observed after single (Day 1) and repeated BID (Day 14) oral administrations of X003 are presented in Fig. 11. X003 plasma concentrations were higher following a 10-day repeated BID oral administration (Day 14) compared to the single dose at Day 1 and reached peak levels approximately 5 h post-dose. Twelve hours post-dose plasma concentrations were not below LOQ.

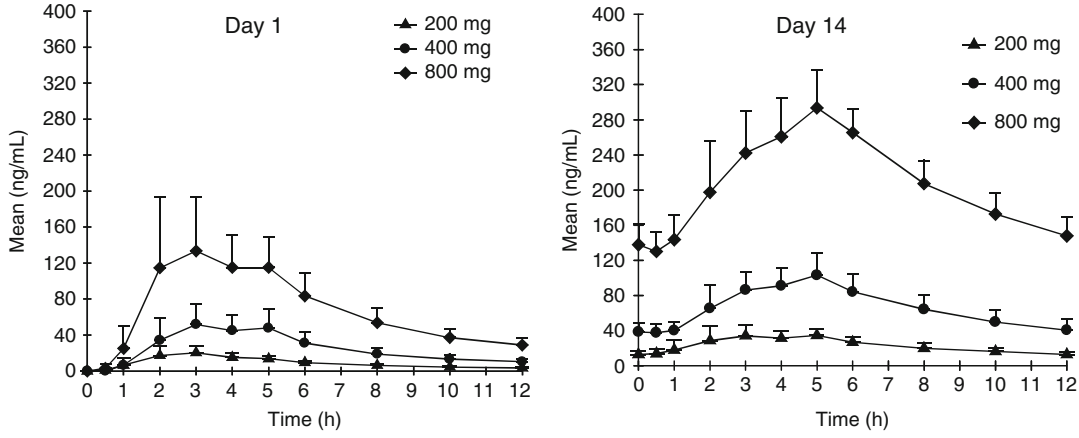


Fig. 11 Mean (SD) X003 plasma concentrations on Day 1 after a single and on Day 14 after repeated BID oral administration of X003 (linear scale)

Table 7 Mean (CV%) X003 pharmacokinetic parameters observed after a single and repeated oral administration of X003 BID

PK parameters	Day	200 mg $N = 17$	400 mg $N = 16$	800 mg $N = 17$
C_{max} (ng/mL)	1	23.1 (38)	67.2 (36)	162 (40)
	14	40.3 (30)	111 (17)	298 (13)
t_{max} (h)[a]	1	3 (2; 3)	3 (2; 5)	3 (2; 6)
	14	5 (2; 5)	5 (3; 6)	5 (2; 6)
$AUC_{0-12\ h}$(ng*h/mL)	1	111 (24)	310 (28)	846 (27)
	14	276 (23)	798 (19)	2,510 (12)
$t_{1/2z}$ (h)	1	9.81 (33)	17.6 (56)	19.6 (33)
	14	26.9 (32)	30.0 (29)	31.2 (32)
AUC (ng*h/mL)	1	160 (27)	474 (33)	1,310 (26)

[a]Median value (min; max)

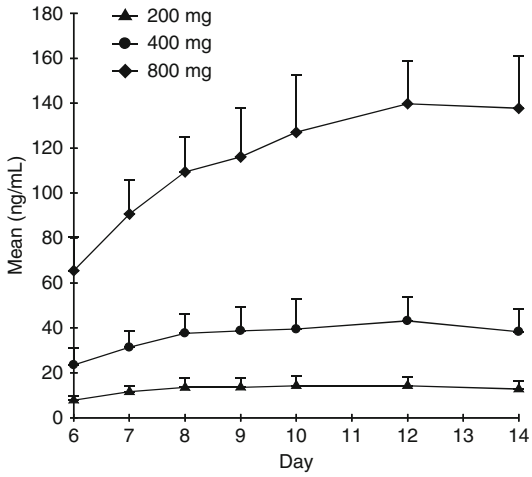


Fig. 12 Mean (SD) X003 trough concentrations from Days 6 to 14 during repeated BID oral administration of X003

A summary of X003 main pharmacokinetic parameters observed on Day 1 after a single oral administration and on Day 14 after a 10-day repeated BID oral administration of X003 is presented in Table 7.

Steady State

Mean (SD) X003 C_{trough} observed from Days 6 to 14 during repeated BID X003 administrations are graphically summarized in Fig. 12. Individual steady state, as expressed by 90th percentile, was reached after 3–5 treatment days. Average steady state, as expressed by 50th percentile, was reached after 3–4 treatment days, regardless of the dose.

Table 8 Accumulation ratio (R_{ac}) with 95% CI for X003 C_{max} and $AUC_{0-12\ h}$

PK parameters	R_{ac} estimate	95% CI
C_{max} (ng/mL)	1.84	1.65; 2.04
$AUC_{0-12\ h}$ (ng*h/mL)	2.72	2.52; 2.94

Accumulation

In the accumulation assessment, the dose-by-dose interaction was not significant for C_{max} and $AUC_{(0-12\ h)}$, allowing the assessment of a single accumulation ratio across doses for each PK parameter of X003. After a 10-day repeated BID oral administration, an accumulation ratio of 1.84 (95% CI, 1.65–2.04) in C_{max} and 2.72 (2.52–2.94) in $AUC_{(0-12\ h)}$ was observed, regardless of the administered dose (Table 8).

Dose Proportionality

Results of the dose proportionality assessment, at Days 1 and 14, are summarized in Table 9.

C_{max} and t_{max}: After a 10-day repeated BID oral doses of X003, doses ranging from 200 mg to 800 mg X003 C_{max} values were reached 5 h after drug intake, and no significant dose, day, or dose-by-day interaction effects were observed on T_{max}. As measured by ratio estimate and associated 90%

Table 9 Ratio estimates and associated 90% CIs for dose proportionality assessed for a twofold increase in dose

PK parameters	Day	Estimate	95% CI
C_{max} (ng/mL)	1	2.62	2.39; 2.88
C_{max} (ng/mL)	14	2.77	2.62; 2.93
AUC (ng/mL)	1	2.86	2.67; 3.06
$AUC_{0–12\ h}$ (ng*h/mL)	14	3.06	2.92; 3.20

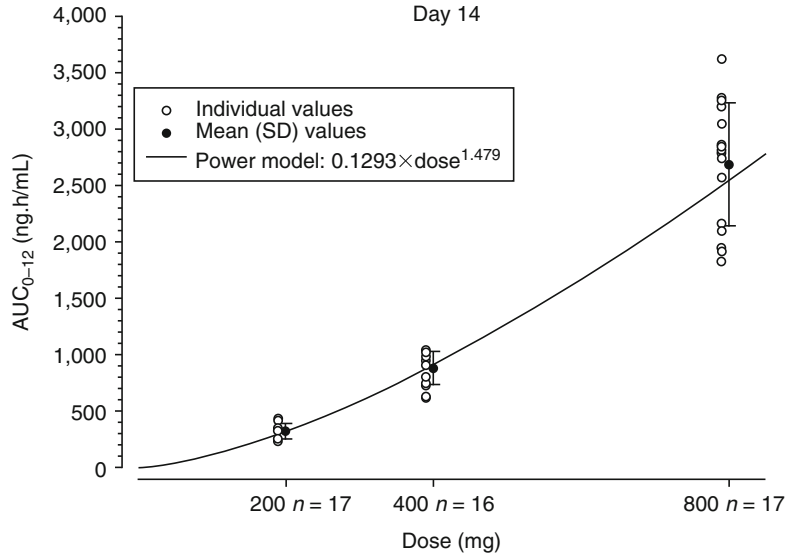


Fig. 13 Individual and mean (SD) X003 $AUC_{(0–12\ h)}$ values observed after repeated BID oral administration of X003

CI, X003 C_{max} increased more than expected by dose proportionality: a twofold increase in dose led to a 2.62 (2.39–2.88) and 2.77 (2.62–293) increase in X003 C_{max} on Days 1 and 14, respectively.

$AUC_{(0–12\ h)}$: Individual and mean (SD) values of X003 $AUC_{(0–12\ h)}$ on Day 14 are presented in Fig. 13. As measured by ratio estimate and associated 90% CI, X003 AUC increased more than expected by dose proportionality (Table 9): a twofold increase in dose led to a 2.86 (2.67–3.06) and 3.06 (2.92–3.20) increase in X003 AUC on Days 1 and 14, respectively.

$t_{1/2z}$: The p-value for the dose-by-day interaction was significant in the analysis of $\log(t_{1/2z})$ for X003 ($p > 0.001$). There was a dose effect at Days 1 and 14, and a Day effect for X003 at 200 mg, 400 mg, and 800 mg. From 200 mg to 800 mg, the mean $t_{1/2z}$ increased significantly from 9.8 h to 19.6 h at Day 1 and from 26.9 h to 31.2 h at Day 14 (Table 7).

Discussion

The described evaluation provides confirmatory data on the lack of dose linearity/proportionality for X003, which allows dose adjustment recommendations in the submission package of a drug. This type of study also supports bracketing approaches in bioequivalence studies in which different formulations and dose strengths are to be tested. In several cases it was accepted by authorities that bioequivalence for only the lowest and highest dose strength had to be demonstrated (US FDA 2014). In a three-period crossover design, the investigator should recruit approximately 50% additional subjects in order to have enough subjects completing all three periods and to guarantee appropriate PK evaluation. Overall, the crossover design is preferable for the assessment of dose linearity/proportionality because it minimizes the variability in PK parameters.

In case steady-state conditions are known to be achieved within a specified dosing regimen, it might be sufficient to evaluate dose proportionality at steady state or alternatively single dose conditions only. The investigator should also consider evaluating dose proportionality for major, active metabolites.

References and Further Reading

Cawello W, Brett M, Weimann H-J, Zimmerman H, Pabst G, Sierakowski B, Gieschke R, Baumann A (1999) Parameters for compartment-free pharmacokinetics: standardisation of study design, data analysis, and reporting. Shaker Verlag, Aachen

EMEA CPMP (2010) Guideline on the investigation of bioequivalence (CPMP/EWP/QWP/1401/98 Rev. 1/Corr)

EMEA CPMP (2014) Guideline on the pharmacokinetic and clinical evaluation of modified release dosage forms (EMA/CPMP/EWP/280/96 Corr1)

Frick A, Marshallsay C, Steinstraesser WR (2006) Clinical studies – typical designs. In: Vogel HG, Maas J, Mayer D, Hock F (eds) Drug discovery and evaluation: safety and pharmacokinetic assays. Springer, Berlin/Heidelberg

Gough K, Hutchison M, Keene O, Byrom B, Ellis S, Lacey L, McKellar J (1995) Assessment of dose proportionality: report from the Statisticians in the Pharmaceutical Industry/Pharmacokinetics UK Joint Working Party. Ther Innov Regul Sci 29(3):1039–1048

Hummel J, McKendrick S, Brindley C, French R (2009) Exploratory assessment of dose proportionality: review of current approaches and proposal for a practical criterion. Pharm Stat 8:38–49. https://doi.org/10.1002/pst.326

Ludden TM (1991) Nonlinear pharmacokinetics: clinical implications. Clin Pharmacokinet 20:429–446

Sethuraman VS, Leonov S, Squassante L, Mitchell TR, Hale MD (2007) Sample size calculation for the Power Model for dose proportionality studies. Pharm Stat 6:35–41. https://doi.org/10.1002/pst.241

Smith BP (2004) Assessment of dose proportionality. In: Bonate P (ed) Pharmacokinetics in drug development: vol 1, clinical study design and analysis. AAPS Press, Arlington

Smith BP, Vandenhende FR, DeSante KA, Farid NA, Welch PA, Callaghan JT, Forgue ST (2000) Confidence interval criteria for assessment of dose proportionality. Pharm Res 17:1278–1283

US FDA (2001) Guidance for industry: statistical approaches to establishing bioequivalence. Office of Training and Communications, Division of Communications Management, Maryland

US FDA (2014) Draft guideline for industry: bioavailability and bioequivalence studies submitted in NDAs or INDs – general considerations, Mar 2014

Effects of Food Intake

38

Teodora Handjieva-Darlenska

Contents

The management of many diseases requires drug treatment, particularly involving the use of multiple drugs. Food-drug interactions can affect or even change the effects of drugs, and the therapeutic effects or side effects of medications can affect the nutritional status. Alternatively, the diet and the use of supplements or the nutritional status of the patient can decrease a drug's efficacy or increase its toxicity.

The terms drug-nutrient interaction and food-drug interaction are often used interchangeably. In fact, drug-nutrient interactions are some of the many possible food-drug interactions. Drug-nutrient interactions include specific changes to the pharmacokinetics of a drug caused by a nutrient(s) or changes to the kinetics of a nutrient(s) caused by a drug. A food-drug interaction is a broader term that also includes the effects of a medication on nutritional status. Nutritional status may be impacted by the side effects of a drug, which could include an effect on appetite or the ability to eat. Food-drug interaction studies are important to evaluate appropriate dosing, timing, and formulation of new drug candidates. It is crucial that healthcare professionals take into consideration the drug-nutrient interactions in order to optimize the effectiveness and minimize the toxicities of medications.

Awareness of these interactions enables the healthcare professional and patient to work together to avoid or minimize problems (Table 1).

T. Handjieva-Darlenska (✉)
Department of pharmacology and toxicology, Medical Faculty, Medical University of Sofia, Sofia, Bulgaria
e-mail: teodorah@abv.bg

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_6

Table 1 Benefits of minimizing drug interactions. (Adapted from Pronsky 2004)

Medications achieve their intended effects
Patients do not discontinue their drugs
The need for additional medication is minimized
Fewer caloric or nutrient supplements are required
Adverse side effects are avoided
Optimal nutritional status is preserved
Accidents and injuries are avoided
Disease complications are minimized
The cost of healthcare services is reduced
There is less professional liability

Pharmacological Aspects of Food-Drug Interactions

Medication is administrated to produce a pharmacological effect in the body or, more specifically, in a target organ or tissue. To achieve this goal, the drug must move from one site of administration to the bloodstream and eventually to the site of drug action. In due course, the drug may be changed to activate or inactivate metabolites and ultimately eliminated from the body. An interaction between the drug and food, a food component, or a nutrient can alter this process at any point. Food-drug interactions may be divided into two broad types: (1) **pharmacodynamic interactions**, which affect the pharmacological action of the drug, and (2) **pharmacokinetic interactions**, which affect the movement of the drug into, around, or out of the body.

Risk Factors for Food-Drug Interactions

Patients must be assessed individually for the effect of food on drug action and the effect of drugs on nutritional status. Interactions can be caused or complicated by **polypharmacy**, nutritional status, genetics, underlying illness, special diets, nutritional supplements, tube feeding, herbal or phytonutrient products, alcohol intake, drugs of abuse, nonnutrients in food, excipients in drugs or food, allergies, or intolerances. Poor patient's compliance and physician's prescribing patterns further complicate the risk. It is well established that in older people food-drug interactions could often occur due to taking of multiple drugs (prescribed and over-the-counter). In this age group, the risk of such interactions is higher than in young patients. All this is a result of physical changes related to aging such as an increase in the ratio of fat tissue to lean body mass, a decrease in liver mass and blood flow, an impairment of kidney and liver functions, illness, and endocrine and cognitive dysfunctions. Malnutrition and dehydration affect pharmacokinetics. It is known that the use of herbal and/or phytonutrient products has increased significantly by older people.

Central nervous system adverse effects of drugs can interfere with the ability or desire to eat. Recognition of these problems as a drug side effect, rather than a consequence of disease or aging, is often overlooked. It is important to evaluate the intake of nutrients that could interact when a specific medication is used. These could be vitamin K with warfarin, calcium and vitamin D with alendronate, and potassium, sodium, and magnesium with loop diuretics as furosemide.

Pharmacogenomics

Pharmacogenomics is defined as "genetically determined variations that are revealed solely by the effects of medications" (Tischio 1995). A well-known examples of a food-drug interaction ramification are G6PD (glucose-6-phosphate dehydrogenase) enzyme deficiency, warfarin resistance, and slow inactivation of isoniazid or phenelzine.

Acetylation is a conjugation reaction that metabolizes and inactivates amines, hydrazines, and sulfonamides. There are two types of acetylators: the "slow acetylators" who metabolize these medications more slowly than average due to inherited lower levels of the hepatic enzyme acetyl transferase. As a result, the unacetylated drug levels stay higher for long periods of time; the "rapid acetylators" show the opposite pattern of metabolization (Zastrow 2012). The half-life of isoniazid for rapid

acetylators is about 70 min, whereas the half-life for slow acetylators is more than 3 h (Roth 1995). More importantly, a dose prescribed normally for rapid acetylators could be toxic for slow acetylators. Thereafter, the elevated blood levels of affected drugs on slow acetylators increase the potential risk for food-drug interactions. Slow inactivation of phenelzine, a monoamine oxidase (MAO) inhibitor, increases the risk for hypertensive crisis if foods high in tyramine (e.g., cheese) are used. Dapsone and hydralazine are also metabolized by acetylation and affected by inherited differences in acetylase enzymes.

Another example is the deficiency of G6PD which is an X-chromosome-linked deficiency of G6PD enzyme in red blood cells. It is demonstrated that it can lead to neonatal jaundice, hemolytic anemia, or acute hemolysis. It is called favism, and it is generally common in African, Middle Eastern, and Southeast Asian populations. Fava beans or pollen could cause acute hemolysis in some G6PD-deficient persons, particularly those of Mediterranean origin. Aspirin, sulfonamides, and antimalarial drugs can cause hemolysis and acute anemia. Rees et al. (1993) demonstrated acute hemolysis induced in G6PD deficiency due to high doses of vitamin K or vitamin C.

Effects of Food on Drug Treatment

Drug Absorption

It has been well demonstrated that the presence of food in the stomach or lumen of the intestinal tract increases the risk to reduction of drug absorption. The absorption of alendronate and risedronate (anti-osteoporosis drugs) is almost negligible if these medications are coadministrated with food, particularly with coffee or orange juice.

The absorption of the iron from supplements can be declined by 50% when coadministrated with food. It is prescribed to take iron with 200 ml of water on an empty stomach. There are several food or nutrients that could affect iron absorption. Thus, iron should not be taken together with bran, eggs, high-phytate foods, fiber supplements, tea, coffee, dairy products, or calcium supplements.

Various mechanisms may contribute to the decrease in the rate or extent of drug absorption in the presence of food or nutrients. The presence and type of meal or food ingested influence the rate of gastric emptying. Gastric emptying could be delayed by the consumption of high-fiber meals and meals with high fat content. In general, a delay in drug absorption is not clinically significant as long as the extent of absorption is not affected. Absorption of antibiotics or analgesics may be clinically significant problems with delayed absorption. Chelation reactions occur between certain drugs and divalent or trivalent cations such as iron, calcium, magnesium, zinc, or aluminum. As a result, the absorption of drugs could be reduced by chelation with one of these ions. A well-known example is the antibiotics ciprofloxacin and tetracycline which form insoluble complexes with calcium in dairy products; calcium, magnesium, zinc, or iron supplements; or aluminum in antacids (Neuhofel et al. 2002). To minimize these interactions, it is advisable to give the drug at least 2 h before 6 h after the mineral administration.

Adsorption or adhesion to food or a nutrient is another mechanism by which drug absorption is altered. For example, a high-fiber diet could decrease the absorption of amitriptyline (a tricyclic antidepressant). Thus, the therapeutic effect is affected because of the adsorption of the drug to the fiber. It was established that digoxin, a cardiovascular drug, should not be coadministrated with wheat bran or oatmeal, both high-phytate foods.

Gastrointestinal pH is an important factor for drug absorption. Any situations that alter the gastric pH (e.g., achlorhydria or hypochlorhydria) may decrease drug absorption. An example is the defect of ketoconazole to treat *Candida* infection in patients with human immunodeficiency virus (HIV). Ketoconazole achieves best absorption in an acid medium. Welage et al. (1995) showed that achlorhydria is a common status in HIV patients, and ketoconazole tablets could not be dissolved leading to impaired drum absorption and no efficacy. This is also a concern in patients with hypochlorhydria, who are taking antacids,

histamine 2 receptor antagonists (famotidine), or proton-pump inhibitors (omeprazole). It was established that ketoconazole taken together with acid-containing beverage (cola) may improve bioavailability in these patients.

The presence of food in the stomach could alter the absorption of some drugs. The antibiotic cefuroxime axetil and the antiretroviral drug saquinavir are prescribed to be taken after a meal to reach an effective level.

Drug Distribution

Albumin is the most important drug-binding protein in the blood. Low serum albumin levels, often the result of inadequate protein intake and poor nutrition, provide fewer binding sites for highly protein-bound drugs. Fewer binding sites mean that a larger free fraction of drug will be present in the serum. Only the free fraction (which is unbound) of a drug is able to exert a pharmacologic effect at the target organ. Patients with albumin levels below 3.0 g/dl are at increased risk for adverse effects from highly protein-bound drugs. Usual adult doses of highly protein-bound drug in such patients may produce more pronounced pharmacologic effects than the same dosage in persons with normal albumin levels. A lower dose of such drug is recommended in patients with low albumin levels.

Anticoagulant warfarin, which is 99.9% serum protein bound, and anticonvulsant phenytoin, which is greater than 90% protein bound, are common drugs used in older individuals. Low albumin levels tend to be more common in elderly. In the case of warfarin, higher levels of free drug lead to risk of excessive anticoagulation and bleeding. Phenytoin toxicity can result from higher levels of free phenytoin.

Drug Metabolism

Enzyme systems in the intestinal tract and the liver, although not the only sites of drug metabolism, account for a large portion of the drug-metabolizing activity in the body. Food can both inhibit and enhance the metabolism of medication by altering the activity of these enzyme systems. A diet high in protein and low in carbohydrates can increase the hepatic metabolism on the antiasthma drug theophylline. The suspected mechanism of increased clearance of this drug is the induction of the hepatic enzyme system responsible for metabolizing the drug (Walter-Sack and Klotz 1996).

Grimm et al. (2018) have demonstrated that Grapefruit juice can cause slower gastric emptying and increase intestinal filling and thus could potentially affect drug pharmacokinetics (enhanced or altered absorption). A substance found in grapefruit and grapefruit juice can inhibit the intestinal metabolism of drugs such as calcium channel antagonists that are dihydropyridine derivatives (Bailey et al. 1994) and/or some HMG-CoA reductase inhibitors such as Simvastatin (Lilja et al. 2000). Grapefruit inhibits the cytochrome P-450 3A4 enzyme system responsible for the oxidative metabolism of many orally administrated drugs. The interaction appears to be clinically significant for drugs with low oral bioavailability, which are substantially metabolized and inactivated in the intestinal tract by the cytochrome P-450 3A4 enzyme in the intestinal wall. When grapefruit juice is ingested, the metabolizing enzyme is irreversibly inhibited, which reduces the normal metabolism of the drug. This reduction in metabolism allows more of the drug to reach the systemic circulation, and the increase in blood levels of unmetabolized drug results in a greater pharmacologic effect and possible toxicity. Unfortunately, the effects of grapefruit on intestinal cytochrome P-450 last up to 72 h, until the body can reproduce the enzyme (Lilja et al. 2000). Thus, separating the ingestion of the grapefruit and the drug does not appear to alleviate the interaction. Kogure et al. (2014) developed a method that could predict the AUC ratio, along with its interindividual variation, from the pharmacokinetic profile in the absence of grapefruit juice. This tool could be used in the daily medical practice. The grapefruit juice-calcium channel blocker interaction has been known since 1989. This interaction is related to both flavonoid and nonflavonoid components of grapefruit juice

interfering with enterocyte CYP3A4 activity. In the process, presystemic clearance of susceptible drugs decreases and bioavailability increases. The most prominent interaction occurs with felodipine (Sica 2006).

Another well-known example is the Seville oranges which are used in some marmalades, but not in commercial orange juice production; pomelos and tangelos may also cause similar reactions (Egashira et al. 2003; Malhotra et al. 2001). The interaction is not significant in drugs that are not metabolized by cytochrome P-450 3A4 in the intestinal wall, such as the HMG-CoA reductase inhibitors pravastatin and fluvastatin.

Competition between food and drugs such as propranolol and metoprolol for metabolizing enzymes in the liver may alter the first-pass metabolism of these medications. Drugs absorbed from the intestinal tract by the portal circulation are first transported to the liver before they reach the systemic circulation. Because many drugs are highly metabolized during the first pass through the liver, only a small percentage of the original dose is actually available to the systemic circulation and the target organ. When food and drug compete for the same metabolizing enzymes in the liver, more of the drug is likely to reach the systemic circulation, which can lead to toxic effect if the dose of the drug is titrated to an optimal level in the fasting state.

Drug Excretion

Food and nutrients can alter the reabsorption of drugs from the renal tubule. Reabsorption of the antimanic agent lithium is closely associated with the reabsorption of sodium. When sodium intake is low or when a patient is dehydrated, the kidneys will reabsorb more sodium. In the person treated with lithium, the kidney will reabsorb lithium as well as sodium under these conditions. Higher Lithium levels and possible toxicity will result. When excess sodium is ingested, the kidneys will eliminate more sodium in the urine and likewise more lithium. This will produce lower lithium levels and possible therapeutic failure.

Drugs that are weak acids or basis are reabsorbed from the renal tubule into the systemic circulation only in nonionic state. An acidic drug is largely in the nonionic state in the urine with an acidic pH, whereas a basic drug is largely in a nonionic state in the urine with an alkaline pH. A change in urinary pH by food may change the amount of drug existing in the nonionic state. In this way, this increases or decreases the amount of drug available for tubular reabsorption. Foods such as milk, most fruits (including citrus fruits), and most vegetables are urinary alkalinizers. The interaction is most likely to be clinically significant when the diet is composed exclusively of a single food or food group. Patients should be advised by a healthcare professional if one starts a fad diet.

Food or nutrients can alter the intended pharmacological action of a medication by enhancing the medication effects or by opposing it. The classic example of such changed drug effect is the interaction between the monoamine oxidase inhibitors (MAOIs) and neurotransmitters such as dopamine, histamine, and tyramine. These biologically active amines are normally present in many foods. They are rapidly deaminated by MAO and diamine oxidases. Inhibition of MAO by medication prevents the breakdown of tyramine and other pressor agents. Tyramine is a vasoconstrictor and, thus, raises blood pressure. It was established that a high consumption of food containing tyramine (e.g., cheeses and meats) together with MAOI antidepressants could potentially cause a hypertensive crisis and even death (Gardner et al. 1996).

Another example is caffeine in foods and beverages which increases the side effects of stimulant drugs such as methylphenidate, amphetamines, and theophylline. This interaction results in nervousness, tremor, and insomnia. In addition, caffeine could oppose or even interact the anti-anxiety effect of tranquilizers (e.g., lorazepam).

Warfarin is an oral anticoagulant which mechanism of action is to reduce the hepatic production of vitamin K-dependent clotting factors. In order to achieve an optimal level of anticoagulation, a balance should be maintained between the dose of the drug and the ingestion of vitamin K. It is mandatory to counsel a person taking warfarin to

prohibit to consumption of high vitamin K foods such as dark green, leafy vegetables (Booth et al. 1997). It was established that the ingestion of other foods or ingredients could also alter the anticoagulant effect of warfarin. Such ingredients are found in onions, garlic, vitamin E supplements in doses greater than 400 UI, and certain herbal products (dong quai, which contains coumarin-like substances, and ginseng which has an anti-platelet activity).

Alcohol

It is well-known that Ethanol together with certain medications will produce additive toxicity which could affect various organs and systems. One of the examples is the combination between ethanol and central nervous system (CNS) depressant medications such as benzodiazepines (diazepam, etc.) and barbiturates. This interaction leads to excessive drowsiness, incoordination, and other signs of CNS depression. Ethanol irritates the stomach mucosa. It was demonstrated that combining ethanol with agents that can cause the same side effects such as the nonsteroidal anti-inflammatory drugs, and particularly aspirin, will increase the risk of gastrointestinal ulceration and bleeding. Drugs that exert hepatotoxic effect (acetaminophen, amiodarone, methotrexate) should not be coadministrated with ethanol due to its high hepatotoxic potential (Lieber 1994). It is established that ethanol inhibits the gluconeogenesis in the liver and particularly when administrated in the fasting state. This leads to prolonged hypoglycemia and thus should not be taken together with insulin or oral antidiabetic drugs. The coadministration of ethanol with disulfiram could potentially lead to life-threatening disulfiram reaction characterized by rapid heartbeat, flushing, palpitations, and elevation of blood pressure. Disulfiram inhibits the acetaldehyde dehydrogenase, an enzyme which catabolizes ethanol in the liver. As a result an increase of acetaldehyde presents in the blood, and within 15 min of alcohol consumption, symptoms of flushing, nausea, and headache occur. These unpleasant effects of disulfiram are commonly used in the treatment of alcoholism or prevention of alcoholics to return to drinking. Other drugs when coadministrated with ethanol could lead to disulfiram-like reactions. These drugs are some antibiotics such as metronidazole and cefoperazone, the oral hypoglycemic drug chlorpropamide, and the antineoplastic agent procarbazine.

References and Further Reading

Bailey G, Arnold M, Spence D (1994) Grapefruit juice and drugs. How significant is the interaction? Clin Pharmacokinet 26:91–98

Booth L, Charnley M, Sadowski A et al (1997) Dietary vitamin K1 and stability of oral anticoagulation: proposal of a diet with constant vitamin K1 content. Thromb Haemost 77:504–509

Egashira K, Fukuda E, Onga T (2003) Pomelo-induced increase in the blood level of tacrolimus in a renal transplant patient. Transplantation 75:1057

Gardner M, Shulman I, Walker E et al (1996) The making of a user friendly MAOI diet. J Clin Psychiatry 57:99–104

Grimm M, Koziolek M, Saleh M et al (2018) Gastric emptying and small bowel water content after administration of grapefruit juice compared to water and isocaloric solutions of glucose and fructose: a four-way crossover MRI pilot study in healthy subjects. Mol Pharm 15:548–559

Kogure N, Akiyoshi T, Imaoka A (2014) Prediction of the extent and variation of grapefruit juice-drug interactions from the pharmacokinetic profile in the absence of grapefruit juice. Biopharm Drug Dispos 35:373–381

Lieber S (1994) Mechanisms of ethanol-drug-nutrition interactions. J Toxicol Clin Toxicol 32:631–681

Lilja J, Kivistö T, Neuvonen J (2000) Duration of effect of grapefruit juice on the pharmacokinetics of the CYP3A4 substrate simvastatin. Clin Pharmacol Ther 68:384–390

Malhotra S, Bailey G, Paine F et al (2001) Seville orange juice-felodipine interaction: comparison with dilute grapefruit juice and involvement of furocoumarins. Clin Pharmacol Ther 69:14–23

Neuhofel L, Wilton H, Victory M et al (2002) Lack of bioequivalence of ciprofloxacin when administered with calcium-fortified orange juice: a new twist on an old interaction. J Clin Pharmacol 42:461–466

Pronsky M (2004) Food medication interactions, 13th edn. Food Medication Interactions, Birchrunville

Rees C, Kelsey H, Richards D (1993) Acute hemodialysis induced by high dose ascorbic acid in glucose-6 phosphate dehydrogenase deficiency. BMJ 306:841

Roth A (1995) Drug metabolism. In: Smith M, Reynard M (eds) Essentials of pharmacology. WB Saunders, Philadelphia

Sica A (2006) Interaction of grapefruit juice and calcium channel blockers. Am J Hypertens 19:768–773

Tischio P (1995) In: Gennaro R (ed) The science and practice of pharmacology, 19th edn. Mack Publishing, Easton

Walter-Sack I, Klotz U (1996) Influence of diet and nutritional status on drug metabolism. Clin Pharmacokinet 31:47–64

Welage S, Carver L, Revankar S et al (1995) Alterations in gastric acidity in patients infected with human immunodeficiency virus. Clin Infect Dis 21: 1431–1438

Zastrow M (2012) Drug receptors and pharmacodynamics. In: Katzung B, Masters S, Trevor A (eds) Basic and clinical pharmacology, 13th edn. McGrawHill Medical, New York. International edition

Special Populations: Profiling the Effect of Obesity on Drug Disposition and Pharmacodynamics

39

Kenneth T. Moore

Contents

Abstract

The health implications of obesity remain a global issue with approximately 13% of the world's population categorized as obese (body mass index [BMI] $\geq$30 kg/m^2) in 2016 (World Health Organization, Obesity and overweight. http://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight. Accessed 8 Oct 2018, 2018). Obesity is typically a consequence of either excess food intake, lack of physical activity, genetic predisposition, or a combination thereof. Its impact is multifaceted, not only on the patient's health leading to a myriad of disease states directly related to obesity, but also on the management of these diseases and other common medical conditions that frequently occur. Obesity-related changes in normal physiology, such as alterations in lipid content, plasma proteins, drug metabolizing enzymes, drug transporters, and blood flow, can affect the disposition (absorption, distribution, metabolism, and excretion) and pharmacodynamics of commonly prescribed drugs, thereby altering their pharmacologic profiles. Hence, an understanding of these pharmacologic changes is necessary to ensure proper treatment is exercised. Unfortunately, our understanding of obesity-related changes in drug pharmacology in addition to the overall safety and efficacy is limited,

K. T. Moore (✉)
Cardiovascular and Metabolism Medical Affairs, Janssen Pharmaceuticals Inc., Titusville, NJ, USA
e-mail: tmoore17@its.jnj.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_7

as clinical trials rarely focus specifically on this population. Therefore, the purpose of this chapter is to provide a review of the available literature assessing the effects of obesity on the disposition and pharmacodynamics of some of the most commonly prescribed drugs. This chapter is a review only; careful clinical decision making should always be used when applying literature from the population to individual patients and scenarios.

Introduction

According to the World Health Organization (WHO), the number of individuals defined as obese has tripled over the last 40 years, and this global trend is expected to continue (World Health Organization 2018). Obesity, defined as a Body Mass Index [BMI = body weight (kg)/height $(m)^2$] $\geq$30 kg/m^2, is considered a consequence of factors, alone or in combination, that include excess food intake, lack of physical activity, and genetic predisposition. Measuring ones BMI is still the most widely used metric for overweight and obesity gradation in adults and is independent of gender (Reflection paper 2018). While obesity is defined as a BMI $\geq$30 kg/m^2, this definition can be further defined by different classes of obesity. Individuals that are categorized as class I (moderately obese) have a BMI of 30 to 34.9 kg/m^2, those within class II (severely obese) have a BMI of 35 to 39.9 kg/m^2, and those assigned to class III (very severely obese/morbid obesity) have a BMI $\geq$40 kg/m^2 (World Health Organization 2018).

Obesity is a major risk factor for several chronic diseases, which includes type 2 diabetes, cardiovascular diseases (i.e., heart disease, coronary artery disease, stroke) osteoarthritis, major depression disorder, obstructive sleep apnea, and cancer (Hanley et al. 2010; Knibbe et al. 2015). It is also considered a proinflammatory state that, among other things, can predispose individuals to arterial and venous thrombosis through prothrombotic mechanisms (Lentz 2016). These risks increase as BMI increases. Considering the prevalence of these disease states associated with obesity, several leading medical associations have taken the step to now classify obesity as an actual disease (Jensen et al. 2014).

Obesity leads to a myriad of physiological and pathophysiological changes that can affect drug disposition. Some physicians may assume that a 100 kg individual is the same as a 70 kg individual, just with an additional 30 kg of adipose; however, this is incorrect (Zuckerman et al. 2015). An increase in weight correlates with more than just an increase in body fat. Lean body mass and organ size also increase with weight and this increase tends to be disproportionate. Changes in gastrointestinal permeability and emptying, changes in cardiac, liver, and renal functions, and changes in endogenous hormones are commonly seen (Knibbe et al. 2015; Smit et al. 2018; Jusko 2017). These changes, both individually and collectively, influence the clinical pharmacology of many drugs. For example, increases in gastrointestinal permeability and gastric emptying may potentially modify a drug's absorption properties, while changes in cardiac output, hepatic, and capillary blood flow may potentially change the metabolism and clearance of a drug (Smit et al. 2018).

Unfortunately, the effects of obesity on the pharmacology of commonly prescribed drugs are lacking. The fact that many drug manufactures do not routinely assess the effects of obesity on either the pharmacokinetics (PK), pharmacodynamics (PD), safety or efficacy of their drugs can be seen as a shortcoming of the drug development process. Perhaps greater direction or influence should be given by the United States Food & Drug Administration (FDA) and the European Medicines Agency (EMA) to help resolve this lack of data. That being said, a recent positive step forward was the development of the EMA's "Reflection Paper" which focuses this topic (Reflection paper 2018).

The following review will address on some of the changes observed in the disposition and pharmacodynamics of commonly prescribed drugs in obese individuals. This review is not intended to encompass all drugs or all the possible changes in the pharmacology of these compounds but to provide a general review in order for the reader to be

better informed on the possible pharmacological changes that can occur when treating obese patients.

Obesity and Drug Disposition

Absorption

Absorption is the process in which a drug is transferred from the site of administration to where it is measured within the body (Rowland and Tozer 1995). The transport of drugs into the systemic circulation where they are typically measured can occur via molecular mechanisms that include passive diffusion, active transport, facilitated transport, endocytosis, exocytosis, ion pair transport, filtration, and bulk flow. Passive diffusion, which is the natural movement of molecules down a concentration gradient, is the most common mechanism for drug transport (Rowland and Tozer 1995). To better understand the process of drug absorption, understanding the different routes of drug administration is needed.

Drug administration can be classified as either intravascular or extravascular. Intravascular drug administration is the process of placing drug directly into the systemic blood circulation via intravenous (i.v.) or intra-arterial routes, while extravascular administration can include oral, sublingual, buccal, intramuscular, subcutaneous, transdermal, pulmonary, vaginal, and rectal routes (Rowland and Tozer 1995). This review will focus on extravascular routes of administration leading to drug absorption.

Oral administration is the most common route for drug absorption. Physiological changes due to obesity that affect this route generally include increases in gastrointestinal blood perfusion, higher cardiac output, increased splanchnic blood flow, changes in enterohepatic recirculation, accelerated gastric emptying, and increased gut permeability, all of which can alter both the rate and extent of drug absorption (Knibbe et al. 2015; Smit et al. 2018; Shank and Zimmerman 2015; Jain et al. 2011). However, there are only a few human-based clinical pharmacology studies that report a change in oral drug absorption in this population. Levothyroxine (a synthetic version of the T4 hormone) as treatment for hypothyroidism is one example (Smit et al. 2018; Cho et al. 2013). In this clinical pharmacology study, it was observed that the corrected area under the curve (AUC) and the maximum T4 concentration (C_{max}) after levothyroxine administration were lower, whereas the time to maximum concentration (T_{max}) and estimated plasma volume was higher in severely obese subjects when compared to the normal weight control group. This led to the conclusion that severely obese individuals may need higher doses than what is administered to normal-weight individuals. The authors suggest that this difference was potentially attributed to higher plasma volume and/or delayed gastrointestinal absorption of the drug in the severely obese (Michalaki et al. 2011). Midazolam, which is a benzodiazepine used as a premedicant/sedative/anesthetic agent, has also shown some PK changes in obese individuals. While drug clearance (CL) does not appear to be impacted, the oral bioavailability and volume of distribution (V_d) increases substantially (Brill et al. 2014a). The increase in bioavailability may be a result of increased splanchnic blood flow or increased paracellular absorption through the gut wall, or a combination of both (Knibbe et al. 2015). Most studies, though, have shown that the rate and extent of oral absorption does not significantly differ between obese and nonobese patients (Smit et al. 2018; Shank and Zimmerman 2015; Cho et al. 2013; Bowman et al. 1986).

While there does not appear to be a significant impact of obesity on oral drug absorption, some other extravascular routes of administration do show some changes. Since obesity is associated with a significant increase in subcutaneous fat, routes of administration that include subcutaneous, transdermal, and intramuscular administration may all be affected by changes in the quantity of fat tissue (Jain et al. 2011). These potential changes in absorption are due to the dependency that these routes have on blood flow to the skin, subcutaneous fat, and muscle (Smit et al. 2018). While cardiac output is generally increased in this population, the blood flow rate per gram of fat tissue is significantly lower when

compared to the nonobese patient (Smit et al. 2018). This is significant, considering that fat tissue typically receives approximately 5% of cardiac output, compared to 73% for viscera and 22% for lean tissues. Drug administration via these routes may ultimately be affected by the percentage of fat tissue present in obese individuals due to alterations in perfusion (Shank and Zimmerman 2015; Cheymol 2000).

One example of this pharmacological change observed in subcutaneous administration is with enoxaparin, a low-molecular-weight heparin. Subcutaneous dosing in obese individuals leads to slower absorption of the drug. This was measured by assessing the maximum PD activity postinjection. Both antifactor X_a and antifactor II_a levels took, on average, 1 h longer to reach maximum activity in obese individuals. However, ultimately the extent of absorption was complete in both the obese and nonobese patients leading the investigators to conclude dose modification in obese patients was unnecessary (Jain et al. 2011).

Human Chorionic Gonadotropin (hCG) is a hormone produced by the placenta after embryo implantation in pregnancy. Injections of hCG are commonly used during in vitro fertilization (IVF) cycles. Depending on the formulation, hCG injections can be either subcutaneous or intramuscular. When assessing the two routes of administration in obese and normal-weight individuals, C_{max}, AUC, and average hCG concentration were higher after intramuscular injection as compared with subcutaneous injection (Shah et al. 2014). When comparing these values between obese versus nonobese individuals, obese women had markedly lower C_{max}, AUC, and average hCG concentration after subcutaneous injection and similar PK values after intramuscular injection when compared with normal-weight women (Shah et al. 2014). An interesting finding in this study was that approximately 1/3 of the obese women studied had an excess of subcutaneous fat that prevented the use of a standard 1.5 in. needle for intramuscular injections (Shah et al. 2014). Therefore, the increased subcutaneous fat associated with obesity leads to inadvertent subcutaneous injection intended to be administered intramuscularly, altering the pharmacokinetics.

Sumatriptan (Imitrex) is a commonly used triptan for treatment of migraines. Following subcutaneous injection, increasing weight was associated with a decrease in sumatriptan systemic exposure. The PK values $AUC_{0–2}$ and C_{max} were approximately 1.2 times and 1.3 times higher, respectively, than those observed in individuals that weighed greater than or equal to the median weight value for the study population (Munjal et al. 2016).

Cefazolin is a cephalosporin class antibiotic used as a prophylactic for postoperative surgical site infection. Using microdialysis, it was found that unbound cefazolin concentrations were lower in morbidly obese compared with nonobese patients following subcutaneous administration. Cefazolin tissue distribution reduces with increasing body weight and this is believed to be a consequence of altered blood perfusion in obese individuals (Brill et al. 2014b). While these are four examples where obesity may affect the pharmacology of a compound, obesity does not appear to affect either the PK or PD of other agents administered subcutaneously like the insulin Lispro (de la Peña et al. 2015), recombinant follicle-stimulating hormone (rFSH) (Steinkampf et al. 2003), or moxifloxacin (Kees et al. 2011).

Considering that pulmonary function is uniformly altered in obese individuals, primarily due to reduced lung volumes, there is an unfortunate scarcity of published data on the topic (Cheymol 2000). A study conducted in overweight patients with persistent asthma administered inhaled corticosteroids has shown attenuated symptom and fractional exhaled nitric oxide (FeNO) dose responses when compared to normal weight individuals. There were no differences in forced expiratory volume in 1 s (FEV1) or methacholine PC20 between the groups. The investigators hypothesize that attenuated cortisol suppression in the overweight group, secondary to reduced peripheral lung absorption, may be the cause (Anderson and Lipworth 2012).

In summary there is a lack of data concerning drug absorption in obese patients. While the few available published examples seem to indicate orally administered drugs are not likely to show a change in absorption, other routes of

extravascular administration like transdermal, subcutaneous, and intramuscular routes may be affected as weight and fat content increases. Other extravascular routes including sublingual, buccal, rectal, and vaginal routes in obese individuals have not been widely studied and are not mentioned in this review.

Distribution

Distribution is the process of reversible drug transfer to and from the blood and other tissues in the body (Hanley et al. 2010; Rowland and Tozer 1995). Organ size and the extent to which they are perfused with blood, drug binding in both blood and tissues, and its permeability all influence a drug's distribution properties (Rowland and Tozer 1995; Shank and Zimmerman 2015). The rate of drug distribution is determined by the blood flow to the tissues and the ability of the drug to cross the capillary wall and enter the cells of the tissue.

Drug distribution can be quantified by determining the volume of distribution (V_d). This represents the degree to which a drug is distributed into the tissues; therefore, the higher the V_d, the greater a drug is distributed. To properly measure this value, the concentration of the drug needs to be determined. Measurement of this distribution in humans is generally limited to those tissues easily accessible, for example, blood or plasma. Note, once a drug is absorbed, it is delivered simultaneously to all tissues. To calculate V_d, an equilibrium needs to occur in the distribution of the drug between the tissues and plasma (Rowland and Tozer 1995). Once this occurs, the following equation can be used:

$$V_d = A/C$$

where (V_d) is the volume of distribution, (A) represents the total amount of drug in the body, and (C) represents the concentration of the drug in the plasma.

It is important to note that V_d is not a physiologic value but a mathematical concept and it rarely corresponds to a real volume. It is more of a reflection of how a drug will distribute throughout the body, depending on its physiochemical properties (e.g., polarity, molecular size, and degree of ionization) and plasma protein binding (Shank and Zimmerman 2015). An easy way of determining the V_d is with an i.v. bolus of the drug in question. That way the amount of drug in the body is known immediately following the completion of the bolus.

$$V_d = D_{i.v.}/C_0$$

where (V_d) is the volume of distribution, ($D_{i.v.}$) represents the amount of drug administered via i.v., and (C_0) represents the concentration of the drug in the plasma at time 0.

Different disease states may influence the V_d of a drug and obesity is one such example. These changes can be a result from both physiological changes observed with increased body weight (i.e., increases in adipose tissue, changes in protein binding, and reduced tissue perfusion) and the intrinsic physiochemical properties of the drug, described previously (Knibbe et al. 2015).

When considering the physiological changes, alterations in either the concentrations of plasma binding proteins or in their binding affinity could ultimately affect the movement of drugs into the tissues and affect its distribution. Obesity has been associated with changes in both characteristics. Research has shown that serum albumin tends to be unaltered in obesity (Jain et al. 2011). Therefore, those compounds that primarily bind to albumin (most acidic drugs) do not appear to be affected and thus the V_d remains unaltered. For example, both thiopental (used for the induction of general anesthesia) and phenytoin (an antiepileptic) bind primarily to albumin and are generally unaltered by increases in weight (Jain et al. 2011). However, those compounds binding to α_1-acid glycoprotein (most basic drugs) appear at times, to be affected. That is, the degree of protein-binding increases, reducing the free plasma concentrations and decreasing their V_d (Leykin et al. 2011). Yet, this affect when assessed across basic drugs is inconsistent (Jain et al. 2011).

Physiological changes to the microvascular system due to obesity may ultimately impact

tissue blood flow and perfusion, therefore altering drug distribution (Levy et al. 2008). This impairment may be due to the increased levels of oxidative stress and/or inflammatory cytokines generally associated with this disease state (Levy et al. 2008). Proper tissue penetration is particularly important for antibiotics used for localized infections or perioperative prophylaxis, where certain tissue concentrations need to be achieved (Smit et al. 2018). Microdialysis studies with cefuroxime (a cephalosporin antibiotic) and ciprofloxacin (a fluoroquinolone antimicrobial) have shown that tissue penetration in obese patients was significantly reduced (Smit et al. 2018).

Lipophilicity (a molecule's affinity to bind to adipose tissue) is one example of a physiochemical property that may influence the distribution of a drug in these patients, as there is a marked increase in the ratio of adipose tissue to lean body mass (Shank and Zimmerman 2015). Therefore, intuitively, highly lipophilic molecules may accumulate more in fat stores and display a larger V_d in these individuals. For example, Diazepam (Valium), a benzodiazepine derivative commonly used for anxiety, is a highly lipophilic drug and shows a dramatic increase in V_d when studied in obese patients (Knibbe et al. 2015). However, the literature is not always consistent in this regard, and it should be noted that the size of V_d does not always correlate with the degree of lipophilicity, nor can it be predicted based on lipophilicity alone (Knibbe et al. 2015; Cheymol 2000). Propofol, an i.v. sedative-hypnotic commonly used for initiation and maintenance of monitored anesthesia care, is a classic example of this dichotomy as it is also highly lipophilic but does not show an increase in V_d when dosed in obese patients. Another physiochemical property that may affect drug distribution is the polarity of the molecule. Polar molecules have more difficulty passing into cells and are not likely to deposit into fat (Shank and Zimmerman 2015).

In summary, changes in drug distribution (V_d) are highly influenced by the intrinsic physiochemical properties of a drug and the physiological changes typically associated with obesity. Understanding these changes is critical in drug therapy, particularly when determining the proper loading dose for antimicrobial agents or treatment with sedatives. Unfortunately, our current understanding of these changes is limited.

Metabolism

Metabolism is a biochemical mechanism in which drugs are altered and eliminated from the body (Rowland and Tozer 1995). These mechanisms (or pathways) consist of chemical reactions which include oxidation, reduction, hydrolysis, and conjugation (Rowland and Tozer 1995). Most drug metabolism occurs in the liver; however, the kidneys, skin, lungs, blood, and gastrointestinal wall may also play a role (Rowland and Tozer 1995). These biotransformation pathways convert drugs into products termed metabolites, which are generally polar molecules that can be easily excreted by the body. Oxidation, reduction, and hydrolysis pathways are typically referred to as Phase I pathways. Oxidation-based reactions primarily rely on the superfamily of enzymes called cytochrome (CYP) P450 monooxygenases (most common enzymes include: CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4) (Morrish et al. 2011; Atkinson et al. 2007). Reduction-based reactions rely on reductase enzymes found in intestinal anaerobic bacteria and hydrolysis-based reactions rely on esterases, amidases, and proteases (Atkinson et al. 2007). Once completing this type of metabolism, the original drug and its metabolites can become either completely pharmacologically inactive; the original drug becomes inactive, but one or more of the metabolites are active (however to a lesser degree) or the original drug, which was inactive to start, produces one or more active metabolites (Atkinson et al. 2007). For the last scenario, the original inactive drug substance is called a prodrug.

Phase II metabolism refers to chemical reactions through the covalent bonding of an endogenous molecule to the drug (Atkinson et al. 2007). This type of reaction is termed conjugation and can come in the form of glucuronidation, sulfation, acetylation, methylation, and amino acid conjugation (Atkinson et al. 2007). These

reactions bond a hydrophilic molecule (e.g., glucuronic acid, sulfate, or glycine) to either the parent drug or the metabolite of the parent drug to form a water-soluble compound that can be excreted. The metabolites (or conjugates in this case) produced from this reaction are unlikely to be pharmacologically active. Glucuronidation is the most common type of phase II reactions. The products, glucuronides, are formed by uridine diphosphate (UDP)-glucuronosyltransferases (UGTs) enzymes located in the liver, kidneys, and brain (Atkinson et al. 2007). While some drugs undergo either phase of metabolism, most undergo phase I followed by phase II metabolism.

While these reactions can occur in a few different tissues, the liver is the main organ of metabolism. This enzyme activity determines what is called the intrinsic liver clearance (CL_{int}) of the drug. This intrinsic clearance of a drug is a theoretical value of CL in the absence of any dependence on blood flow or protein binding. CL_{int} together with hepatic blood flow (Q_H) and fraction of unbound drug (f_u) determines the hepatic clearance (CL_H) of the drug (Smit et al. 2018). The following equation can be applied (well-stirred model):

$$CL_H = Q_H \cdot (f_u \cdot CL_{int})/(Q_H + f_u \cdot CL_{int})$$

Changes in these parameters in addition to the hepatic extraction ratio (E_H) (i.e., the efficiency of the liver to clear the drug from the blood) can therefore alter the hepatic clearance (Smit et al. 2018).

$$CL_H = Q_H \cdot E_H$$

A drug's extraction ratio can be either high or low (Range from 0 to 1). High extraction drugs ($E_H > 0.7$) are primarily dependent on Q_H and less on enzyme capacity or f_u as it is nonrestrictively cleared (High CL_{int}). Low extraction drugs ($E_H < 0.3$) are primarily dependent on hepatic enzyme capacity, therefore restrictively cleared (Low CL_{int}) (Smit et al. 2018).

Since $CL_H = Q_H \cdot E_H$, the extraction ratio can be determined with the following equation:

$$Q_H \cdot (f_u \cdot CL_{int})/(Q_H + f_u \cdot CL_{int}) = Q_H \cdot E_H$$

$$E_H = (f_u \cdot CL_{int})/(Q_H + f_u \cdot CL_{int})$$

In obese individuals, liver pathologies are common. Fat deposition and inflammation in the liver can result in hepatic steatosis (fatty infiltration), changes in Phase I or II enzyme systems, and hepatic blood flow (Jain et al. 2011). It is hypothesized that chronic low-grade inflammation in the liver can result in decreased enzyme expression of certain CYP enzymes (Smit et al. 2018). For example, obesity has been found to correlate to a decrease in CYP3A4-mediated metabolism for low E_H drugs (Jusko 2017). A good example of this can be seen with the corticosteroid methylprednisolone, in which the absolute clearance was decreased by approximately 40% in the obese individuals (Jusko 2017; Dunn et al. 1991). Another example of decreased CYP3A4 metabolism in obese individuals can be seen with the use of triazolam, a benzodiazepine class drug. In a small clinical pharmacology study with 12 obese patients and 12 pair-matched normal weight subjects, the administration of triazolam was found to result in a much lower CL in the obese subjects (340 mL/min) compared to the matched controls (531 mL/min) (Abernethy et al. 1984). A final example of decreased CL in this phase I CYP3A4 mediated pathway is alfentanil, an opioid antagonist. The CL of alfentanil was shown to decrease approximately 46% from 321 mL/min in nonobese subjects to 179 mL/min in obese subjects (Brill et al. 2012).

Alternatively, the activity of CYP2E1 enzymes has been observed to increase in obese individuals thereby increasing drug clearance. This can be seen with the use of acetaminophen (Fig. 1) in morbidly obese patients. When administered in this population, the median AUC of acetaminophen was significantly smaller, while the AUCs of the glucuronide, sulfate, and cysteine metabolites were significantly higher (van Rongen et al. 2016).

Additionally, the muscle relaxant chlorzoxazone was found to have an unbound CL values that was ~3-times higher (CL_{obese} = 27.5 L/min vs.

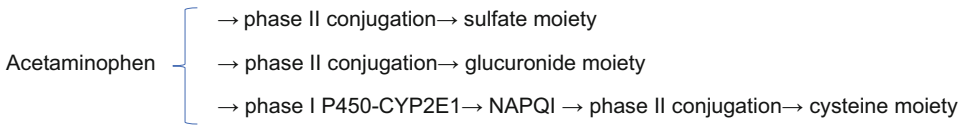


Fig. 1 Phase I and II Metabolism of Acetaminophen

$CL_{non\text{-}obese}$ = 9.9 L/min) in morbidly obese subjects indicating a significant induction of the CYP2E1 enzyme system (Brill et al. 2012). With the exception of the changes reported for the CYP3A and CYP2E1 enzymes, and a possible trend toward higher clearance of CYP1A2, CYP2C9, CYP2C19, and CYP2D6 substrates, phase I pathway enzymes do not appear to be highly impacted by obesity (Knibbe et al. 2015) (Fig. 1).

The phase II pathway, specifically conjugation reactions, seems to be consistently elevated in morbidly obese individuals, leading to an increase in drug CL (Knibbe et al. 2015). This was observed with acetaminophen, oxazepam, and lorazepam, the last two being benzodiazepines, and notably, all three are examples of low-to-medium extraction ratio drugs (Smit et al. 2018; Jain et al. 2011). Mean acetaminophen CL values were CL_{obese} = 484 mL/min and $CL_{non\text{-}obese}$ = 323 mL/min; mean oxazepam CL values were CL_{obese} = 156.8 mL/min and $CL_{non\text{-}obese}$ = 50.4 mL/min; and mean lorazepam CL values were CL_{obese} = 102 mL/min and $CL_{non\text{-}obese}$ 62.9 mL/min for obese and nonobese subjects, respectively (Brill et al. 2012).

Interestingly, recent studies with morphine, also mainly glucuronidated, show somewhat conflicting results. While these studies have demonstrated higher morphine glucuronide concentrations in obese and nonalcoholic steatohepatitis (NASH), patients compared to nonobese individuals, a couple of these studies have also shown similar morphine concentrations together with increased glucuronide concentrations (Smit et al. 2018). This then suggests that there was not a significant increase in glucuronidation, but likely a decrease in glucuronide clearance (Smit et al. 2018). A possible explanation for this difference may be attributed to the fact that morphine is actually a medium-to-high extraction ratio drug. Assuming hepatic blood flow is unchanged in this population, that may explain the difference observed between drugs that have high versus low extraction ratios and are primarily glucuronidated (Smit et al. 2018). As previously mentioned drugs with high extraction ratios depend almost entirely on hepatic blood flow, rather than intrinsic metabolic clearance.

Whether hepatic blood flow changes in obese individuals remains unclear. Some have reported that obesity and NASH lead to an increase in fat accumulation in the liver, in turn causing sinusoidal narrowing and ultimately a reduction in hepatic blood flow (Knibbe et al. 2015; Cho et al. 2013; Farrell et al. 2008). While others argue that hepatic blood flow is not necessarily reduced in these patients because of the increased blood volume and cardiac output observed with an increase in body mass (Cho et al. 2013; Casati and Putzu 2005).

Interestingly, when assessing propofol, which is extensively metabolized by phase II UGT enzymes and considered a high hepatic extraction ratio drug (having a CL that is limited only by Q_H), drug CL was significantly increased (~10.0 L/min) in morbidly obese subjects compared to nonobese subjects (4.1 L/min) (Wu 2016; van Kralingen et al. 2011). The authors suggest that this increase in CL is likely a result of increase in Q_H secondary to the typical increase in cardiac output generally observed in obese patients. While the literature remains sparse, there have been other reports of higher CL values observed for a small number of high-extraction-ratio drugs, again suggesting an increase in Q_H in obese patients (Knibbe et al. 2015).

In summary, changes in drug clearance via phase I metabolism are dependent on the specific enzymatic pathway. In the current published literature, CYP3A4 mediated clearance was found to be consistently lower in obese subjects, while CYP2E1 mediated clearance tends to show higher

activity in this population. Other pathways, including CYP2C9, CYP2C19, and CYP2D6, show a trend for higher clearance in obese subjects; however, in most cases these changes were not statistically significant. In contrast, the phase II pathways, specifically conjugation reactions, appear to be elevated in morbidly obese individuals and hence an observed increase in CL. Lastly, there appears to be a trend for higher CL values for drugs considered to have high hepatic extraction ratios. While the exact reason for these higher CL values may not yet known, many have speculated that higher Q_H with increasing body mass may explain these observations.

Excretion

Excretion is the process of eliminating a drug from the body. Drugs can be excreted unchanged as the parent moiety or following metabolism as active and/or inactive metabolites. There are several physiological routes in which a drug can be eliminated from the body; however, most drugs are eliminated via the kidneys or liver. In some cases, drug excretion can occur through the lungs, breast milk, sweat, tears, skin, hair, or saliva, though these are considered secondary pathways.

Renal excretion plays a critical role in elimination of both unchanged drugs and metabolites from the body. Although the ability to eliminate these compounds by renal processes would require them to be polarized and water-soluble first. Therefore, normally highly lipid soluble drugs generally require them to first go through hepatic metabolism (phase I/II) to increase their water solubility before they can be excreted. It is important to note that drugs entering the hepatic circulation may also be excreted into the bile. For these compounds, the drug re-enters the intestinal tract and can be eliminated via the feces or reabsorbed back into the systemic circulation, which is termed enterohepatic recycling.

For the purposes of this review, we will focus on renal elimination. The basic functional unit of the kidney is the nephron. Collectively these units are responsible for the removal of metabolic waste and the maintenance of water and electrolyte balance. The nephron can be broken down into five basic anatomic units: the glomerulus, proximal tubule, loop of Henle, distal tubule, and collecting tubule (Rowland and Tozer 1995). Blood is first filtered through the glomerulus with a rate of approximately 120 mL/min. The resulting filtrate is passed down the proximal tubule then through the loop of Henle, the distal tubule and finally whatever filtrate is left is passed into the collection tubules and eliminated from the body (Rowland and Tozer 1995). There are four processes involved in drug elimination via these nephrons, they include glomerular filtration, tubular secretion, tubular reabsorption, and renal metabolism with glomerular filtration playing a primary role (Morrish et al. 2011; Brill et al. 2012; Fan and de Lannoy 2014).

Clearance by glomerular filtration (CL_{FILT}) is a passive process that relies on normal physiologic hydrostatic pressure in the glomerulus and can only filter unbound drug into the urine. It can be represented as the following equation:

$$CL_{FILT} = f_u \cdot GFR$$

where f_u is the fraction of unbound drug and GFR is the glomerular filtration rate (GFR is approximately 125 mL/min in an average adult)

Tubular secretion of drugs is an active transport process that occurs in the proximal tubule of the nephron. Since this is a carrier mediated system, its capacity is limited and can become saturated. Clearance by tubular secretion (CL_{SEC}) relies on renal blood flow (Q_R), the unbound fraction (f_u), and the intrinsic clearance (CL_{INT}) of the drug and is similar to hepatic clearance. It is represented by the following equation: (Tucker 1981)

$$CL_{SEC} = (Q_R \cdot f_u \cdot CL_{INT})/(Q_R + f_u \cdot CL_{INT})$$

Tubular reabsorption (F_{REAB}) of drugs occurs in the distal tubule after the drug is filtered through the glomerulus and passes the proximal tubule and loop of Henle. Reabsorption can be either a passive or active process and is dependent on the f_u, urine flow and pH.

Renal clearance (CL_R) is therefore the sum of these four processes and can be represented by the equation below: (Fan and de Lannoy 2014)

$$CL_R = (f_u \cdot GFR + CL_{SEC}) \cdot (1 - F_{REAB}) + CL_{MET}$$

where GFR is the glomerular filtration rate, f_u is the fraction unbound in plasma, CL_{SEC} is clearance by tubular secretion, F_{REAB} is the fraction reabsorbed, and CL_{MET} is clearance by renal metabolism.

The influence of obesity on renal function and drug excretion is not straight forward nor clear. There is some evidence that obesity is related to a state of glomerular hyperfiltration and estimated GFRs have been seen to increase up to 62% (Cho et al. 2013). This type of increase mimics those observed in early stage diabetic nephropathy or end-stage renal disease, both of which are commonly observed in obese individuals (Cho et al. 2013; Brill et al. 2012). However, less is known about the effect of obesity on renal tubular secretion and tubular reabsorption. Ultimately an increase in drug CL has been observed in obese individuals who have normal kidney function. This increase is likely secondary to increased kidney size and renal blood flow generally associated with obesity (Knibbe et al. 2015; Atkinson et al. 2007; Brill et al. 2012).

With drugs that are primarily excreted through the kidneys via filtration, for example, vancomycin and low molecular-weight heparins (LMWH), they generally display an overall increase in CL with an increase in weight. This can be seen with the use of vancomycin, having a greater than 50% increase in CL in obese versus nonobese patients (CL_{obese} = 197 mL/min vs. $CL_{non\text{-}obese}$ = 77 mL/min) (Brill et al. 2012). A similar trend is observed with Dalteparin (CL_{obese} = 1.30 L/h vs. $CL_{non\text{-}obese}$ = 1.11 L/h) (Cho et al. 2013; Brill et al. 2012). In both cases, the increase in CL is likely related to an increase in GFR.

There are several drugs that are at least partly eliminated by tubular secretion, procainamide, ciprofloxacin, and digoxin to name a few (Blouin and Warren 1999). In general, a trend towards higher CL has been observed with these compounds. Procainamide, an antiarrhythmic drug, is eliminated unchanged through a combination of glomerular filtration and tubular secretion. The approximate 50% increase in CL observed (CL_{obese} = 4.19 mL/min vs. $CL_{non\text{-}obese}$ = 2.68 mL/min) is likely a result of increased tubular secretion as there was no difference in observed in creatinine clearance (Brill et al. 2012). Studies with oseltamivir, an antiviral medication used to treat and prevent influenza A and B, that undergoes both renal tubular secretion and filtration, have shown a consistent increase in renal excretion, suggesting renal secretion may be augmented in these patients (Smit et al. 2018).

In a similar fashion, the renal CL of lithium, a compound used for bi-polar and major depression disorders, primarily involves glomerular filtration and tubular reabsorption. When dosed in obese individuals, renal CL increased without a change in glomerular filtration, supporting the premise that tubular reabsorption was decreased in obese individuals (Blouin and Warren 1999).

In summary, obesity has a significant impact on kidney function and renally eliminated compounds. Changes include compensatory hyperfiltration to meet increased metabolic demands, alterations in tubular secretion and reabsorption, and potential changes in some of the metabolic pathways of the kidney. In general, studies have shown that clearance of renally eliminated drugs is higher in obese patients. This increase in CL is likely due to increased glomerular filtration and tubular secretion. There is less published evidence on the changes to tubular reabsorption and renal metabolism with obesity and hence, just briefly mentioned in this review. It should be noted that though renal clearance is initially enhanced by compensatory hyperfiltration and hyperperfusion in obese individuals, this effect is believed to eventually decrease over-time or with a continued increase in weight, as a result of persistently elevated intraglomerular pressure.

Obesity and Pharmacodynamic Changes

Obesity and morbid obesity is more than just excess weight gain; it is associated with both an altered anatomy and physiological state. Many of

the pathophysiological manifestations of obesity can be characterized by a low-level chronic inflammation leading to dysregulation of metabolic homeostasis, dyslipidemia, altered blood pressure, diabetes, cardiovascular disease, chronic kidney disease, cancer, and thrombosis (Lentz 2016). This low level of chronic inflammation is reflective of the over expression of cytokines like TNF-α, interleukin-6, and interleukin-1b secreted by adipose tissue, leading to an increase in the concentration of macrophages (Lentz 2016; Gandhi et al. 2012). Additionally, nutritional and genetic changes related to obesity can affect receptor expression and/or receptor affinity to many drugs (Shank and Zimmerman 2015; Jain et al. 2011). For example, cytokine tumor necrosis factor alpha is produced in excessive amounts in obese patients; this in turn perpetuates insulin resistance (Shank and Zimmerman 2015). This is compounded by the fact that adipose tissue also has greater intrinsic insulin cleaving activity, hence the need for more insulin to produce the same pharmacodynamic (PD) response (Shank and Zimmerman 2015).

Other receptor changes observed in obese patients include a decreased sensitivity to acetylcholine and increased sensitivity to benzodiazapines (e.g., triazolam), thereby increasing the psychomotor response (Shank and Zimmerman 2015; Jain et al. 2011; Blouin and Warren 1999). Polymorphisms in the μ-opioid receptor gene, the P-glycoprotein gene (ABCB1), and the catechol-O-methyltransferase gene (COMT) seem to be responsible for the variability observed in morphine PD and PK, as previously discussed (Jain et al. 2011). Additionally, adipocytes secrete leptin, which reduces macrophage and T-cell differentiation and activity, in turn affecting the immune system (Smit et al. 2018). In fact, when obese patient contract infectious diseases, they are generally associated with worse outcomes compared to the normal weight patients (Smit et al. 2018).

Whole organ systems are also affected. The cardiovascular system is significantly impacted in obese individuals. Increasing weight leads to an increase in blood volume and a subsequent increase in cardiac output, and an increase in baseline heart rate (Zuckerman et al. 2015). Some of the anatomical changes observed in obese individuals include the development of cardiomyopathies (left ventricular hypertrophy), atherosclerosis, and endothelial dysfunction (decrease in nitric oxide production) leading to hypertension (Zuckerman et al. 2015). While obesity causes an initial hyperdynamic change leading to increased cardiac output and blood volume, the long-term effects of obesity lead to increases in vascular resistance and to structural changes in the heart and ultimately heart failure (Martin et al. 2012). While serum albumin and total protein concentrations do not appear to be altered, concentrations of alpha-1-acid glycoprotein are generally increased (Hanley et al. 2010; Knibbe et al. 2015; Smit et al. 2018).

Obesity impacts respiratory function as well. Changes in large airway anatomy lead to changes in lung volumes, compliance, reserve capacity, airway resistance, ventilatory drive, and the work of breathing (Zuckerman et al. 2015; Smit et al. 2018). Obese patients are also at increased risk for obstructive sleep apnea and asthma and may also suffer from obesity-hypoventilation syndrome (Zuckerman et al. 2015; Smit et al. 2018).

Physiologic changes to the gastrointestinal track typically include accelerated gastric emptying and increases in gut permeability, blood perfusion, and splanchnic blood flow (Zuckerman et al. 2015; Smit et al. 2018). Intestinal motility is increased in the upper intestinal tract and decreased in the lower intestinal tract and colon (Zuckerman et al. 2015). Changes in enterohepatic recirculation commonly occur along with changes in the expression of metabolic enzymes and transporters (Zuckerman et al. 2015). Other changes include an increased risk for a hiatal hernia and increased intra-abdominal pressure which increases the risk for aspiration (Zuckerman et al. 2015).

Excessive accumulation of fat in the liver can cause functional changes in morphology (Knibbe et al. 2015). Nonalcoholic fatty liver disease (NAFLD) resulting in steatosis or steatohepatitis (NASH) commonly occurs (Knibbe et al. 2015). Together with the increased rate of sinusoidal narrowing, it may cause liver blood flow to decline over time (Smit et al. 2018). Liver volume in obese and morbidly obese individuals is

generally increased and alterations in some enzyme systems (e.g., CPY3A and CYP2E) have been observed (Knibbe et al. 2015).

Changes in renal function have also been observed in obese and morbidly obese individuals. These changes can include alterations in glomerular filtration, tubular secretion, and tubular reabsorption. Similar to the trend in hepatic changes observed over time, it is generally believed that renal clearance, though initially enhanced via hyperfiltration and hyperperfusion, eventually declines as a result of constantly elevated intra-glomerular pressure (Smit et al. 2018).

Obesity is also associated with mood disorders, specifically depression, and changes in dopamine-mediated reward mechanisms (which regulate both substance abuse behaviors and binge eating behaviors) (Zuckerman et al. 2015).

The following are some selected examples of compounds from different classes of therapeutics where the effects of obesity on the PK and/or PD were assessed.

Obesity and Pharmacokinetic and Pharmacodynamic Changes

Cardiovascular System

Antihypertensive Drugs

Obesity is an underlying cause of hypertension, with approximately 75% of obese individuals ultimately diagnosed with the condition (Landsberg et al. 2013). Together they are recognized as a significant risk factor for cardiovascular disease. The relationship between increasing body weight and blood pressure (BP) was demonstrated in the 1960s with the Framingham Heart Study (Landsberg et al. 2013). While the exact cause is not known, insulin-mediated sympathetic nervous system (SNS) stimulation appears to be a significant factor. Hypertension in lean patients appear to be a result of peripheral vasoconstriction, while obesity-related hypertension seems to depend on SNS hyperactivation, with downstream effects that include increases in cardiac output and renin and aldosterone release (Cataldi et al. 2016). The mechanism behind this SNS stimulation appears to be related to the release of adipokines, inflammatory cytokines, and free fatty acids from adipose tissue. The release of these substances then affects insulin sensitivity and in turn SNS stimulation (Cataldi et al. 2016). This mechanism is supported by clinical studies that demonstrated the concomitant decrease in BP and SNS activity when insulin was lowered by low energy diets in obese patients (Landsberg et al. 2013).

As far as treatment of hypertension is concerned, some of the more popular antihypertensive medications include the β-adrenergic receptor antagonists due to their efficacy in the treatment of hypertension, ischemic heart disease, congestive heart failure, and certain arrhythmias. Propranolol is an example of a nonselective β-blocker, as it has equal infinity for β_1 and β_2 adrenergic receptors. Other β-blockers like metoprolol, atenolol, acebutolol, nebivolol, bisprolol, metoprolol, and esmolol have a greater affinity for β_1 receptors than β_2 (Brunton et al. 2006). While several studies have been conducted, assessing the PK and PD in obese patients taking propranolol, results have not been consistent across these studies. A study conducted by *Cheymol* et al. that assessed a 8 mg i.v. dose of propranolol in obese and nonobese subjects reported a significant increases in AUC ($AUC_{obese} = 161.0\ h \cdot \mu g/L$ vs. $AUC_{non\text{-}obese} = 109.6\ h \cdot \mu g/L$) and lower V_d ($V_{d\text{-}obese} = 208.9$ L vs. $V_{d\text{-}non\text{-}obese} = 318.6$ L) and CL ($CL_{obese} = 57.5$ L/h vs. $CL_{non\text{-}obese} = 75.9$ L/h) values for obese subjects compared to nonobese subjects (Cheymol et al. 1987).

A slightly older study conducted by *Bowman* et al. assessed both an i.v. dose (10 mg) and oral dose (40 mg) of propranolol in obese and nonobese subjects. Following i.v. administration propranolol CL was unchanged, however the V_d was greater ($V_{d\text{-}obese} = 3391$ vs. $V_{d\text{-}non\text{-}obese} = 1981$ L). Similarly, following oral administration, there was no difference observed in CL, while V_d was higher in the obese subjects and there was a trend towards higher systemic availability in the obese group (35% vs. 27%) however this was not statistically significant (Bowman et al. 1986). Both propranolol (80 mg) and atenolol (100 mg), β_1 selective blockers, were assessed in lean normolipemic, obese normolipemic, and

obese hyperlipidemic patients. After completion of this crossover study, there was no statistically significant difference in propranolol serum concentrations across the three groups, while the concentrations of atenolol were significantly lower in both normolipemic obese and hyperlipidemic obese subjects. Propranolol displayed a trend towards increases in both V_d and CL in the obese patients with hyperlipidemia. In comparison, Atenolol displayed a significantly lower systemic exposure (AUC), C_{max} and CL in both normolipemic and hyperlipidemic obese patients. Changes in the PD effects (heart rate and systolic blood pressure) for both compounds were similar across patient groups (Wójcicki et al. 2003). Studies with other β_1 selective blockers (nebivolol and metoprolol) have shown higher V_d and CL values in obese subjects when compared to nonobese subjects however with similar reductions in HR and BP (Galletti et al. 1989; Cheymol et al. 1997).

While there was some diversity in changes observed in the PK of these β-blockers, the PD effects were generally the same across the different weight groups assessed. That being said, there is a lack of data on the simultaneous administration of multiple antihypertensive drugs when treating patients that display poor responsiveness to single drug therapy. Important to note is that β-blockers have been linked to insulin resistance and associated weight gain as well as decreased diet-induced thermogenesis and fat oxidation. Therefore, many have suggested that this class of antihypertensives should be limited to obese patients with specific cardiovascular indications, like postmyocardial infarction and heart failure (Landsberg et al. 2013). Based on this rationale, many physicians have suggested the use of ACE inhibitors that antagonize the renin-angiotensin-aldosterone system (RAAS) as first-line agents (Landsberg et al. 2013).

Antiarrhythmic Drugs

Antiarrhythmic drug therapy can have two goals: (1) the termination of an ongoing arrhythmia or (2) the prevention of an arrhythmia. Since antiarrhythmic drugs can have various mechanisms of action, the Vaughan-Williams classification system was developed to better organize these compounds. Most antiarrhythmic drugs are grouped into four main classes based on their dominant cellular electrophysiologic effect. Class I drugs are sodium channel blockers, which can be broken down into three subclasses (a, b, c,) based on their kinetic effects of the sodium channel; Class II drugs are beta-blockers; Class III drugs are primarily potassium channel blockers; and Class IV drugs are the non-dihydropyridine calcium channel blockers (Brunton et al. 2006).

Amiodarone is a Class III antiarrhythmic but noted to have properties of the other classes (Shank and Zimmerman 2015). It is a preferred antiarrhythmic for patients with structural heart disease. Following a literature search, only one study was identified that assessed the effects of weight on the pharmacology of this compound. A study conducted by *Fukuchi* et al. evaluated the influence of obesity on pharmacology of amiodarone using PK Modelling in Japanese patients treated with oral therapy. The model indicated that the total clearance of amiodarone was influenced by BMI and age, specifically the clearance decreased by approximately 22.3% with a BMI >25 kg/m (Fukuchi et al. 2009).

Procainamide is a Class Ia antiarrhythmic. The PK of i.v. procainamide was studied in a small clinical pharmacology study in obese and normal-weight subjects. The V_d was similar between both groups, while CL, corrected for body surface area, was greater in obese subjects. While metabolic clearance was similar between the groups, renal clearance was found to be significantly increased, which is likely due to increased tubular secretion of PA in the obese group (Christoff et al. 1983).

Verapamil is a Class IV antiarrhythmic. Verapamil (0.15 mg/kg) was administered by 10-min i.v. infusion to 12 obese (127 +/− 8 kg) and 11 normal weight (74 +/− 4 kg) hypertensive patients. Pharmacodynamic measures which included electrocardiographic P-R interval, mean BP, and HR were recorded with simultaneous PK blood sampling. Elimination half-life ($t_{1/2}$) was prolonged in obese patients ($t_{1/2obese} = 10.1$ h vs. $t_{1/2non\text{-}obese} = 3.6$ h) and a marked increase in V_d ($V_{d\text{-}obese} = 713$ L vs. $V_{d\text{-}non\text{-}obese} = 301$ L) was observed with no change in total CL

(CL_{obese} = 1339 mL/min vs. $CL_{non-obese}$ = 1250 mL/min). Additionally, verapamil plasma protein binding was similar between groups (percent unbound, 4.8% obese vs. 5.1% nonobese). Through PD modeling, the E_{max} (maximal prolongation in the P-R interval) was unchanged with obesity ($E_{max-obese}$ = 53.7 ms vs. $E_{max-non-obese}$ = 45.9 ms). However, the EC_{50} (verapamil concentration required to achieve 50% of E_{max} prolongation in the P-R interval) was greater in obese patients ($EC_{50-obese}$ = 45.9 ng/mL vs. $EC_{50\ -\ non-obese}$ = 22.6 ng/mL) (Abernethy and Schwartz 1988).

Digoxin is a cardiac glycoside, typically used for the treatment of heart failure and arrhythmias. Digoxin PK was studied in obese (mean weight = 100.2 kg) and nonobese (mean weight = 64.6 kg) subjects. After administration of 0.75 mg digoxin via i.v., serial plasma samples were obtained. Elimination $t_{1/2}$ was similar between groups ($t_{1/2-obese}$ = 36 h vs. $t_{1/2-non-obese}$ = 41 h), along with the V_d ($V_{d-obese}$ = 981 L vs. $V_{d-non-obese}$ = 937 L), and total CL (CL_{obese} = 328 mL/min vs. $CL_{non-obese}$ = 278 mL/min) of digoxin (Abernethy et al. 1981).

Anticoagulants/Antiplatelets

Obesity is a risk factor for both arterial and venous thrombosis, along with ischemic stroke, deep vein thrombosis (DVT), and pulmonary embolism (PE) (Shank and Zimmerman 2015; Lentz 2016). The obesity-related mechanisms that appear to be associated with this risk of thrombosis are chronic inflammation, impaired fibrinolysis, immobility, obstructive sleep apnea, heart failure, and venous stasis (Lentz 2016). Additionally, obese patients appear to have a greater risk of recurrent VTE compared with normal-weight patients (Alquwaizani et al. 2013).

Unfractionated heparin (UFH) is one of several options for VTE prophylaxis in medical patients with known risk factors or for both general and bariatric surgery (Shank and Zimmerman 2015). Heparin reversibly binds to antithrombin, which in turn inhibits activated coagulation factors in both the intrinsic and common coagulation pathways, including thrombin (factor IIa), factor Xa, and factor IXa (Brunton et al. 2006). Both i.v. and subcutaneous (s.c.) injections are routes for UFH administration. When administered subcutaneously, doses need to be large enough (>30,000 U/day) to overcome UFH's low bioavailability; therefore, i.v. administration is generally preferred, as therapeutic plasma concentrations can be quickly achieved and effectively monitored (Alquwaizani et al. 2013). Several clinical trials have been published suggesting that modified dosing regimens in obese patients may be necessary to rapidly achieve the desired PD effect. This is done by measuring the activated partial thromboplastin time (aPTT) when administering UFH (Shank and Zimmerman 2015). Several studies suggest that larger than normal standard dosing of UFH may be warranted in these cases to provide the desired therapeutic aPTT levels (Freeman et al. 2010). Many of these same studies have also suggested that this can be done without risking excessive anticoagulation leading to bleeding. However, while this may allow the target aPTT level to be reached quicker, administering a higher dose of UFH to obese patients may not lead to additional efficacy in reducing the incidence of VTE, and the potential increase in the risk for bleeding should not be ignored (Joy et al. 2016).

Because of the unpredictable bioavailability and inconsistent anticoagulant effects of UFH, low molecular weight heparins (LMWHs) with their predictable dose response (peak anti-Xa activity occurring 3–5 h after injection) have replaced UFH in many treatment paradigms. LMWHs are derived by depolymerization of UFH, with isolation and extraction of low molecular weight fragments. The most consistent and widely used laboratory test for LMWH has been the anti-FXa activity assay, although monitoring is typically not used, it remains an option for high-risk patients (renal insufficiency, obesity, pregnancy, noncompliance) where dosing adjustments may be required. In these cases, anti-Xa plasma levels are typically drawn 4 h after administration, and subsequent dosing adjusted to target levels (Alquwaizani et al. 2013).

The enoxaparin package insert recommends a 1 mg/kg dose; however, many times obese

patients receive an arbitrary lower dose in practice. This seems to be out of a concern that if the recommended dose is followed in obese patients, there is an increased likelihood of supratherapeutic anti-Xa levels and therefore an increased risk of bleeding. However, it is important to note that anti-Xa levels have not been prospectively correlated with any clinical outcomes (Thompson-Moore et al. 2015). That said, a recent study assessed the i.v. administration of enoxaparin sodium (1.5 mg/kg Total Body Weight) as a 6-h infusion and resulted in higher observed maximum PD activity (E_{max}) and overall systemic PD activity (effect)-time curve from time zero to infinity ($AUEC_1$) values for both anti-Xa and anti-IIa levels in the obese patients. The absolute CL and V_d for anti-Xa activity were significantly increased in obese subjects (0.99 L/h vs. 0.74 L/h and 5.77 L vs. 4.37 L, respectively) (Hanley et al. 2010).

Another study conducted by *Thompson-Moore* et al. prospectively assessed enoxaparin dosing in hospitalized morbidly obese patients (Thompson-Moore et al. 2015). The dosing practices observed in the hospital setting seemed to be reflective of other studies reported in the literature. Approximately 53% of patients received less than the recommended 1.0 mg/kg dose of enoxaparin. While 15 patients weighed >150 kg, only 1 patient was dosed with 1.0 mg/kg, again, reflective of an arbitrary dosing limit for these types of patients. Interestingly, despite this under-dosing, greater than 50% of the patients still had supratherapeutic anti-Xa levels. The actual median dose that produced a therapeutic anti-Xa level was 0.83 mg/kg actual body weight) (Thompson-Moore et al. 2015). The increased PD activity observed in this study is likely reflective of the drug's poor distribution into adipose tissue and corresponding increased V_d (Thompson-Moore et al. 2015). While no recommendations are provided in the enoxaparin package insert for dosing in obese patients, considering this increase in PD activity, it is generally recommended that initiation of therapy should occur at a lower initial dose and anti-Xa levels should be monitored and used for dose adjustments as needed (Thompson-Moore et al. 2015).

Warfarin is still the most commonly prescribed anticoagulant globally. Warfarin's mechanism of action involves the inhibition of the vitamin K epoxide reductase complex in the liver thereby depleting the body's vitamin K dependent coagulation Factors II, VII, IX, and X and anticoagulant proteins C and S (Brunton et al. 2006). Warfarin is dosed to a therapeutic PD effect by measuring the international normalized ratio (INR), a value derived from the patient's prothrombin time (PT) laboratory value. Compared to normal weight patients, obese and morbidly obese patients have been found to take a significantly longer median time to achieve therapeutic INR (8 and 10 days vs. 6 days) values and higher average daily doses (6.6 ± 0.3 and 7.6 ± 0.5 vs. 5 ± 0.3 mg) and mean discharge doses (6.7 ± 0.5 and 6.7 ± 0.7 vs. 4.4 ± 0.5 mg). In summary, compared to normal weight patients, obese and morbidly obese patients had a lower initial response to warfarin (Wallace et al. 2013).

Rivaroxaban is an oral, direct Factor Xa inhibitor that targets free and clot-bound Factor Xa and Factor Xa in the prothrombinase complex. Rivaroxaban is the first of a new class of compounds termed direct oral anticoagulants that specifically target a single coagulation factor (such as Factor Xa or thrombin). These compounds were developed in recent years to overcome the limitations of established anticoagulants, particularly warfarin. Factor Xa plays a central role in blood coagulation as it is activated by both the intrinsic and common coagulation pathways. Factor Xa directly converts prothrombin to thrombin via the prothrombinase complex, leading to fibrin clot formation and activation of platelets by thrombin (Mueck et al. 2014). Rivaroxaban does not require routine PD monitoring; however, there is a linear relationship between PT and rivaroxaban concentrations when using a sensitive PT reagent (e.g., Neoplastin+). Additionally, antifactor Xa values derived from an assay using rivaroxaban calibrators can be used to indirectly measure rivaroxaban concentrations.

In a clinical pharmacology study conducted during drug development, there was no clinically significant difference between rivaroxaban PK and PD observed when assessed in obese vs.

nonobese individuals. The C_{max} and AUC values were unaffected in subjects weighing >120 kg. Additionally, rivaroxaban inhibited FXa activity to a similar extent in obese and nonobese individuals. FXa maximum values (E_{max}) occurred 3 to 4 h after rivaroxaban administration for both groups. However, the E_{max} for prolongation of PT decreased significantly (P <0.001) with increasing body weight and was observed 2 to 3 h after administration of rivaroxaban (Kubitza et al. 2007).

As previously mentioned, obesity is a risk factor for cardiovascular disease (CVD), many times caused by hyperlidemia and platelet hyperactivity that leads to atherosclerosis. Excess adipose tissue promotes a prothrombotic and inflammatory state and in turn increases plaque formation. Antiplatelets are commonly used in this CVD. However, observations from several clinical trials have shown that obese individuals have higher platelet reactivity and ultimately display a blunted response to antiplatelet agents (Beavers et al. 2015). Aspirin reduces prostaglandin biosynthesis and irreversibly inhibits cyclooxygenase-1 & -2 (COX) enzyme activity (Brunton et al. 2006). Aspirin exerts its antiplatelet effects through its inhibition of COX-1, which in turn decreases the downstream production of platelet activation by thromboxane A2, a potent platelet activator (Shank and Zimmerman 2015). In a study conducted by *Bordeaux* et al., obese patients consistently showed greater platelet reactivity both before and after administration of low-dose aspirin (81 mg daily). Increasing the dose of aspirin did not appear to increase platelet activity suppression (Bordeaux et al. 2010). These results suggest that obese individuals have an innate hyperaggregable state, which have been suggested in other studies (Bordeaux et al. 2010).

Clopidogrel is an irreversible $P2Y_{12}$ receptor antagonist. Due to its improved tolerability, reduced incidence of hematological side effects, more rapid onset of action, and a convenient dosing regimen, this second-generation thienopyridine has largely replaced ticlopidine (a first-generation thienopyridine with similar efficacy). Clopidogrel asserts its PD effects by specifically and irreversibly binding to $P2Y_{12}$, a subtype of the adenosine diphosphate (ADP) receptor, on the surface of platelets (Jiang et al. 2015). This PD effect is measured by ADP-induced platelet aggregation in platelet-rich plasma which has been the gold standard for assessing platelet function in relation to the clinical outcome (Jiang et al. 2015). Obesity has been shown to significantly affect clopidogrel response. Several studies have reported that BMI or body weight is associated with high platelet reactivity (HPR) in both patients and healthy subjects.

Measurement of platelet aggregation utilizing the VerifyNow $P2Y_{12}$ assay is a fast, standardized point-of-care method that determines platelet-induced aggregation in whole blood by using ADP and prostaglandin E1. A recent clinical pharmacokinetic study reported that, compared with patients with lower body weight, patients with higher body weight had approximately 30% lower AUC values for the pharmacologically active metabolite R-130964 (clop-AM), which ultimately led to higher on treatment platelet reactivity in these obese patients. This can be seen when assessing the VerifyNow $P2Y_{12}$ reaction reading in patients, which was 207 $P2Y_{12}$ reaction units (PRU) for those considered obese and 152 PRU for those considered of normal weight (Jiang et al. 2015). The increased platelet reactivity observed in obese patients is likely a result of the inflammatory state created by obesity. This state may set in motion a number of mechanisms that lead to increased reactivity, suppression of some CYP-P450 activity, and platelet turnover, that all contribute to this poor clopidogrel response (Jiang et al. 2015).

Endocrine System

Estrogens/Progestins/Contraceptives

Obesity adds to the overall risk of pregnancy. Those females of reproductive age who are obese have an increased risk of spontaneous abortion, preeclampsia, gestational diabetes, shoulder dystocia, and cesarean section (Robinson and Burke 2013). While obese females have the same potential risks for altered drug PK as any

other obese individual (i.e., the potential for changes in drug absorption, distribution, metabolism, and excretion), there are also potential PD changes in the outcomes of contraceptive therapy (Robinson and Burke 2013). For example, the suppression of ovarian follicular development, alterations in cervical mucus characteristics, higher serum estrogen concentrations, lower progesterone concentrations, decreased luteinizing hormone concentrations, and altered rhythmic patterns of some of these during the menstrual cycle (Robinson and Burke 2013).

There are various forms of contraceptives, pill forms, intrauterine devices, subdermal implants, and injectables. Additionally, there are single hormonal and combined hormonal formulations. Progestin-only contraceptives can be found in each of these forms. The two intrauterine systems (IUS) available, release progestin levonorgestrel (LNG) come in both 3- and 5-year LNG-IUS formulations (Robinson and Burke 2013). The main mechanism of action of the LNG-IUS is that it delivers hormone directly to the endometrium where it leads to endometrial thinning and decidualization. Considering that the mechanism of action occurs locally on the uterus, there is no reason to think that these systems would be less effective in obese versus nonobese women (Robinson and Burke 2013). While some authors have noted that insertion of the IUS in obese individuals may be problematic, the efficacy that is preserved across BMIs makes these systems ideal for obese women wishing to delay or avoid pregnancy.

Currently marketed contraceptive subdermal implants include etonogestrel (ENG) and levonorgestrel (LNG). A study by *Mornar* et al. assessed the PK of ENG implants in obese (BMI $\geq$30 kg/m^2) women and nonobese women that were consistent with historical controls (Mornar et al. 2012). The researchers observed that the obese women had approximately 50% lower ENG AUC values than the nonobese women. Additionally, this lower systemic exposure led to an estimated 40% lower ENG exposure over the life of the implant; however, none of these women were projected to have an ENG level below the estimated cut-off concentration for reliable ovulation suppression (90 pg/mL) (Mornar et al. 2012).

Depot medroxyprogesterone acetate (DMPA-i.m.) is an intramuscular injectable contraceptive that is widely used around the world at a dose of 150 mg and provides highly effective contraception for approximately 3 months. The subcutaneous formulation (DMPA-s.c.) delivers a lower dose of DMPA (104 mg) and also provides comparable contraceptive effect. Neither the efficacy of the i.m. nor s.c. formulations are known to be decreased in obese women. While one study that evaluated the PK and PD of DMPA-s.c. did observe lower serum concentrations in obese subjects, levels remained above the threshold needed for ovulation suppression (Robinson and Burke 2013).

The most frequently used agents are the combination oral contraceptives containing both an estrogen and a progestin. The theoretical efficacy of this combination is approximately 99.9%. Both ethinyl estradiol (EE) and mestranol are the two estrogens used in the various formulations, while there is a greater variety of progestins that are included (Brunton et al. 2006). The estrogen component contributes to ovulation suppression and control of irregular bleeding, while the inclusion of a progestin suppresses the hypothalamic–pituitary–ovarian axis leading to inhibition of ovulation. The transdermal patch and vaginal ring also contain a combination of estrogen and progestin, and their mechanism of action is the same as the oral formulations (Robinson and Burke 2013).

In previous studies that assessed the various PK parameters in obese women taking the combination oral contraceptive, 20-mg EE/100-mg levonorgestrel (LNG), the $t_{1/2}$ was prolonged, systemic exposure measured by AUC values were larger and CL was lower. At the same time, it was observed that obese women had higher levels of ovarian hormone production (estradiol and progesterone), suggesting greater ovarian activity and the potential for decreased contraceptive efficacy (Robinson and Burke 2013; Simmons and Edelman 2016). These PK changes led to longer times to reach steady-state concentrations and therefore a longer time to reach levels sufficient for ovulatory suppression. In one study, steady state concentrations in one obese individual was not achieved even after the full 21-day pill cycle

(Robinson and Burke 2013; Simmons and Edelman 2016).

However, other studies have demonstrated very different results. For example, one study assessing the use of a 30 μg EE/150 μg LNG combined oral contraceptive in nonobese and obese women found obese individuals had lower AUC, C_{max} and T_{max} values as well as a prolonged t½ for both hormones, compared with nonobese women. Results are somewhat contradictory to the previously mentioned studies. Furthermore, PD assessments indicated that there was no significant difference in ovarian follicular activity across BMI, supporting the conclusion that despite the PK changes observed, conceptive efficacy was maintained in obese individuals. When considering the limited data available on the use of oral contraceptives in obese women, most studies tend to demonstrate that there is not a significant difference in effectiveness across weight or BMI. However, it is important to note that these studies rarely included women that were morbidly obese, and the total amount of data is insufficient.

Again, there is limited data on the PK, PD, or efficacy of transdermal contraceptive patches. A study by *Westhoff* et al. stratified participants according to three BMI groups (Group 1 $\leq$30 kg/m^2; Group 2, n >30 kg/m^2 and $\leq$35 kg/m^2; and Group 3 >35 kg/m^2) (Westhoff et al. 2014). Each participant received a transdermal patch containing 0.55-mg EE and 2.1-mg gestodene (GSD). Each patch was used weekly for three 28-day cycles and its PD effect was measured by the Hoogland score, which is a composite score that comprises of transvaginal ultrasound and estradiol (E_2) and progesterone levels every 3 days in Cycles 2 and 3. Additionally, PK and EE, GSD, and sex hormone-binding globulin were assessed (Westhoff et al. 2014).

Study results reported that only six ovulations occurred during the study, and no participant ovulated in both study cycles. The ovulations observed occurred across the different weight groups and was unaffected by differences in BMI. A majority of participants had Hoogland scores of 1 or 2 regardless of BMI grouping. Follicle-like structures <13 mm were reported in ~80% (Cycle 2) and ~86% (Cycle 3) in Group 1; ~61% (Cycle 2) and ~75% (Cycle 3) in Group 2; and ~78% (Cycle 2) and ~73% (Cycle 3) in Group 3. Hormone levels (follicle-stimulating hormone (FSH), luteinizing hormone (LH), E2, and progesterone) were similar across BMI groups. The authors concluded that the EE/GSD patch provided effective ovulation inhibition across BMI strata (Westhoff et al. 2014).

Emergency contraceptive pills (ECPs) are effective at reducing the risk of pregnancy up to 5 days after unprotected intercourse. Their mechanism of action involves the inhibition of delaying of ovulation and is most effective if taken soon after unprotected intercourse. The LNG-based regimens are provided as either a single dose of 1.5 mg LNG or two doses of 0.75 mg LNG taken 12 h apart and are effective in reducing the risk of pregnancy up to 72 h after unprotected intercourse, while the ulipristal acetate (UPA) ECP, which is a progesterone agonist/antagonist, is a one-time 30 mg dose that is effective up to 5 days after unprotected intercourse (Glasier 2013).

Similar to the other contraceptives discussed, there are limited data on the effects of obesity on these treatments. A meta-analysis performed by *Glasier* et al. looked at the combined results from two randomized clinical trials comparing LNG and UPA ECPs (Glasier et al. 2011). The authors reported that the risk of pregnancy was >3-fold for obese women compared to nonobese women whichever ECP was taken. However, this risk was greater for those taking LNG than for those taking UPA. Interestingly, women who had unprotected intercourse after using ECP were more likely to get pregnant than those who did not, regardless of type used (Glasier et al. 2011). Consistent with Glasier et al., a small clinical pharmacology study conducted by *Edelman* et al. assessed the PK parameters of nonobese (median 22.8 kg/m^2) and obese women (median 39.5 kg/m^2) dosed with both a single and double dose of LNG ECP. The single dose of LNG ECP in obese women resulted in a significantly lower C_{max} ($C_{max\text{-}obese}$ = 5.57 ng/mL) than that observed in normal weight women ($C_{max\text{-}non\text{-}obese}$ = 10.30 ng/mL), approximately 50% lower. Doubling the dose of LNG ECP increased the C_{max} significantly ($C_{max\text{-}obese}$ = 10.52 ng/mL) essentially

normalizing the C_{max} level to that of the normal BMI subjects receiving a single ECP dose (Edelman et al. 2016).

Invading Organisms

Antimicrobials

Antimicrobial agents are classified based on chemical structure and proposed mechanism of action. There are those agents that: (1) inhibit synthesis of bacterial cell walls; (2) act directly on the cell membrane, increasing permeability and compromising the structure of the microorganism; (3) disrupt ribosomal function to inhibit protein synthesis and are bacteriostatic; (4) disrupt ribosomal function to alter protein synthesis and are bactericidal; (5) affect bacterial nucleic acid metabolism; and (6) are antimetabolites that block essential enzymes of folate metabolism (Brunton et al. 2006).

Additionally, these mechanisms of action can be further divided into time-dependent or concentration-dependent effects and each antimicrobial class has a unique PK/PD target. For example, β-lactam antibiotics are time-dependent, they produce the most effective PD response when the concentration of the free drug remains above the minimum inhibitory concentration (MIC). Aminoglycosides are concentration-dependent, with the most effective PD response occurring when the C_{max} of the drug is over the MIC (C_{max}/MIC) (3). Then there are those compounds that are both time- and concentration-dependent, therefore when the AUC of the antimicrobial from 0–24 h is over the MIC (AUC_{0-24}/MIC) thereby drives bacterial killing (Alobaid et al. 2016).

The β-lactam antibiotics fall into the first class of agents described that inhibit the synthesis of bacterial cell walls. They include the penicillins, cephalosporins, β-Lactamase inhibitors, and Carbapenems (Brunton et al. 2006). A prospective study conducted by *Hites* et al. assessed the PK and PD parameters of infected obese patients (BMI $\geq$30 $kg \cdot m^{-2}$) who received either meropenem (MEM), piperacillin-tazobactam (TZP), or cefepime/ceftazidime (CEF) β-lactam antibiotics. The primary PD parameter in this study was defined as the "clinical breakpoint" of the microorganisms assessed for different antimicrobial therapies (Hites et al. 2014).

These breakpoints, as defined by the European Committee on Antimicrobial Susceptibility Testing (EUCAST), have been fixed to ensure a good probability of therapeutic success. The breakpoint for β-lactams used by EUCAST is defined as the drug's free fraction (fT) >1× MIC [fT>MIC] during 40–50%, 50–60%, and 60–70% of the dosage interval for MEM, TZP, and CEF, respectively. It should be noted that a higher PD target is frequently used when clinical situations arise where patients are suffering from conditions like septic shock or are neutropenic due to oncological treatment. These situations may justify the use of a total fraction (T)>4×MIC [T>4MIC] for 40%, 50%, and 70% of the dosage interval for MEM, TZP, and CEF, respectively (Hites et al. 2014).

Different pathogens have different PD targets. For example, for infections caused by *Enterobacteriaceae* spp. and *P. aeruginosa*, the EUCAST's clinical breakpoints for these pathogens are: 2 mg/L for MEM, 8 mg/L and 16 mg/L for TZP, and 1 mg/L and 8 mg/L for CEF, respectively. Total serum concentrations and fT at 40%, 50%, and 70% of the dosage intervals for MEM, TZP, and CEF, respectively, were as follows: 6.2 mg/L and 6.1 mg/L for MEM, 36.7 mg/L and 25.7 mg/L for TZP, and 16.1 mg/L and 12.5 mg/L for CEF (Hites et al. 2014).

When evaluating the fT>MIC for *Enterobacteriacae* spp., adequate serum concentrations were obtained in 93% of patients receiving MEM, 84% of patients receiving TZP, and 91% of patients receiving CEF. For infections due to *P. aeruginosa*, adequate fT>MIC were obtained in 93% of patients receiving MEM, 68% of patients receiving TZP, and 73% of patients receiving CEF. According to the EUCAST criteria, any percentage of PD target that is <90% was considered insufficient therapeutic coverage (Hites et al. 2014).

Therefore, when standard doses of these antibiotics were administered to obese, noncritically ill patients, infected with *Enterobacteriaceae* spp., only MEM and CEF reached therapeutic serum concentrations. Additionally, when the

same antibiotics were administered to obese, noncritically ill patients infected with *P. aeruginosa*, only MEM reached therapeutic serum concentrations (Hites et al. 2014).

When the investigators assessed the higher PD target (T>4MIC) for infections from *Enterobacteriacae* spp. and *P. aeruginosa*, much fewer patients reached the PD target. Adequate serum concentrations were only reached for 21% for MEM, 55% for TZP, and 91% for CEF with *Enterobacteriacae* spp. infections and for 21% for MEM, 19% for TZP, and 18% for CEF *P. aeruginosa* infections. Therefore, only CEF with 91% of patients infected with *Enterobacteriacae* spp. met the EUCAST criteria (Hites et al. 2014).

The decreased serum concentrations observed in this study is likely caused from an increase in V_d and CL for all three study drugs, when compared to non-obese individuals. A CrCL of >150 mL/min was observed in approximately 25% of those in the obese cohort. This augmented renal CL appears to be the major risk factor for failure to reach therapeutic concentrations. Together, these data suggest that standard dosage regimens, particularly for TZP, are insufficient in obese, noncritically ill patients (Hites et al. 2014).

Increased drug CL has been described in critically ill patients and has been termed augmented renal clearance (ARC), a condition where renal elimination of circulating solutes is increased. ARC is defined as a CrCL of ≥130 mL/min/1.73 m^2. ARC is associated with subtherapeutic antimicrobial concentrations and worse clinical outcomes in critically ill patients receiving standard doses of antimicrobial therapy.

An increase in renal CL has been observed in obese individuals who have normal kidney function, most likely due to the increased kidney size and renal blood flow associated with obesity. While this altered physiology may result in lower antimicrobial concentrations, higher concentrations may be observed in obese patients with co-morbidities such as diabetic nephropathy. Obese individuals are more likely to have pathologies that cause hepatic dysfunction, such as hepatic steatosis, possibly resulting in decreased drug metabolism. Obesity may also have an impact on different hepatic enzyme systems causing increased (e.g., CYP2E1) or decreased (e.g., CYP3A4) activity.

In another study conducted by *Sturm* et al., the use of piperacillin/tazobactam was examined in critically ill morbidly obese patients (BMI >40 kg/m^2). All patients achieved the PK/PD target of 100% *f*T>MIC for pathogens with an MIC of 16 mg/L using a piperacillin/tazobactam dose of 4.5 g i.v. every 6 h. Morbidly obese patients had a higher piperacillin V_d (31.0 L vs. 22.4 L) and lower CL (6.0 L/h vs. 13.7 L/h) compared with nonobese patients (Atkinson et al. 2007). The net result was a $t_{1/2}$ of 3.7 h compared with 1 h reported in other populations. This longer $t_{1/2}$ likely contributed to an extended % *f*T>MIC for susceptible pathogens. Based on these results, it would appear that the tested 4.5 g i.v. dose every 6 h would be sufficient to attain the desired PK/PD target of %fT>MIC (Sturm et al. 2014).

While the effects of obesity on penicillin PK and PD parameters are sparse to nonexistent in the literature, compounds like ampicillin, penicillin, ticarcillin would likely exhibit similar changes in their PK and PD to that of piperacillin.

Cefoxitin is a second-generation cephamycin antibiotic and classified as a semisynthetic, broad-spectrum, cephalosporin. Cefoxitin is commonly used for perioperative parenteral surgical prophylaxis. *Moine* et al. studied the PK and PD of a 40 mg/kg i.v. dose in morbidly obese individuals (Moine et al. 2016). Although the dose used in this study was substantially higher than the standard cefoxitin doses typically used, the C_{max} observed in this study was similar to those previously reported for nonobese populations receiving a much lower weight-based dose. This lower than expected C_{max} along with a prolonged $t_{1/2}$ is consistent with the approximate twofold higher V_d values observed in this study compared to nonobese populations. Additionally, despite the use of these higher doses, tissue concentrations were poor, with an average tissue/serum ratio of 8%, and below the Clinical and Laboratory Standards Institute (CLSI) breakpoint for anaerobes targeted by this antibiotic. As with the other β-lactams, the time during which unbound drug concentrations are greater than the pathogen MICs (*f*T>MIC) is

the PD parameter best correlated with clinical efficacy (Moine et al. 2016).

The authors noted that the PD target for surgical prophylaxis is largely undefined. Considering that surgical contamination may occur at any point during the procedure, a *f*TMIC of 100% was suggested as an ideal target for cefoxitin and β-lactam antibiotics in general, during surgery. Even though the authors calculated a more conservative *f*TMIC of 70%, tissue concentrations failed to reach the needed level for efficacy. Concentrations were below the susceptibility breakpoint for *S. aureus* and *E. coli*, suggesting inadequate coverage if an intraoperative contamination would have occurred. The authors suggest that although the weight-based dose at 40 mg/kg performed better than a standard 2-g dose, it would likely be inadequate to prevent infection (Moine et al. 2016).

Aminoglycosides are primarily used to treat infections caused by aerobic gram-negative bacteria. They are considered bactericidal as they disrupt ribosomal function to alter protein synthesis. This class of antibiotic includes gentamicin, tobramycin, amikacin, kanamycin, netilmicin, streptomycin, and neomycin (Brunton et al. 2006). The aminoglycoside antimicrobials are hydrophilic weak bases and have a corresponding low V_d (closer to blood volume). These compounds have optimal bactericidal activity when achieving peak concentrations that are 8x to 10x the MIC of the targeted pathogen (Hanley et al. 2010). A study conducted by *Bauer* et al. assessed the steady-state PK of three aminoglycosides – gentamicin, tobramycin, and amikacin – in morbidly obese subjects (Bauer et al. 1983). The investigators observed that the mean V_d values were substantially larger for the morbidly obese subjects compared to the nonobese subjects: gentamicin ($V_{d\text{-obese}}$ = 23.31 L vs. $V_{d\text{-non-obese}}$ = 17.01 L), tobramycin ($V_{d\text{-obese}}$ = 29.01 L vs. $V_{d\text{-non-obese}}$ = 18.31 L), and amikacin ($V_{d\text{-obese}}$ = 26.81 L vs. $V_{d\text{-non-obese}}$ = 18.61 L). Similarly, CL values were larger in the morbidly obese subjects compared to the nonobese subjects: gentamicin (CL_{obese} = 135.8 mL/min vs. $CL_{non\text{-}obese}$ = 95.9 mL/min), tobramycin (CL_{obese} = 162.4 mL/min vs. $CL_{non\text{-}obese}$ = 101.3 mL/min), and amikacin (CL_{obese} = 157.3 mL/min vs. $CL_{non\text{-}obese}$ = 99.2 mL/min). Albeit the significant increases in V_d and CL, there was no significant difference between predicted and measured steady-state concentrations as determined by C_{max} and C_{min} values. This is likely attributed to the larger creatinine clearance values secondary to hyperfiltration commonly observed in obese individuals (Bauer et al. 1983).

Daptomycin is a lipopeptide antibiotic used in the treatment of systemic and life-threatening infections caused by gram-positive organisms (i.e., enterococci, staphylococci, streptococci). The MIC of daptomycin (MIC_{90}) is typically ≤1 μg/mL for staphylococci and streptococci and 2 to 4 μg/mL for enterococcal species. This compound is both time- and concentration-dependent, in which bacterial eradication is dependent upon the ratio of $AUC_{0–24h}$ to the MIC ($AUC_{0–24h}$/MIC). *Dvorchik* and *Damphousse* conducted a study assessing the PK of daptomycin in moderately obese (BMI between 25 and 39.9 kg/m^2) or morbidly obese (BMI ≥40 kg/m^2) compared to matched nonobese (BMI between 18.5 and 24.9 kg/m^2) subjects (Dvorchik and Damphousse 2005).

After administration of a 4-mg/kg total body weight dose, the C_{max} and AUC values for both obese groups were higher compared to their respective matched nonobese controls: Moderately Obese [$C_{max\text{-}obese}$ = 57.75 μg/mL vs. $C_{max\text{-}non\text{-}obese}$ = 46.28 μg/mL and $AUC_{(0\text{-}\infty)obese}$ = 420.53 μg·h/mL vs. $AUC_{(0\text{-}\infty)non\text{-}obese}$ = 322.37 μg·h/mL] and Morbidly obese [$C_{max\text{-}obese}$ = 67.00 μg/mL vs. $C_{max\text{-}non\text{-}obese}$ = 53.22 μg/mL and $AUC_{(0\text{-}\infty)obese}$ = 547.78 μg·h/mL vs. $AUC_{(0\text{-}\infty)non\text{-}obese}$ = 418.76 μg·h/mL]. These differences equate to mean C_{max} values that were ~25% higher in the obese groups than in their matched controls and AUC values were ~30% to 35% greater in the obese groups compared to their matched controls.

Additionally, significant differences in daptomycin V_d were observed between obese and nonobese groups. The obese group saw increases in absolute V_d that were ~25% larger in the moderately obese subjects and ~55% in the morbidly obese subjects compared to the

respective controls. In a similar trend, CL values were also increased in obese vs. nonobese subjects: moderately obese CL_{obese} = 855.80 mL/h vs. $CL_{non\text{-}obese}$ = 732.80 mL/h and morbidly obese CL_{obese} = 1015.83 mL/h vs. $CL_{non\text{-}obese}$ = 696.41 mL/h. Interestingly, renal CL was not substantially different between obese and nonobese subjects. The authors speculate that this may be a consequence of all subjects having an estimated creatinine clearance ≥70 mL/min, however did not provide any information on whether obese subjects had CrCL values representative of abnormally high glomerular filtration rate (GFR) or hyperfiltration (Dvorchik and Damphousse 2005).

Vancomycin is a tricyclic glycopeptide antibiotic that is commonly used to treat Gram positive pathogens, for example, severe staphylococcal and enterococcal infections. Vancomycin exerts a time-dependent antibacterial effect, and its clinical response is a function of the $AUC_{0–24h}$ and MIC ($AUC_{0–24h}$/MIC). Current guidelines recommend maintaining vancomycin trough concentrations >10 μg/mL to avoid *Staphylococcus aureus* resistance and between 15 and 20 μg/mL in complicated infections such as bacteremia, endocarditis, and osteomyelitis. Guidance is also provided for treatment of *methicillin-resistant staphylococcus aureus* (MRSA) in which a dosage of 15–20 mg/kg every 8–12 h without exceeding 2000 mg of vancomycin/dose. This in turn creates a challenge since obese patients would, however, typically require doses that would exceed 2000 mg/dose due to the initial weight-based dosing paradigm.

A study conducted by *Adane* et al. assessed the PK of vancomycin in obese subjects. For the treatment of *S. aureus*-associated lower respiratory tract infections, clinical success was found with an $AUC_{0–24h}$/MIC >315, whereas a successful microbiologic response required a $AUC_{0–24h}$/MIC >866. Clinical practice guidelines state that an $AUC_{0–24h}$/MIC ≥400 is needed for clinical effectiveness. At this level, the likelihood of *S. aureus* resistance is low and ensures adequate penetration in tissues such as the lung with minimizing potential nephrotoxicity. Those enrolled into the study had a median weight of 147.9 kg, BMI of 49.5 kg/m^2 a Cl_{Cr} of 124.8 mL/min/1.73m^2 and received a median vancomycin dose of 4000 mg/day, resulting in median $AUC_{0–24h}$ that was 582.9 mg·h/L. The mean V_d was 0.51 L/kg, and CL was 6.54 L/h. Simulations indicated that 4000–5000 mg/day of vancomycin in this population provided ≥93% probability of a $AUC_{0–24h}$ /MIC ratio of ≥400 leading to an MIC of 1 μg/mL and therapeutic effectiveness (Adane et al. 2015).

Ciprofloxacin is a broad-spectrum fluoroquinolone antibiotic. Quinolones function by inhibiting DNA gyrase and topoisomerase ultimately inhibiting cell division. Its PD effect is both concentration-dependent and time-dependent, with clinical efficacy best described using an $AUC_{0–24}$/MIC ratio. Ciprofloxacin is active against both gram-positive and gram-negative bacteria. A study conducted by *Allard* et al. assessed the PK of ciprofloxacin and its primary metabolite (desethyleneciprofloxacin) in both obese subjects (mean weight = 110.7 kg; mean BMI = 36.4 kg/m^2) and normal weight subjects (mean weight = 71.8 kg; mean BMI = 23.3 kg/m^2) (Allard et al. 1993). After receiving a single 400 mg iv dose of ciprofloxacin infused over 1 h, ciprofloxacin CL was significantly increased in obese subjects (897.44 mL/min) compared with nonobese subjects (744.44 mL/min), additionally CL_{renal} in obese subjects was 29% higher than in nonobese subjects. Lastly, ciprofloxacin V_d was larger in obese subjects ($V_{d\text{-}obese}$ = 269.17 L) than nonobese subjects ($V_{d\text{-}non\text{-}obese}$ = 219.03 L) (Allard et al. 1993).

Antifungals

The azole antifungals include the imidazole and triazole classes, which share the same antifungal spectrum and mechanism of action. Azoles inhibit the fungal cytochrome P450 enzyme 14α-sterol demethylase thereby inhibiting fungal growth (Brunton et al. 2006). Fluconazole is active against a variety of *Candida* spp., with a PK/PD target of $AUC_{0–24h}$/MIC >25. There is scarcity of literature that addresses the effects of obesity on the PK and PD of fluconazole. Of the few publications available, most report data from case studies.

Lopez and Phillips described a case report of a critically ill morbidly obese patient (BMI = 84 kg/m^2) receiving renal replacement therapy and who was being treated with fluconazole at a dose of 600 mg (Atkinson et al. 2007). The investigators calculated a V_d = 163.3 L and CL = 3.25 L/h, for this patient, which was significantly larger than those previously reported for critically ill nonobese patients with acute renal failure (V_d = 65.6 L and CL = 1.9 L/h) (Atkinson et al. 2007).

References and Further Reading

Abernethy DR, Schwartz JB (1988) Verapamil pharmacodynamics and disposition in obese hypertensive patients. J Cardiovasc Pharmacol 11(2):209–215

Abernethy DR, Greenblatt DJ, Smith TW (1981) Digoxin disposition in obesity: clinical pharmacokinetic investigation. Am Heart J 102(4):740–744

Abernethy DR, Greenblatt DJ, Divoll M et al (1984) The influence of obesity on the pharmacokinetics of oral alprazolam and triazolam. Clin Pharmacokinet 9(2):177–183

Adane ED, Herald M, Koura F (2015) Pharmacokinetics of vancomycin in extremely obese patients with suspected or confirmed *Staphylococcus aureus* infections. Pharmacotherapy 35(2):127–139

Allard S, Kinzig M, Boivin G et al (1993) Intravenous ciprofloxacin disposition in obesity. Clin Pharmacol Ther 54(4):368–373

Alobaid AS, Hites M, Lipman J et al (2016) Effect of obesity on the pharmacokinetics of antimicrobials in critically ill patients: a structured review. Int J Antimicrob Agents 47(4):259–268

Alquwaizani M, Buckley L, Adams C, Fanikos J (2013) Anticoagulants: a review of the pharmacology, dosing, and complications. Curr Emerg Hosp Med Rep 1:83–97

Anderson WJ, Lipworth BJ (2012) Does body mass index influence responsiveness to inhaled corticosteroids in persistent asthma? Ann Allergy Asthma Immunol 108(4):237–242

Atkinson AJ, Abernethy DR, Daniels CE, Dedrick RL, Markey SP (2007) Principles of clinical pharmacology, 2nd edn. Academic, Amsterdam

Bauer LA, Drew Edwards WA, Patchen Dellinger E et al (1983) Influence of weight on aminoglycoside pharmacokinetics in normal weight and morbidly obese patients. Eur J Clin Pharmacol 24:643–647

Beavers CJ, Heron P, Smyth SS et al (2015) Obesity and antiplatelets-does one size fit all? Thromb Res 136:712–716

Blouin RA, Warren GW (1999) Pharmacokinetic considerations in obesity. J Pharm Sci 88(1):1–7

Bordeaux BC, Qayyum R, Yanek LR et al (2010) Effect of obesity on platelet reactivity and response to low-dose aspirin. Prev Cardiol 13:56–62

Bowman SL, Hudson SA, Simpson G et al (1986) A comparison of the pharmacokinetics of propranolol in obese and normal volunteers. Br J Clin Pharmacol 21:529–532

Brill MJ, Diepstraten J, van Rongen A et al (2012) Impact of obesity on drug metabolism and elimination in adults and children. Clin Pharmacokinet 51(5):277–304

Brill MJ, van Rongen A, Houwink AP et al (2014a) Midazolam pharmacokinetics in morbidly obese patients following semi-simultaneous oral and intravenous administration: a comparison with healthy volunteers. Clin Pharmacokinet 53:931–941

Brill MJ, Houwink AP, Schmidt S et al (2014b) Reduced subcutaneous tissue distribution of cefazolin in morbidly obese versus non-obese patients determined using clinical microdialysis. J Antimicrob Chemother 69:715–723

Brunton LL, Lazo JS, Parker KL (eds) (2006) Goodman and Gilman's The pharmacological basis of therapeutics, 11th edn. McGraw-Hill, New York

Casati A, Putzu M (2005) Anesthesia in the obese patient: pharmacokinetic considerations. J Clin Anesth 17(2): 134–145

Cataldi M, di Geronimo O, Trio R, Scotti A et al (2016) Utilization of antihypertensive drugs in obesity-related hypertension: a retrospective observational study in a cohort of patients from southern Italy. BMC Pharmacol Toxicol 17:9

Cheymol G (2000) Effects of obesity on pharmacokinetics implications for drug therapy. Clin Pharmacokinet 39(3):215–231

Cheymol G, Poirier J-M, Barre J, Pradalier A, Dry J (1987) Comparative pharmacokinetics of intravenous propranolol in obese and normal volunteers. J Clin Pharmacol 27(11):874–879

Cheymol G, Woestenborghs R, Snoeck E et al (1997) Pharmacokinetic study and cardiovascular monitoring of nebivolol in normal and obese subjects. Eur J Clin Pharmacol 51(6):493–498

Cho S-J, Yoon I-S, Kim D-D (2013) Obesity-related physiological changes and their pharmacokinetic consequences. J Pharm Investig 43:161–169

Christoff PB, Conti DR, Naylor C, Jusko WJ (1983) Procainamide disposition in obesity. Drug Intell Clin Pharm 17(7–8):516–522

de la Peña A, Yeo KP, Linnebjerg H et al (2015) Subcutaneous injection depth does not affect the pharmacokinetics or glucodynamics of insulin lispro in normal weight or healthy obese subjects. J Diabetes Sci Technol 9(4):824–830

Dunn TE, Ludwig EA, Slaughter RL et al (1991) Pharmacokinetics and pharmacodynamics of methylprednisolone in obesity. Clin Pharmacol Ther 49(5):536–549

Dvorchik BH, Damphousse D (2005) The pharmacokinetics of daptomycin in moderately obese, morbidly obese, and matched non-obese subjects. J Clin Pharmacol 45:48–56

Edelman AB, Cherala G, Blue SW et al (2016) Impact of obesity on the pharmacokinetics of levonorgestrel-based emergency contraception: single and double dosing. Contraception 94(1):52–57

Fan J, de Lannoy IA (2014) Pharmacokinetics. Biochem Pharmacol 87:93–120

Farrell GC, Teoh N, McCuskey R (2008) Hepatic microcirculation in fatty liver disease. Anat Rec Adv Integr Anat Evol Biol 291(6):684–692

Freeman AL, Pendleton RC, Rondina MT (2010) Prevention of venous thromboembolism in obesity. Expert Rev Cardiovasc Ther 8(12):1711–1721

Fukuchi H, Nakashima M, Araki R et al (2009) Effect of obesity on serum amiodarone concentration in Japanese patients: population pharmacokinetic investigation by multiple trough screen analysis. J Clin Pharm Ther 34(3):329–336

Galletti F, Fasano ML, Ferrara LA et al (1989) Obesity and beta-blockers: influence of body fat on their kinetics and cardiovascular effects. J Clin Pharmacol 29 (3):212–216

Gandhi A, Moorthy B, Ghose R (2012) Drug disposition in pathophysiological conditions. Curr Drug Metab 13 (9):1327–1344

Glasier A (2013) Emergency contraception: clinical outcomes. Contraception 87(3):309–313

Glasier A, Cameron ST, Blithe D et al (2011) Can we identify women at risk of pregnancy despite using emergency contraception? Data from randomized trials of ulipristal acetate and levonorgestrel. Contraception 84(4):363–367

Hanley MJ, Abernethy DR, Greenblatt DJ (2010) Effect of obesity on the pharmacokinetics of drugs in humans. Clin Pharmacokinet 49(2):71–87

Hites M, Taccone FS, Wolff F et al (2014) Broad-spectrum β-lactams in obese non-critically ill patients. Nutr Diabetes 4(6):e119

Jain R, Chung SM, Jain L et al (2011) Implications of obesity for drug therapy: limitations and challenges. Clin Pharmacol Ther 90(1):77–89

Jensen MD, Ryan DH, Apovian CM et al (2014) 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults. J Am Coll Cardiol 63:2985–3023

Jiang X-L, Samant S, Lesko LJ, Schmidt S (2015) Clinical pharmacokinetics and pharmacodynamics of clopidogrel. Clin Pharmacokinet 54:147–166

Joy M, Tharp E, Hartman H et al (2016) Safety and efficacy of high-dose unfractionated heparin for prevention of venous thromboembolism in overweight and obese patients. Pharmacotherapy 36(7):740–748

Jusko WJ (2017) Clarification of contraceptive drug pharmacokinetics in obesity. Contraception 95: 10–16

Kees MG, Weber S, Kees F, Horbach T (2011) Pharmacokinetics of moxifloxacin in plasma and tissue of morbidly obese patients. J Antimicrob Chemother 66:2330–2335

Knibbe CA, Brill MJ, van Rongen A et al (2015) Drug disposition in obesity: toward evidence-based dosing. Annu Rev Pharmacol Toxicol 55:149–167

Kubitza D, Becka M, Zuehlsdorf M, Mueck W (2007) Body weight has limited influence on the safety, tolerability, pharmacokinetics, pharmacodynamics of rivaroxaban (BAY 59-7939) in healthy subjects. J Clin Pharmacol 47:218

Landsberg L, Aronne LJ, Beilin LJ et al (2013) Obesity-related hypertension: pathogenesis, cardiovascular risk, and treatment. A position paper of the obesity society and the American Society of Hypertension. J Clin Hypertens (Greenwich) 15:14–33

Lentz SR (2016) Thrombosis in the setting of obesity or inflammatory bowel disease. Blood 128(20): 2388–2394

Levy BI, Schiffrin EL, Mourad JJ et al (2008) Impaired tissue perfusion a pathology common to hypertension, obesity, and diabetes mellitus. Circulation 118 (9):968–976

Leykin Y, Miotto L, Pellis T (2011) Pharmacokinetic considerations in the obese. Best Pract Res Clin Anaesthesiol 25:27–36

Martin JH, Saleem M, Looke D (2012) Therapeutic drug monitoring to adjust dosing in morbid obesity – a new use for an old methodology. Br J Clin Pharmacol 73 (5):685–690

Michalaki MA, Gkotsina MI, Mamali I et al (2011) Impaired pharmacokinetics of levothyroxine in severely obese volunteers. Thyroid 21(5):477–481

Moine P, Mueller SW, Schoen JA et al (2016) Pharmacokinetic and pharmacodynamic evaluation of a weight-based dosing regimen of cefoxitin for perioperative surgical prophylaxis in obese and morbidly obese patients. Antimicrob Agents Chemother 60(10): 5885–5893

Mornar S, Chan LN, Mistretta S et al (2012) Pharmacokinetics of the etonogestrel contraceptive implant in obese women. Am J Obstet Gynecol 207(2):110. e1–110.e6

Morrish GA, Pai MP, Green B (2011) The effects of obesity on drug pharmacokinetics in humans. Expert Opin Drug Metab Toxicol 7(6):697–706

Mueck W, Stampfuss J, Kubitza D, Becka M (2014) Clinical pharmacokinetic and pharmacodynamic profile of rivaroxaban. Clin Pharmacokinet 53:1–16

Munjal S, Gautam A, Rapoport AM, Fisher DM (2016) The effect of weight, body mass index, age, sex, and race on plasma concentrations of subcutaneous sumatriptan: a pooled analysis. Clin Pharmacol Adv Appl 8:109–116

Reflection paper on investigation of pharmacokinetics and pharmacodynamics in the obese population 25 January 2018 1, EMA/CHMP/535116/2016 2, Committee for Human Medicinal Products (CHMP)

Robinson JA, Burke AE (2013) Obesity and hormonal contraceptive efficacy. Women's Health (Lond Engl) 9(5):453–466

Rowland M, Tozer TN (1995) Clinical pharmacokinetics, concepts and applications, 3rd edn. Lippincott Williams & Wilkins, Philadelphia

Shah DK, Missmer SA, Correia KF, Ginsburg ES (2014) Pharmacokinetics of human chorionic gonadotropin injection in obese and normal-weight women. J Clin Endocrinol Metab 99(4):1314–1321

Shank BR, Zimmerman DE (2015) Demystifying drug dosing in obese patients. American Society of Health System Pharmacists. eBook

Simmons KB, Edelman AB (2016) Hormonal contraception and obesity. Fertil Steril 106(6):1282–1288

Smit C, De Hoogd S, Brüggemann RJM, Knibbe CAJ (2018) Obesity and drug pharmacology: a review of the influence of obesity on pharmacokinetic and pharmacodynamic parameters. Expert Opin Drug Metab Toxicol 14(3):275–285

Steinkampf MP, Hammond KR, Nichols JE, Slayden SH (2003) Effect of obesity on recombinant follicle stimulating hormone absorption: subcutaneous versus intramuscular administration. Fertil Steril 80(1):99–102

Sturm AW, Allen N, Rafferty KD et al (2014) Pharmacokinetic analysis of piperacillin administered with tazobactam in critically ill, morbidly obese surgical patients. Pharmacotherapy 34(1):28–35

Thompson-Moore NR, Wanat MA, Putney DR et al (2015) Evaluation and pharmacokinetics of treatment dose enoxaparin in hospitalized patients with morbid obesity. Clin Appl Thromb Hemost 21(6):513–520

Tucker GT (1981) Measurement of the renal clearance of drugs. Br J Clin Pharmacol 12:761–770

van Kralingen S, Diepstraten J, Peeters MY et al (2011) Population pharmacokinetics and pharmacodynamics of propofol in morbidly obese patients. Clin Pharmacokinet 50:739–750

van Rongen A, Välitalo PAJ, Peeters MYM et al (2016) Morbidly obese patients exhibit increased CYP2E1-mediated oxidation of acetaminophen. Clin Pharmacokinet 55:833–847

Wallace JL, Reaves AB, Tolley EA et al (2013) Comparison of initial warfarin response in obese patients versus non-obese patients. J Thromb Thrombolysis 36(1):96–101

Westhoff CL, Reinecke I, Bangerter K, Merz M (2014) Impact of body mass index on suppression of follicular development and ovulation using a transdermal patch containing 0.55-mg ethinyl estradiol/2.1-mg gestodene: a multicenter, open-label, uncontrolled study over three treatment cycles. Contraception 90(3):272–279

Wójcicki J, Jaroszynska M, Droździk M et al (2003) Comparative pharmacokinetics and pharmacodynamics of propranolol and atenolol in normolipaemic and hyperlipidaemic obese subjects. Biopharm Drug Dispos 24 (5):211–218

World Health Organization (2018) Obesity and overweight. http://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight. Accessed 8 Oct 2018

Wu B (2016) Morbid obesity alters both pharmacokinetics and pharmacodynamics of propofol: dosing recommendation for anesthesia induction (Short communication). Drug Metab Dispos 44: 1579–1583

Zuckerman M, Greller HA, Babu KM (2015) A review of the toxicologic implications of obesity. J Med Toxicol 11:342–354

Special Populations: Renal Impairment 40

Gerard Sanderink and Andreas Kovar

Contents

Abstract

Kidney and liver are the main organs involved in the elimination of drugs. In general, the elimination capacity of the kidney is lower than of the liver, because of the smaller organ size and associated blood flow. Renal excretion can be limited by glomerular filtration rate in case of passive excretion or transporter capacity by total renal blood flow and in case of active secretion. Impaired renal function is a rather common condition in patients. Therefore, both dedicated studies in patients with renal impairment and pharmacokinetic investigations via means of population pharmacokinetics are routinely used in drug development to investigate if a dose adjustment needs to be applied in this vulnerable population.

Purpose and Rationale

Kidney and liver are the main organs involved in the elimination of drugs. Both have a metabolic and a direct excretory capacity, although the first is predominant for drugs eliminated by the liver, while the most frequent mechanism of renal clearance is direct excretion of the unchanged drug or its circulating metabolites. In general, the elimination capacity of the kidney is lower than of the

G. Sanderink (✉)
Translational Medicine and Early Development, Sanofi-Aventis R&D, Vitry-sur-Seine, France
e-mail: Gerard.Sanderink@sanofi-aventis.com

A. Kovar
Translational Medicine and Early Development, Sanofi, Frankfurt, Germany
e-mail: andreas.kovar@sanofi.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_8

liver, because of the smaller organ size and associated blood flow. Renal excretion can be limited by glomerular filtration rate in case of passive excretion or transporter capacity by total renal blood flow and in case of active secretion.

Impaired renal function is a rather common condition, with an estimated 30 million people in the USA having chronic kidney damage (CDC 2018). Kidney function is well known to decrease with age and in our experience mild or moderate renal impairment is the norm in study populations above 75 years.

It is therefore necessary to evaluate the potential impact of changes in renal function on drug pharmacokinetics in many cases, and not only for prescribing information including dose adjustment, but also in order to conduct a clinical development program in large patient populations with adequate exclusion criteria guaranteeing safe drug administration.

The main rationale to evaluate the impact of renal impairment is when a drug or its active metabolites are mainly eliminated by renal excretion, especially when a drug is intended to be given in patients with decreased renal function (e.g., elderly patients) or augmented renal function (e.g., critically ill patients) and exhibits a narrow therapeutic margin. In such situations, dose adjustment may be warranted to reduce the risk of adverse drug reactions/toxicity and therapy failure, respectively. However, impaired renal function can also affect nonrenal drug elimination and has also been associated with other changes, such as changes in absorption, transport, and tissue distribution. Also plasma protein binding might play a role through decreased drug protein binding due to low albumin levels. This is of relevance for drugs with high plasma protein binding (>90%) and that have a high hepatic extraction ratio (>0.7). Another indirect mechanism of impaired renal function is uremic plasma that inhibits enzyme or transporter activity.

For most drugs, the evaluation will focus on the effect of decreased glomerular filtration, but it should be kept in mind that renal (drug) transporters, expressed in the basolateral and apical membrane of renal proximal tubules, play an important role in tubular secretion and reabsorption of drug molecules in the kidney as well. Thus, in case of active renal secretion of a transporter substrate, there is also a potential for clinical significant drug-drug interactions with perpetrators that inhibit those transporters like cimetidine (OCT) or probenecid (OAT and OATP). Finally, the possibility that a drug is eliminated during hemodialysis needs also to be considered.

As a consequence, for most drugs that are likely to be administered to patients with renal impairment – including drugs that are not primarily excreted by the kidney – the respective pharmacokinetics should be assessed in patients with renal impairment to provide appropriate dosing recommendations.

Both guidance documents on renal impairment of the FDA and the EMA (US FDA Guidance for Industry 2010; EMA 2015) outline several approaches to study the effect of renal impairment on drug exposure. One approach is a population PK analysis of sparse data in large scale clinical trials, which allows to compare patients with reduced renal function with the typical patient for a given indication. Some limitations of this approach are that it does not address patients that are voluntary excluded from such studies, which is generally the case for severe and end-stage renal impairment. Also some specific parameters like unbound drug fraction and circulating metabolites may need to be included in the evaluation, and the sample size should provide sufficient sensitivity. However, the approach may be very useful to confirm the absence of a treatment risk for patients with renal disease, especially when it is not easily feasible to conduct a specific study in renally impaired subjects without the clinical indication.

When a specific pharmacokinetic study in subjects with decreased renal function is conducted, it can be done according to a “full” design or according to a “reduced” or “staged” design. In the first case, all degrees of renal impairment are included in the study. In the reduced design, the effect of severely decreased renal function in comparison to normal renal function is investigated first. If the results indicate that also other degrees of renal impairment may alter the pharmacokinetics of the study drug to a clinically relevant extent,

the study should be expanded (staged design). This reduced design approach is considered to be useful only if no relevant effect is expected.

For a full-range study in subjects with renal impairment, participants are classified according to their glomerular filtration rate (GFR) which is considered the best indicator of overall kidney function. The current classification is as follows:

Normal/control: GFR $\geq$ 90 ml/min
Mild impairment: 60−<90 ml/min
Moderate impairment: 30−<60 ml/min
Severe impairment: 15−29 ml/min, not requiring dialysis
End stage renal disease patients: <15 ml/min, requiring hemodialysis are considered a separate group

Ideally, GFR should be measured with a marker that has no other elimination pathways (tubular secretion, extrarenal), does not undergo reabsorption, and can easily be measured. The gold standards are exogenously administered substances as a filtration marker (e.g., inulin, iothalamate, iohexol, 51Cr-EDTA, or 99mTc-DTPA) with quantification analysis in both serum and urine. However, in general, the endogenous marker creatinine is used, and in routine clinical practice mostly by estimation of the GFR calculated from the serum creatinine concentration (SCr). SCr with a molecular weight of about 113 Dalton is an amino acid derivative which is endogenously generated from the breakdown of creatine in muscle, distributed throughout total body water, and excreted by the kidneys primarily by glomerular filtration.

Two methods are established in PK studies and commonly used for estimating the renal function:

(a) The estimated creatinine clearance (Clcr) calculated by the Cockcroft-Gault equation (Cockcroft and Gault 1976). It is based on a correlation analysis between SCr and actual creatinine clearance measured from serum and urine creatinine concentrations. This correction is necessary because SCr depends not only on renal excretion but also on muscle mass and food intake.

$$\mathrm{GFR\,(ml/min)} = (140 - \mathrm{age}) \times \mathrm{weight}/(72 \times \mathrm{SCr}) \quad (\times 0.85 \text{ if female})$$

(b) The estimated glomerular filtration rate (eGFR) from the Modification of Diet in Renal Disease (MDRD) study. It is considered to be a more accurate estimation and is being proposed by the US National Kidney Foundation (2012). The MDRD method was established in a larger population and takes into account additional patient-specific factors. It is standardized on body surface area and several versions of MDRD equations have been published. The example recommended by the Kidney Foundation is also provided in the FDA Guidance document.

$$\begin{aligned}\mathrm{eGFR}\left(\mathrm{ml/min/1.73\ m^2}\right) &= 175 \times (\mathrm{SCr,standardized})^{-1.154} \\ &\times (\mathrm{Age})^{-0.203} \times (0.742 \text{ if female}) \\ &\times (1.212 \text{ if African American}).\end{aligned}$$

Both methods have limitations in patients with low muscle mass or with vegetarian diets, should be used in adults only, and can be biased by the fact that creatinine is also a OATPB1 transporter substrate.

Procedure and Examples

Protocols

The pharmacokinetics of the low-molecular-weight heparin enoxaparin were evaluated in 12 healthy volunteers and 36 patients with mild, moderate, or severe renal impairment (Sanderink et al. 2002). This open-label, multicenter, parallel-group study was conducted at four centers. A total of 48 volunteers were enrolled in the study, 36 of whom had either mild, moderate, or severe renal impairment

according to the FDA and EMA classification (12 in each group) and 12 of whom were healthy volunteers, selected to match the overall age, weight, and gender distribution of the subjects with renal impairment. Subjects were initially selected based on their creatinine clearance calculated by the Cockcroft-Gault formula, which was subsequently established more accurately during the study by collecting 24 h urine on day 4. Serum and urine creatinine levels were assessed and creatinine clearance was estimated using the following formula:

$$\frac{\text{Urine creatinine (mg/dl)} \times \text{24 hours urine volume (ml)}/1440\ \text{min}}{\sum[\text{serum creatinine from days} - 1, 4, \text{and } 5\ \text{(mg/dl)}]/3}$$

In addition, eligible subjects were required to meet the following inclusion criteria: to be 18–75 years old, to have a body mass index (BMI) of 18–30 kg/m^2, to have laboratory coagulation parameters within normal ranges.

Enoxaparin was administered once daily by subcutaneous injections at a dose of 40 mg for 4 days when the drug plasma exposure was anticipated to be at steady state.

Evaluation

Venous blood samples were taken over a 5-day period to determine anti-factor Xa and anti-factor IIa activity, and the activated partial thromboplastin time. The schedule for blood sampling was as follows: Day 1: predose, and 30 min, 1, 1.5, 2, 3, 4, 6, 9, 12, and 16 h postdose; Day 2: predose (i.e., 24 h postdose 1) and 3 h postdose 2; Day 3: predose and 3 h postdose 3; Day 4: predose, and 30 min, 1, 1.5, 2, 3, 4, 6, 9, 12, and 16 h postdose; Day 5: 24, 30, and 36 h after dose 4.

Pharmacokinetic analysis was performed by a noncompartmental approach with a full evaluation on Day 1 after the first dose and on Day 4; the maximum observed activity (Amax); the time of maximum observed activity (Tmax); the area under the plasma activity-time curve extrapolated to infinity (AUC, only Day 1); the area under the 24-h plasma activity-time curve (AUC(0-24)); the apparent terminal elimination half-life (t1/2λz) and the apparent total body clearance (CL/F).

The differences between healthy volunteers and the three groups of patients with renal impairment were assessed on the logarithmic transformations of Amax, AUC(0-24), AUC, and CL/F for anti-Xa activity; 90% confidence intervals (CI) for log-transformed ratios in pharmacokinetic parameters of patients with renal impairment versus healthy volunteers.

The relationship between total clearance and creatinine clearance was assessed using correlation and regression procedures.

The steady state was established by comparing values on Days 2, 3, 4, and 5 for minimum observed activity (Amin) or Amax.

Critical Analysis

The study design was in agreement with the regulatory guidance at the time of conduct. Since then, the classification of the groups for renal impairment has changed slightly due to guidance revision and guideline replacement, respectively. A repeat-dose regimen was chosen, because previous data had shown that steady-state exposure was not well predicted by single-dose pharmacokinetics. The 4-day duration was longer than necessary to achieve steady state in patients with normal renal function, because it was anticipated that a prolongation of half-life would occur in renal impairment, with a consequently longer time to reach steady state in these subjects. Nevertheless, the pharmacokinetics were also fully evaluated after the first dose in order to provide information for possible dose adjustment regimens. In contrast, the free fraction was not evaluated, because low-molecular-weight heparins act by binding to antithrombin III in plasma and nonspecific binding is known to be low for LMWH. For an application of free drug evaluation in a renal impairment study, see the chapter on

▶ Chap. 42, "Special Populations: Protein Binding Aspects."

Modifications

A full design was implemented for this study, because an effect of renal impairment on enoxaparin was expected. It allowed to select a control group that matched the overall demographic distribution of the patients, which can be more difficult with a staged design, mainly because subjects with severe renal impairment tend to be younger than the other groups. In this study, the mean (+SD) of the severe renal impairment group was 49.2 ± 15.2 years versus 63.9 ± 7.4 and 62.4 ± 13.3 years in the mild and moderate, respectively. The control group was 57.2 ± 14.3 years. Of notice, there was no individual subject-by-subject matching which would have needed 36 healthy volunteers instead of 12. This was given in by the fact that the means of the groups were to be compared and not by a paired test. Interestingly, more healthy volunteers would not have increased the power of the study. The number of subjects of 12 per group was expected to allow detecting clinically meaningful differences. It is generally not necessary to target a bioequivalence approach in special populations and many studies include only 6–8 subjects per group.

The dose regimen was the same in all groups, 40 mg once daily which is the prophylactic regimen, because higher doses of the drug were known to be well tolerated. This would not have been possible with the dose regimen for treatment of deep venous thrombosis (1 mg/kg bid). For an investigational drug, it may sometimes be necessary to give a reduced dose, especially in the severely impaired group, for instance, when the therapeutic margin is low or still unknown, or the effect of renal function difficult to anticipate. In such cases, it is also possible to include the mild, moderate, and severe groups sequentially in order to select an appropriate safe dose for each group based on the finding in the previous group. An advantage can also be that it allows to compare directly the adjusted dose regimen to the reference regimen. But extrapolation to other doses remains problematic in case of nonlinear pharmacokinetics. For a drug with dose- and time-independent pharmacokinetics, a single dose administration can be selected. An accurate evaluation of elimination half-life and AUC is in that case mandatory.

Results

PK Study

The plasma pharmacokinetic profile on Day 4 of the study described above is shown in Fig. 1 and the associated pharmacokinetic parameters in Table 1.

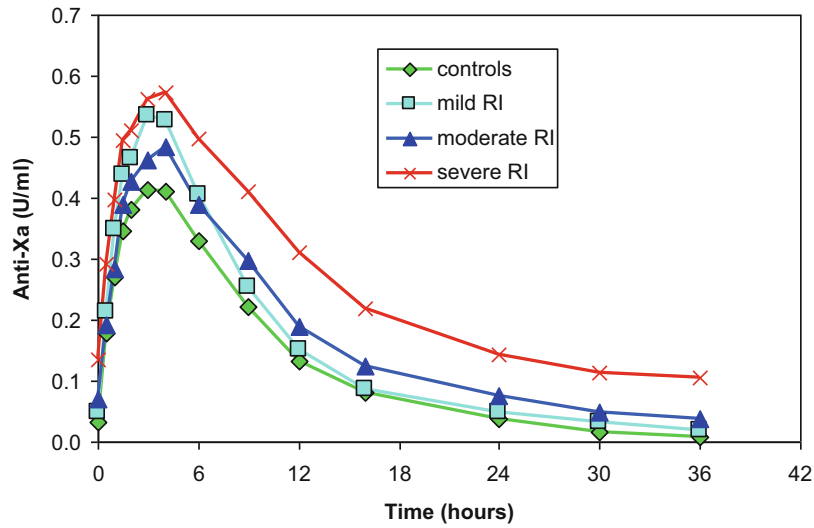


Fig. 1 Enoxaparin plasma anti-Xa pharmacokinetics on Day 4 in renally impaired subjects (RI) and normal subjects (controls) after 40 mg od dosing. (From Sanderink et al. 2002)

Table 1 Enoxaparin plasma anti-Xa pharmacokinetics parameters at steady state in renally impaired and normal subjects after 40 mg od dosing

	Controls ($n = 12$)	Mild RI ($n = 12$)		Moderate RI ($n = 12$)		Severe RI ($n = 12$)	
CrCL (ml/min) Mean ± SD Range	120.7 ± 11.3 84–217	66.4 ± 2.8 51–81		38.5 ± 1.4 30–46		19.3 ± 2.0 5–28	
PK parameter	Mean (CV%)	Mean (CV%)	Ratio[b] (90%CI) p-value	Mean (CV%)	Ratio[b] (90%CI) p-value	Mean (CV%)	Ratio[b] (90%CI) p-value
AUC(0-24) (h.IU/ml)	4.31 (26)	5.20 (32)	120.3 (99.8–145.0) 0.1035	5.53 (22)	120.9 (100.2–146.0) 0.0971	7.88 (36)	165.2 (136.8–199.4) 0.0001
Amax (IU/ml)	0.421 (26)	0.562 (29)	134.6 (111.9–161.9) 0.0088	0.497 (20)	111.7 (92.8–134.6) 0.3244	0.584 (30)	127.3 (105.6–153.3) 0.0341
Amin (IU/ml)[a]	0.036 (<−0.048)	0.046 (<−0.096)	– – 0.0225	0.071 (0.043–0.111)	– – 0.0001	0.132 (0.057–0.245)	– – 0.0001
Tmax (h)[a]	3.0 (2.0–4.0)	3.0 (1.5–4.0)	– – 0.9029	4.0 (2.0–4.0)	– – 0.4373	4.0 (1.5–4.0)	– – 0.3725
t1/2λz (h)[a]	6.87 (3.97–13.2)	9.94 (3.67–20.2)	– – 0.0117	11.3 (5.53–20.0)	– – 0.0011	15.9 (9.66–23.0)	– – 0.0001
CL/F (l/h)	0.98 (25)	0.87 (41)	83.1 (68.7–100.5) 0.1083	0.76 (22)	83.0 (68.5–100.5) 0.1093	0.58 (44)	60.8 (50.2–73.6) 0.0001

[a]median and range
[b]exp.(ln ratio)%
– not calculated
< below quantitation limit
90% CI 90% confidence interval

Rate of absorption was not different among groups, as shown by a similar Tmax (4 h on day 1 and 3–4 h on day 4). Amax was relatively high in the group with mild renal impairment on both Day 1 and Day 4. Because of slower elimination, predose trough values on Day 4 were increased with the degree of renal impairment, and Amax values in RI subjects became significantly increased after repeated dosing.

Exposure on Day 4 was quite well predicted by Day 1 pharmacokinetics in healthy volunteers (+10% accumulation). The accumulation in patients with severe renal impairment was 29%, suggesting some underestimation of AUC and half-life on day 1 in the groups with longer half-lives because of a too short sampling duration (<24 h).

The mean log-transformed anti-Xa AUC(0-24) was 20% and 21% higher in mildly and moderately renally impaired subjects than in normal volunteers. In severe RI patients, mean AUC (0-24) was on average 65% higher on day 4 ($p < 0.001$).

The median apparent elimination half-life increased significantly with the degree of renal impairment (6.87, 9.94, 11.3, 15.9 h).

On Day 1, CL/F was similar between healthy volunteers and subjects with mild or moderate renal impairment (1.00, 0.99, and 0.90 l/h), but there was a statistically significant 27% difference between healthy volunteers and severe RI (0.73 l/h). In the absence of accumulation, apparent plasma clearance was similar on Day 4 and Day 1 in healthy volunteers. In the other groups, estimations of CL/F tended to decrease over time. Differences across groups were therefore more pronounced on Day 4: −17% in mild and moderate RI and up to −39% in severe RI based on mean log-transformed values.

On Day 4, anti-Xa CL/F was linearly although quite poorly correlated with the degree of the renal impairment as depicted in Fig. 2 ($p = 0.0002$, $r^2 = 0.2625$). A more sigmoid relationship might have been more appropriate.

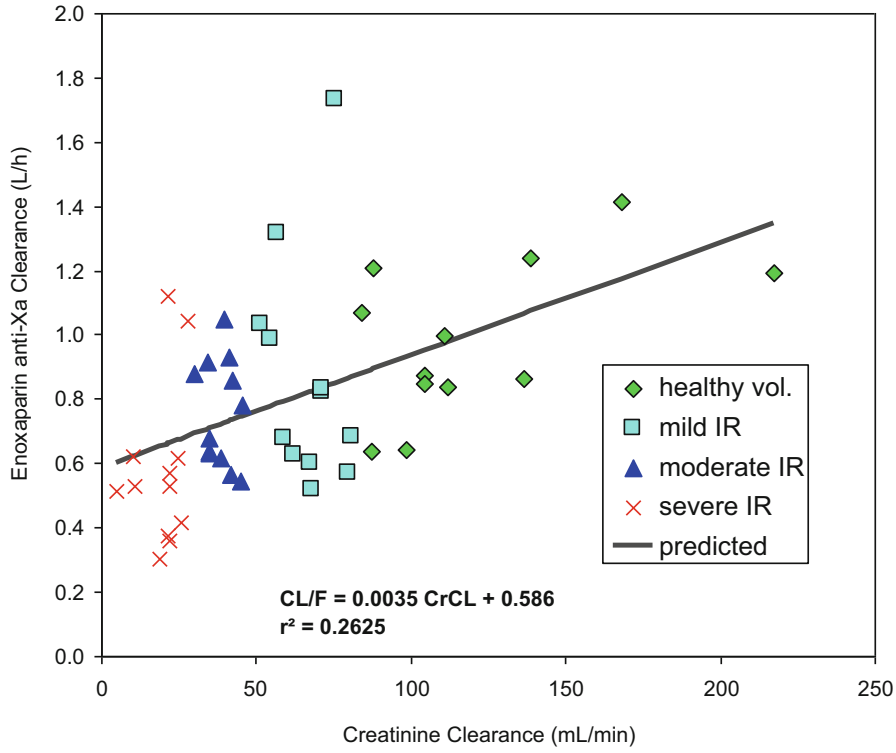


Fig. 2 Correlation between creatinine clearance and enoxaparin anti-Xa plasma clearance. (From Sanderink et al. 2002)

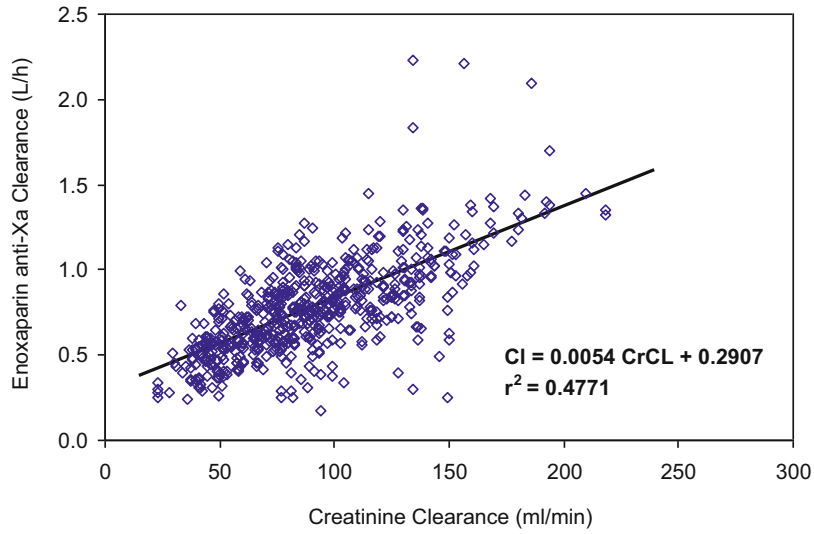


Fig. 3 Effect of renal function on enoxaparin anti-Xa plasma clearance in study TIMI-11A. (From Bruno et al. 2003)

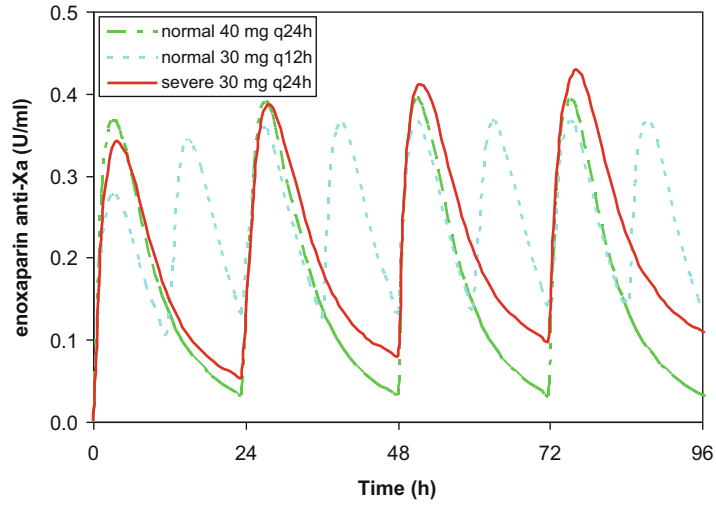


Fig. 4 Comparison of once-daily 30 mg enoxaparin dose regimen in severe renally impaired patients with standard prophylactic regimens in normal patients

Population PK Analysis of Sparse Data

The effect of renal impairment was also evaluated in a population PK study at a higher dose (Bruno et al. 2003). Study TIMI 11A was a multicenter trial of the safety and tolerability of two doses of enoxaparin in patients with unstable angina and non-Q-wave myocardial infarction. The study population consisted of 630 patients with unstable angina or non-Q-wave myocardial infarction. The only exclusion criterion relevant for the analysis was creatinine $\geq$2.0 mg/dL. The median calculated creatinine clearance was 85.7 ml/min (5th–95th percentiles 41.2–152 ml/min). Fifty-one patients had moderate renal impairment (<50 ml/min), but only four severe (<30 ml/min). The effect of renal impairment on the clearance of enoxaparin was consistent with the observations in the PK trial (Fig. 3).

Dose Adjustment

Although the number of subjects in each group was rather low in the pharmacokinetic study, the effect of renal impairment was quite well

estimated as confirmed by the population PK study and the residual variability in each group was low. This justifies the implementation of a model-based dose adjustment for renally impaired subjects rather than individual therapeutic drug monitoring and subsequent dose adaptation. For a low-clearance drug like enoxaparin, a decrease in clearance mainly translates into a prolongation of half-life, so dose adjustment can be accomplished by a reduction in dose or prolongation of dosing interval. In the prophylactic setting and depending on the indication and country, enoxaparin is approved with a 40 mg once daily or a 30 mg twice daily regimen. The data from the PK study were used to simulate different adjusted dose regimens in severe renal impairment (Fig. 4), and it was concluded that 30 mg once daily would lead to early and steady-state exposure parameters (Amax, Amin, and AUC0-24) similar to that in subjects with normal renal function treated with either of one of the standard regimens.

This dose adjustment recommendation is now part of the Lovenox®/Clexane® labelling in several countries, among them the USA (FDA 2007). Similarly, a 1 mg/kg once daily dosing regimen is recommended for treatment of deep vein thrombosis in severe renally impaired patients, instead of 1 mg/kg bid or 1.5 mg/kg od.

References and Further Reading

Bruno R, Baille P, Retout S, Vivier N, Veyrat-Follet C, Sanderink GJ, Becker R, Antman EM (2003) Population pharmacokinetics and pharmacodynamics of enoxaparin in unstable angina and non-ST-segment elevation myocardial infarction. Br J Clin Pharmacol 56:407–414

Centers for Disease Control and Prevention (2018) Chronic kidney disease surveillance system website. http://www.cdc.gov/ckd. Accessed 6 Dec 2018

Cockcroft DW, Gault MH (1976) Prediction of creatinine clearance from serum creatinine. Nephron 16:31–41

EMA/CHMP/83874/2014 (2015) Guideline on the evaluation of the pharmacokinetics of medicinal products in patients with decreased renal function. https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-pharmacokinetics-medicinal-products-patients-decreased-renal-function_en.pdf

Sanderink GJ, Guimart CG, Ozoux ML, Jariwala NU, Shukla UA, Boutouyrie BX (2002) Pharmacokinetics and pharmacodynamics of the prophylactic dose of enoxaparin once daily over 4 days in patients with renal impairment. Thromb Res 105:225–231

US FDA (2007) Lovenox® (enoxaparin sodium injection) for subcutaneous and intravenous use; highlights of prescribing information. http://www.accessdata.fda.gov/drugsatfda_docs/label/2007/022138,020164s075lbl.pdf

US FDA (2010) Guidance for industry: pharmacokinetics in patients with impaired renal function – study design, data analysis, and impact on dosing and labelling. March 2010, revision 1. https://www.fda.gov/downloads/drugs/guidances/ucm204959.pdf

US National Kidney Foundation (2012) KDIGO clinical practice guideline for the evaluation and management of CKD. https://www.kidney.org/content/mdrd-study-equation

Special Populations: Influence of Hepatic Impairment

41

Gerard Sanderink and Andreas Kovar

Contents

Abstract

Liver and kidney are the main organs involved in the elimination of drugs. Both have a metabolic and a direct excretory capacity, while the liver is generally the main organ responsible for drug metabolism and also metabolite excretion. Since a variety of processes is involved in drug elimination by the liver, liver disease can affect the pharmacokinetics of drugs by several mechanisms; for instance, reduced metabolic enzyme activity altered hepatic uptake or biliary excretion by transporters and more generally reduced liver blood flow. Therefore, pharmacokinetic studies are usually performed during drug development if the liver contributes to more than 20% in the elimination of the parent drug or an active metabolite and for any drug with a narrow therapeutic range. A case study is provided with the antihistamine ebastine.

Purpose and Rationale

Liver and kidney are the main organs involved in the elimination of drugs. Both have a metabolic and a direct excretory capacity, while the liver is generally the main organ responsible for drug metabolism and also metabolite excretion. Direct biliary excretion also occurs, but is sometimes a futile pathway because of enterohepatic cycling (reabsorption). Since a variety of processes are involved in drug elimination by the liver,

G. Sanderink (✉)
Translational Medicine and Early Development, Sanofi-Aventis R&D, Vitry-sur-Seine, France
e-mail: Gerard.Sanderink@sanofi.com

A. Kovar
Translational Medicine and Early Development, Sanofi, Frankfurt, Germany
e-mail: andreaskovar@web.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_9

liver disease can affect the pharmacokinetics of drugs by several mechanisms, for instance, reduced metabolic enzyme activity (Frye et al. 2006), altered hepatic uptake or biliary excretion by transporters, and more generally reduced liver blood flow (Verbeeck 2008). Major plasma proteins are synthesized by the liver and thus, drug pharmacokinetics can be affected by decreased plasma protein binding. Liver disease can also affect other organs such as the kidney.

Both the FDA (2003) and EMA (2005) have issued quite similar guidance documents for studies to investigate the effect of hepatic impairment on the pharmacokinetics of medicinal products. For a new drug candidate, the effect should be studied if the drug is likely to be administered in patients with hepatic impairment, if the condition is likely to affect the PK of the drug, and if in that case it would be necessary to proceed with a dose-adjustment in such patients. The FDA recommends a study to be performed if the liver contributes to more than 20% in the elimination of the parent drug or an active metabolite and for any drug with a narrow therapeutic range. An exception is made for drugs intended for single-dose administration.

Moreover, for drugs with a high hepatic extraction ratio (>70%) and high plasma protein binding (>90%), the unbound drug pharmacokinetics should also be evaluated. In some cases, the total drug concentration can mask an effect on the unbound drug concentration (see also ▶ Chap. 42, "Special Populations: Protein Binding Aspects").

In contrast to renal impairment, and maybe because there are several mechanisms potentially involved in the effect of liver dysfunction on drug pharmacokinetics, it has proven difficult to find an adequate marker to classify liver disease with a good predictive value. Both guidance documents recommend the Child-Pugh score as the preferred classification system, acknowledging its limitations and the need for further exploration. The Child-Pugh score is a composite of five parameters: three biochemical (serum albumin, serum bilirubin, prothrombin time) and two clinical (encephalopathy and ascites) (Table 1). The classification is "A" or mild hepatic impairment for a score of five to six points, "B" or moderate for a score of seven to nine, and "C" or severe for ten and more.

Table 1 Child-Pugh score used for classification of liver disease

Parameter	Points		
	1	2	3
Serum albumin (g/dl)	>3.5	2.8–3.5	<2.8
Serum bilirubin (mg/dl)	<2	2–3	>3
Prothrombin time (s > control)	<4	4–6	>6
Encephalopathy (grade)	None	1 or 2	3 or 4
Ascites	Absent	Slight	Moderate

It is important to apply this score only for patients diagnosed with liver disease, because a score of 5 is the minimum value, even in the absence of any abnormalities. On the other hand, some of the parameters can also be changed in other clinical conditions. Even in patients with liver disease, the underlying condition can also affect the pharmacokinetics of drugs differently (e.g., alcoholic cirrhosis, cholestasis, viral hepatitis, liver metastases).

There are different approaches to study the effect of hepatic impairment on drug elimination. One approach is a population PK screen in large-scale clinical trials, which allows to compare the PK in patients with liver disease with the PK of a typical patient population in the proposed indication. Some limitations of this approach are that it does not address patients that are voluntarily excluded from such studies, which is generally the case for severe and sometimes also moderate hepatic impairment. Also, some specific parameters like unbound drug fraction and circulating metabolites may need to be included in the evaluation and the sample size should provide sufficient sensitivity. However, the approach may be very useful to confirm the absence of a risk of liver impairment, or to investigate the effect when it is not feasible to conduct a specific study in subjects without potential clinical benefit, for instance, anticancer drugs (Bruno et al. 1998). In case of a risk, an alternative approach can be a dose-escalation study in patients with liver dysfunction (Raymond et al. 2002). If a population PK approach is being used as part of a submission package for a

new molecular entity, the analysis needs to be prespecified.

For therapeutic proteins, a dedicated pharmacokinetics trial in patients with hepatic impairment is usually not necessary, but population pharmacokinetic analyses have been performed for some of them (Yang et al. 2013).

When a specific hepatic impairment study is conducted, this can be done according to a full design or a reduced design. In the first case, all relevant degrees of hepatic impairment are included in the study. In the reduced design, only the moderately impaired group is initially evaluated in comparison to normal subjects. If necessary, the study can be completed with subjects from the mild and severe groups. In both approaches, the effect of severe hepatic impairment is sometimes not conducted. In such situations, severe hepatic impairment automatically constitutes a contraindication in the labeling. If there is an effect of moderate impairment on the pharmacokinetics of a drug and the mild Child-Pugh category has not been investigated, the findings in the moderate group need to be applied to the mild one as well.

Procedure

Protocol

The antihistamine ebastine is a drug exhibiting a high first-pass effect after oral dosing and the main circulating active metabolite, carebastine, is also partly eliminated by the liver. In addition, the metabolite is >99% bound to plasma proteins in normal subjects.

The effect of hepatic impairment on the pharmacokinetics of ebastine was evaluated in an open-label, multicenter, parallel-group study conducted at three centers (Lasseter et al. 2004). Subjects were classified according to their Child-Pugh score. Indocyanine clearance was also determined as a liver function test. Eight subjects with a Child-Pugh classification A and eight subjects with classification B were treated with 20 mg ebastine once daily for 7 days. For safety reasons, four patients with a Child-Pugh C (severe) status were initially given only one single reduced dose of 10 mg. After evaluation of these patients, the option to treat four more Child Pugh C subjects for 7 days was implemented into the clinical protocol. Twelve healthy volunteers were selected to match the overall age, weight, and gender distribution of the subjects with hepatic impairment.

Evaluation

Venous blood samples for pharmacokinetic profiling of ebastine and carebastine in plasma were taken on day 1 (predose and 0.5, 1, 1.5, 2, 3, 5, 8, 12, and 23.5 h postdose) and on day 7 (predose and 0.5, 1, 1.5, 2, 3, 5, 8, 12, 23.5, 48, 72, and 96 h postdose). Predose samples were also taken on days 3, 4, 5, and 6 to determine the steady-state achievement. For Child-Pugh C patients receiving a single dose, the sampling times were predose and 0.5, 1.0, 1.5, 2, 3, 5, 8, 12, 23.5, 48, 72, and 96 h postdose. Samples were analyzed by validated LC-MS/MS methods with lower limits of quantification of 0.05 ng/ml for ebastine and 1 ng/ml for carebastine. For the active metabolite, the unbound fraction was measured at the anticipated Tmax of 5 h using equilibrium dialysis.

Pharmacokinetic analysis was performed by a noncompartmental approach with a full evaluation on Day 1 after the first dose and on Day 7 at steady state; the main parameters were Cmax, Tmax, Cmin, AUCinf (only day 1); AUC0–24, t1/2λz, and the apparent total body clearance (CL/F). For the active metabolite, the unbound AUC0-24 was calculated by multiplying the total AUC0-24 by the unbound fraction at Cmax.

The differences between healthy volunteers and the three groups of patients with hepatic impairment were assessed on ratios and 90% confidence intervals (CI) of log-transformed pharmacokinetic parameters. Given the high variability of ebastine PK (>70%), the predefined no-effect criteria was a < 2-fold increase as compared to the PK parameters of the matched healthy population.

Indocyanine green was administered intravenously on day 2 of the study at 0.5 mg/kg. The relationship between drug exposure and indocyanine green clearance was assessed by a correlation analysis.

Critical Analysis and Procedure

The study design was in agreement with the regulatory guidance regarding the classification of subjects with hepatic impairment. A repeat-dose regimen was chosen, because previous data had shown that steady-state exposure was not well predicted by the PK determined after a single-dose pharmacokinetics. The 7-day duration was longer than necessary to achieve steady state in patients with normal liver function, to account for a potential occurrence of half-life prolongation in patients with hepatic impairment. In order to mitigate any risk, only four subjects with Child-Pugh class C were initially treated at half the therapeutic dose and only for 1 day. Therefore, the full PK was also evaluated on day 1 in the other groups for comparison. The free fraction was only evaluated at Tmax of the active metabolite, because the methods were not sensitive enough for ebastine-free fraction at any time point and metabolite-free fraction at Ctrough.

A full design was implemented for this study, because an effect of hepatic impairment on ebastine or its metabolite was expected due to the high oral clearance for the parent drug and significant extraction ratio and protein binding of the metabolite. In the past, only mild and moderate impairment have been studied for many drugs because the respective drug is unlikely to be used in severe hepatic impairment and/or can be declared as contraindicated. Also, pharmacokinetic studies in this vulnerable patient population can be quite difficult to perform because of the poor overall health status and prognosis. In practice, it often takes also more time to recruit these patients compared to those from other groups.

A limited control group of 12 subjects that matched the overall demographic distribution of the patients was selected, rather than individual subject-by-subject matching which would have resulted in 24 instead. This was justified by the fact that the means of the groups were to be compared, and not the individual data by a paired test. Actually, the first eight control subjects were recruited when 2/3 of the patients had been recruited and then the control group was completed at the end of the study. Healthy volunteers were selected in a way that for each demographic variable (age, weight), there were eventually two controls in each quartile of the distribution and also with a similar gender distribution.

The number of subjects of eight per hepatic impairment group was expected to allow detecting clinically meaningful differences. Targeting a bioequivalence approach in this special population to exclude any effect on the rate and extent of exposure is generally not feasible and does not make sense.

For a drug with dose- and time-independent pharmacokinetics, a single dose administration could have been selected. An accurate evaluation of the elimination half-life and AUCinf is mandatory in such a case.

For a discussion regarding options for plasma protein binding in this kind of studies, see also the ▶ Chap. 42, "Special Populations: Protein Binding Aspects."

Example

PK Study Results

The results of the study described above did not show a significant impact of hepatic impairment on ebastine and its total metabolite pharmacokinetics (Fig. 1).

At steady state after 20 mg once daily dosing, ebastine AUC was on average 86% higher in mild and 51% in moderate hepatically impaired patients, which was considered not clinically relevant given the high interindividual variability. For the active metabolite carebastine, which achieved higher exposure levels than the parent drug, the AUC0-24 in mild and moderate impaired patients represented 84% and 94%, respectively, of that in normal controls, suggesting a similar formation rate and turnover. The free metabolite concentrations were also similar in these two groups, and no effect of the impairment status on free drug exposure was concluded for the mild and moderate impaired patients (Table 2).

In the severe hepatic impairment group, the exposure after a 10 mg single dose was 27% of

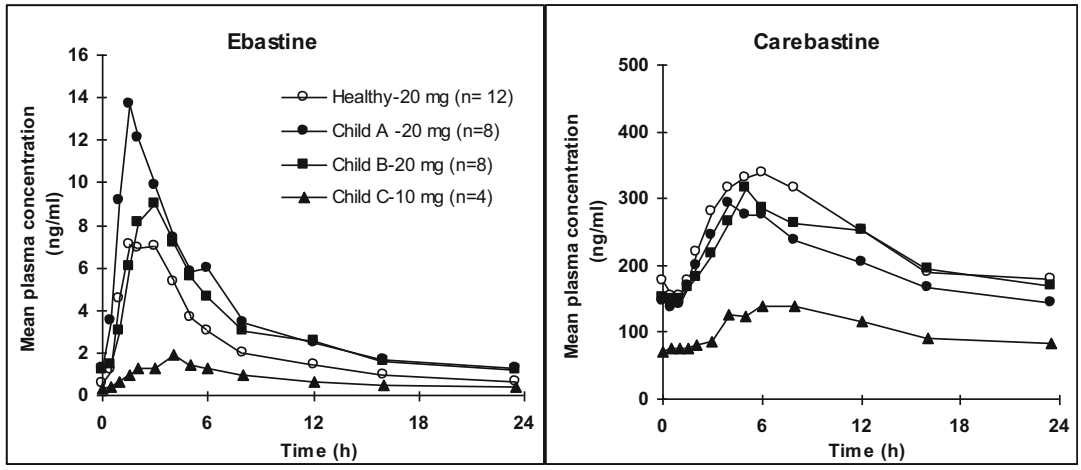


Fig. 1 Mean ebastine and metabolite (carebastine) concentration-time profiles on day 7 after repeated dosing in subjects with hepatic impairment and healthy controls. (Adapted from Lasseter et al. (2004))

Table 2 Ebastine total drug and carebastine total and free metabolite exposure on day 7 after 20 mg once daily dosing (10 mg in Child-Pugh C)

	Ebastine	Carebastine		
	AUC	Total AUC	Free fraction (%)	Free AUC
Healthy volunteers ($n = 12$)	51.9 (65)	5608 (40)	0.54 (97)	30.7 (85)
Child Pugh A ($n = 8$)	89.5 (70)	4661 (40)	0.55 (74)	26.7 (90)
Child Pugh B ($n = 8$)	73.5 (73)	5191 (43)	0.67 (65)	32.1 (64)
Child Pugh C ($n = 4$)	18.2 (64)	2443 (15)	1.84 (66)	46.1 (72)

Mean (CV%). AUC0–24 in ng.h/ml

that of parent drug and 40% for the metabolite. This administration was well tolerated; therefore, four other subjects were treated at the same dose for 1 week. Total parent drug and metabolite exposure at steady state were 37% and 47% of that in healthy volunteers, respectively, close to what would be expected given the lower dose. However, the free metabolite fraction was about threefold higher, and thus the free metabolite AUC was in fact slightly higher than in the control group (Table 2).

Indocyanine clearance decreased with worsening Child-Pugh score, although values were overlapping between the groups, showing that this parameter is driven by different aspects of liver function. However, no negative correlation between ebastine AUC and indocyanine clearance could be established either (Fig. 2).

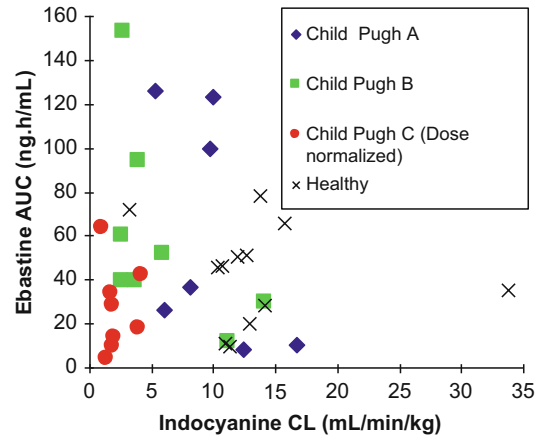


Fig. 2 Correlation between ebastine AUC0-24 on day 7 and indocyanine clearance in subjects with hepatic impairment and healthy controls

References and Further Reading

Bruno R, Hille D, Riva A, Vivier N, Ten Bokkel-Huinink WW, Van-Oosterom AT, Kaye SB, Verweij J, Fossella FV, Valero V, Rigas JR, Seidman AD (1998) Population pharmacokinetics/pharmacodynamics of docetaxel in phase II studies in patients with cancer. J Clin Oncol 16:187–196

EMA CPMP/EWP/2339/02 (2005) Guideline on the evaluation of the pharmacokinetics of medicinal products in patients with impaired hepatic function. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-pharmacokinetics-medicinal-products-patients-impaired-hepatic-function_en.pdf

FDA (USA) (2003) Guidance for industry: pharmacokinetics in patients with impaired hepatic function: study design, data analysis, and impact on dosing and labeling. May 2003, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pharmacokinetics-patients-impaired-hepatic-function-study-design-data-analysis-and-impact-dosing-and

Frye RF, Zgheib NK, Matzke GR, Chaves-Gnecco D, Rabinovitz M, Shaikh OS, Branch RA (2006) Liver disease selectively modulates cytochrome P450-mediated metabolism. Clin Pharmacol Ther 80: 235–245

Lasseter KC, Dilzer SC, Vargas R, Waldman S, Noveck RJ (2004) Pharmacokinetics and safety of ebastine in patients with impaired hepatic function compared with healthy volunteers: a phase I open-label study. Clin Pharmacokinet 43(2):121–129

Raymond E, Boige V, Faivre S, Sanderink GJ, Rixe O, Vernillet L, Jacques C, Gatineau M, Ducreux M, Armand JP (2002) Dosage adjustment and pharmacokinetic profile of irinotecan in cancer patients with hepatic dysfunction. J Clin Oncol 20:4303–4312

Verbeeck RK (2008) Pharmacokinetics and dosage adjustment in patients with hepatic dysfunction. Eur J Clin Pharmacol 64:1147–1161

Yang J, Shord S, Zhao H, Men Y, Rahman A (2013) Are hepatic impairment studies necessary for therapeutic proteins? Clin Ther 35(9):1444–1451

Special Populations: Protein Binding Aspects

42

Italo Poggesi

Contents

Abstract

The aim of this chapter is to provide a basic understanding of the effect of protein binding and its alterations on the pharmacokinetics of drugs and their pharmacological and clinical effects. This has been matter of controversies in the last decades: this chapter calls for scientific comprehension of the underlying phenomena and recommends a reasoned experimental program aiming to the characterization of the protein binding of new compounds and the effects of its changes in different conditions. Considerations and data are also reported concerning the potential changes of drugs' protein binding in some pathophysiological conditions.

Introduction

Background

In addition to proteins specifically acting as transport proteins (e.g., sex hormone binding protein, cortisol-binding globulin), many proteins can bind ligands, such as endogenous compounds or xenobiotics. Among the most important proteins able to bind drugs, there are albumin, α_1-acid glycoprotein (AAG), and lipoproteins (Hervé et al. 1994). Human albumin is responsible for the binding of drugs with different physical-chemical characteristics, but it has

I. Poggesi (✉)
Global Clinical Pharmacology, Quantitative Sciences/Model-Based Drug Development, Cologno Monzese, Italy
e-mail: ipoggesi@its.jnj.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_10

affinity especially for acidic drugs. It is characterized by a molecular weight of 67 kDa and accounts for more than one half of the plasma proteins in humans (normal values for albumin concentration are 3.5–5.5 g/dL, i.e., approximately 500–750 μM); therefore, it provides the higher binding capacity for ligands (Mehvar 2005; Zhang 2012). AAG, with molecular weight 42 kDa, binds preferentially basic or neutral drugs and it is present in human plasma at much lower concentrations (1–3% of total plasma proteins; normal values are 0.04–0.1 g/dL, i.e., approximately 9–23 μM) (Mehvar 2005); for this reason, the binding of drugs to this protein can be more easily saturated. Lipoproteins are proteins able to bind basic or neutral drugs characterized by high lipophilicity. They are classified based on their density (high, low, and very low) and characterized by variable molecular weight (≥200,000). Their plasma concentrations are variable, lower than 0.5 g/dL (Mehvar 2005).

All drugs, when they are in the systemic circulation of animals or human subjects, are characterized by some degree of interaction with plasma proteins. The interaction can be negligible (for instance, the case of acethaminophene, atenolol, carboplatin, ethosuximide, metformin, and ribavirin), therefore providing the total drug concentrations essentially in the unbound form (Rowland et al. 2011a; Zhang et al. 2012). Other drugs are avidly bound to plasma proteins, which, in turn, results in very small unbound concentrations of the compound in the systemic circulation (<1% of the total drug concentrations; for instance, amiodarone, diclofenac, isotretinoin, ketoprofen, nabumetone, naproxen, and teniposide) (Rowland et al. 2011a; Zhang et al. 2012).

Kinetics and Dynamics of the Drug-Protein Interactions

The interaction between drugs and protein is occurring typically in milliseconds and a variety of parameters or metrics can be used to describe the extent of binding (Rowland et al. 2011b). The fraction unbound (fu, expressed as such or as %), defined as the ratio between the free and the total (i.e., the sum of free drug and protein-drug complex concentrations) drug concentration:

$$fu = \frac{[\,free\ drug]}{[\,free\ drug] + [protein - drug]}$$

should be considered the preferred metric to be used to describe protein binding. In this ratio, [*free drug*] is the concentration of unbound drug and [*protein − drug*] is the concentration of the protein-drug complex; the complement of fu to unit (fraction bound or protein binding, also expressed as such or as %) is also used.

It should be considered that the interaction between proteins and drugs can be described in the light of the mass balance law: the relationships between free protein and ligand concentrations and the complex protein-ligand concentration can be described as a chemical equilibrium and therefore ruled by a dissociation (or association) constant.

$$protein - drug \underset{\leftarrow}{\overset{\rightarrow}{}} free\ protein + free\ drug$$

$$k_D = \frac{k_{\rightarrow}}{k_{\leftarrow}} = \frac{[\,free\ protein]\cdot[\,free\ drug]}{[\,protein - drug]},$$

where k_D is the dissociation equilibrium constant, which can also be seen as the ratio of the kinetic micro-rate constants, [*free protein*] is the concentration of free protein, and the other terms were defined above. This equation is appropriate in case the protein-drug complex has a 1:1 stoichiometry; other forms of the mass balance equation can be devised depending on the stoichiometry of the protein-drug complex. Based on the above equation, it can be easily understood that fu is not, in principle, invariant with total drug concentrations: it can be considered invariant when conditions far from saturation are realized (i.e., when total drug concentrations are much lower than total protein concentrations). Saturation, which is rarely approached for drugs binding to albumin, can sometimes be achieved when drugs are avidly bound to AAG.

Impact of Altered Unbound Exposure

It is important to notice that protein binding deals with the ability of crossing membranes; the higher the unbound concentrations the higher can be the interaction with the target responsible for drug effects (both clinical efficacy and clinical safety) and the interaction with entities responsible for drug disposition (for instance, the drug-metabolizing enzymes in hepatocytes). In general, however, when unbound drug concentrations account for $\geq$10% of total drug concentrations, the effect of protein binding on the membrane diffusion is relatively unimportant, as there is always a sufficient amount of free drug that can cross membranes and interact with targets (Rolan 1994). While this should be applied to actual measurements of unbound and total drug concentrations, this is often applied to unbound data assessed in separate experiments. Thus, unbound concentrations can be considered equal to the product fu·total drug plasma concentration or when the systemic exposure is expressed in terms of area under the plasma concentration-time curves (AUC), $AUC_{unbound} = fu \cdot AUC_{total}$.

Changes in plasma protein binding are considered important in most cases by many scientists (Musteata 2012; Ascenzi et al. 2014); other groups, not without reason, are instead stating that protein binding changes are clinically meaningful in a minority of specific cases (Rolan 1994; Benet 2002; DeVane 2002). The latter opinion is largely based on physiological considerations related to the disposition of compounds. In this regard, we can assume, in first approximation, that the unbound AUC is responsible for the drug effects.

$$AUC_u = f_u \cdot AUC = f_u \cdot \frac{Dose}{\frac{CL}{F}},$$

where CL is the elimination clearance of the drug and F is the absolute bioavailability of the drug (CL/F is thus the apparent clearance of the oral drug); the other terms are defined previously in the text. Assuming also that drugs are eliminated exclusively via hepatic metabolism, hepatic CL (CL_H~CL) can be accurately described by the well-stirred model (Wilkinson and Shand 1975; Pang and Rowland 1977)

$$CL_H = \frac{Q_H \cdot f_u \cdot CL_{int}}{Q_H + f_u \cdot CL_{int}}$$

where Q_H is the hepatic blood flow, fu is the fraction unbound (here expressed in blood), and CL_{int} is the hepatic intrinsic clearance, which represents the intrinsic ability of the liver to remove the drug from the systemic exposure in the absence of blood flow limitations and binding (Wilkinson and Shand 1975). It can also be assumed that the drug is completely absorbed from the gastrointestinal tract and that there is no metabolism in the gut (i.e., the absolute bioavailability is equal to the hepatic bioavailability, F~F_H). Hepatic bioavailability (i.e., the fraction of dose that escapes the hepatic first pass) can be written as:

$$F_H = 1 - \frac{CL_H}{Q_H} = \frac{Q_H}{Q_H + fu \cdot CL_{int}},$$

where all the terms were already defined above.

For drugs given orally, rearranging the above relationships, it is easy to demonstrate that the unbound exposure is independent of fu:

$$AUCu = fu \cdot \frac{Dose}{\frac{Q_H \cdot f_u \cdot CL_{int}}{Q_H + f_u \cdot CL_{int}} \Big/ \frac{Q_H}{Q_H + fu \cdot CL_{int}}} = \frac{Dose}{CL_{int}}$$

For drugs given intravenously, F = 1 by definition, so that the following relationship can be written:

$$AUCu = fu \cdot \frac{Dose}{\frac{Q_H \cdot f_u \cdot CL_{int}}{Q_H + f_u \cdot CL_{int}}}$$

In this case, the outcome is different for drugs characterized by high hepatic extraction (i.e., drugs for which fu.CLH> > QH) versus low

hepatic extraction (i.e., drugs for which fu. CLH $<<$ Q). In the first case, Q_H in the $Q_H + f_u \cdot CL_{int}$ term can be considered negligible, so that the terms fu.CL_{int} can be simplified and the unbound exposure is effectively dependent on fu:

$$AUC_u = fu \cdot \frac{Dose}{Q_H}$$

Vice versa, in the second case, the term fu. CL_{int} can be considered negligible, so that Q_H and fu can be simplified and the unbound exposure is effectively not dependent on fu:

$$AUC_u = \frac{Dose}{CL_{int}}$$

Therefore, the binding of a drug is relevant in terms of its influence on the clinical effects only in case of intravenous administration of drugs with high extraction ratio. Other cases in which meaningful clinical effects are expected for low hepatic extraction drugs are in case of transient increase in free drug concentrations (Benet 2002; Rolan 1994). Benet (2002) indicates that the above considerations can be extended to drugs characterized by nonhepatic clearance. Also in this case, the effect of protein binding is clinically relevant in case of drugs with high extraction ratio given intravenously. The same consideration should hold true for drugs with high nonhepatic extraction ratio given orally, but the author claimed that no drugs from a consulted list of 452 drugs, eventually met these criteria.

With respect to the volume of distribution (V, assuming expressed in terms of blood concentrations), it can be written (Rowland et al. 2011c):

$$V = V_B + \frac{f_u}{f_T} \cdot V_{TW},$$

where V_B is the volume of blood, f_u is the fraction unbound in blood, f_T is the fraction unbound in tissues, and V_{TW} is the aqueous volume outside the blood. For drugs with sufficiently large V, V_B can be neglected and V becomes proportional to f_u. Based on the fact that the half-life ($t_{1/2}$) can be expressed as:

$$t_{1/2} = \frac{\ln(2) \cdot V}{CL}$$

the considerations done above for CL can be further extended.

To some extent, however, these are theoretical pharmacokinetic considerations, sometimes valid for asymptotic situations, supported more by the lack of observed clinical effects. A general recommendation is that the pharmacokinetics of new candidate drugs, especially those with high plasma protein binding, should always be performed measuring both total and unbound species in the clinical pharmacology trials, to increase the clinical PK understanding of the new molecule.

Change of Protein Concentrations in Special Populations

In the following section, information on the protein concentration and protein binding in different special populations will be given. The relevancy of such changes should though be assessed in the light of the above-mentioned considerations.

Pediatric and Elderly Subjects

Neonates and infants tend to have reduced concentration of plasma proteins compared to adult subjects (Lu and Rosenbaum 2014). Serum albumin concentrations at birth are approximately 75%–80% of those measured in adults; AAG concentrations, instead, account for approximately one-half of the concentrations in adults (McNamara and Alcorn 2002). Protein concentrations eventually achieve adult levels at 1–3 years of age. In addition to changes in plasma protein concentrations, fetal albumin is characterized by a lower binding affinity to drugs compared to human adults, indicating that also the ability of albumin to bind drugs may be lower in pediatric subjects (Lu and Rosenbaum 2014). Because of these finding, the fraction unbound of drugs are typically higher in pediatric subjects than in adults and protein binding is, therefore, reduced. A paper

(McNamara and Alcorn 2002) reports fu data for a list of 39 drugs bound to either albumin or AAG. The average fraction unbound in adults of the drugs reported in this dataset is lower (approximately 67%) compared to the average of those observed in neonates at term. The same paper reports an interesting computational model, based on simple mass balance equation concepts, that was able to predict with reasonable accuracy fu in plasma of pediatric subjects based on the fu obtained in adults and the relative concentrations of proteins (P) in pediatric and adult subjects:

$$fu_{pediatric} = \frac{1}{1 + \frac{P_{pediatric}}{P_{adult}} \cdot \frac{(1 - fu_{adult})}{fu_{adult}}}$$

This approach could be easily extended also to the other special populations.

In the aging subjects, albumin concentrations are slowly decreasing with age. Albumin concentrations in an 80-year-old subjects are approximately 20% lower than those observed in 20-year subjects (Wallace and Verbeeck 1987). Applying the above relationship indicated by (McNamara and Alcorn 2002) to a drug with fraction unbound of 0.10 in adults, the fraction unbound in an 80-year subjects would be predicted to increase approximately to 0.12, i.e., likely within the inter-subject variability of fu in adult subjects. Although the dependency of AAG is affected by larger variability, it seems that the plasma concentrations of this protein are instead slightly increasing with age. In an extensive review, the protein binding values for elderly were tabulated for more than 50 drugs in elderly subjects (Grandison and Boudinot 2000); selected data from this report are shown in Table 1.

It is interesting to notice that, for diazepam, oral clearance is relatively independent of age; however, this is the result of the counterbalancing of the increased fu with a decreased CLu, likely due to the decreased metabolic clearance in elderly subjects, as reported in Rowland et al. (2011d).

Only approximately 30% of the (Grandison and Boudinot 2000) dataset was composed by drugs with fu <0.1. From the table reported in the original paper, it can be appreciated that more than 50%, approximately 28%, and less than 10% of the compounds showed no relevant changes, an increase and a decrease in the fu values in elderly compared to adults, respectively.

Pregnant Women

Numerous physiological changes occur during pregnancy. Plasma volume may expand up to 50%. The plasma expansion is quicker than the albumin production which leads to physiological hypoalbuminemia due to dilution and thus decreased binding. Plasma albumin concentrations were reported to decrease on average from 44 g/L in the third month of gestation to 32 g/L in the ninth month of gestation (Notarianni 1990). This effect may be further enhanced by the decreased binding capacity due to the increased circulating concentrations of steroid and hormones. As indicated in the introduction, however, the increased free drug concentration is compensated by an increased clearance in many cases (Loebstein et al. 1997).

Some conflicting results have been reported in terms of the effect of pregnancy on AAG concentrations: in some cases, no changes have been reported (Loebstein et al. 1997; Chu et al. 1981), while in other cases, a decrease, though smaller than that reported for albumin, was observed (Notarianni 1990); also in this case, concentrations declined with the month of gestation (from 0.72 g/L in the third month of gestation to 0.50 g/L in the last month of gestation). In case of pregnancy complications (e.g., preeclampsia), AAG was found to significantly increase (Chu et al. 1981).

Hepatic Impairment

Liver disease may decrease protein binding of drugs (or, in turn, increase fraction unbound) via decreased protein concentrations due to reduced synthesis of albumin and AAG and accumulation of endogenous inhibitors which may compete for the binding (Verbeeck 2008). Fraction unbound of individual compounds may show significant increases, whilst no important changes are

Table 1 Fraction unbound (fu) data in adult and elderly subjects for selected drugs (extracted from (Grandison and Boudinot 2000))

Compound	Population	fu	
		Adult	Elderly
Acetazolamide	Healthy	0.041	0.069
Alfentanil	Gastrointestinal surgery	0.093	0.093
Amitriptyline	Healthy	0.052	0.044
Benazeprilat	Healthy	0.083	0.106
Canrenone	Healthy	0.050	0.060
Diazepam	Healthy males Healthy females	0.0125 0.0134	0.0172 0.0166
Diflunisal	Healthy	0.0012	0.0019
Digitoxin	Healthy	0.041	0.045
Flurazepam	Healthy	0.031	0.037
Haloperidol	Healthy	0.095	0.085
Ibuprofen (S) Ibuprofen (R)	Healthy	0.0059 0.0039	0.0078 0.0040
Naproxen	Healthy	0.00084–0.0023	0.0017–0.0051
Nitrazepam	Healthy males Healthy females	0.178 0.179	0.189 0.190
Triazolam	Healthy males Healthy females	0.213 0.229	0.247 0.228
Verapamil	Healthy	0.09	0.10

Table 2 Protein binding data in studies in subjects with hepatic impairment

Compound	Fraction	%				Reference
		Normal healthy subjects	Mild hepatic impairment	Moderate hepatic impairment	Severe hepatic impairment	
Dabigatran	Bound	28.8 ± 1.55	34.5 ± 3.65[a]			(Stangier et al. 2008)
Everolimus	Bound	73.8 ± 3.6	73.5 ± 2.4			(Kovarik et al. 2001)
Febuxostat	Unbound	0.7	0.7	0.6		(Khosravan et al. 2006)
Lopinavir	Unbound	0.69 ± 0.06	0.89 ± 0.21	0.94 ± 0.10	0.91 ± 0.16	(Peng et al. 2006)
Sildenafil	Unbound	3.46 ± 0.61	3.70 ± 1.34			(Muirhead et al. 2002)
Tiagabine	Unbound	3.59 ± 0.74	3.11 ± 0.32	5.13 ± 0.89		(Lau et al. 1997)
Ziprasidone	Bound	99.92 ± 0.03	99.84 ± 0.11	99.91 ± 0.04		(Everson et al. 2000)

[a]Different from normal healthy subjects, but similar to the value of 35% obtained in other phase 1 clinical trials in healthy volunteers and reported in the label

When available, data are reported as mean ± SD

reported for others. Data for selected drugs are summarized in Table 2.

Renal Impairment

Protein binding to albumin is often decreased in patients suffering from an impaired renal function. As in the case of subjects with liver impairment, hypoalbuminemia, accumulation of endogenous substances competing for the binding sites or conformational alteration of the albumin molecule may be the leading cause of the changes of protein binding. Binding to AAG may instead increase due to the higher α1-acid glycoprotein concentration observed in certain categories of patients (renal transplant, patients undergoing hemodialysis). Unbound fraction data for selected drugs are summarized in Table 3.

Table 3 Fraction unbound (fu) data in subjects with hepatic impairment

Compound	Fraction	%				Reference
		Normal subjects	Mild hepatic impairment	Moderate hepatic impairment	Severe hepatic impairment	
Diazepam	Unbound	4.2 ± 0.5 4.6 ± 0.8	4.4 ± 0.7		5.4 ± 1.1[a]	(Viani et al. 1992)
Digitoxin	Unbound	4.6 ± 0.7 5.2 ± 1.0	4.7 ± 0.8		5.9 ± 0.7[a]	(Viani et al. 1992)
Salycilic acid	Unbound	26.0 ± 5.4 34.2 ± 4.6	29.3 ± 5.1		36.3 ± 5.1[a]	(Viani et al. 1992)
Sildenafil	Unbound	2.7 ± 0.8	2.4 ± 0.7	2.0 ± 0.5	2.2 ± 0.3	(Muirhead et al. 2002)
Valproic acid	Unbound	8.4 ± 2.5	20.3 ± 4.7			(Gugler and Mueller 1978)

[a]Subjects undergoing emodialysis

mean ± SD

Oncology Patients

Oncology patients typically have decreased albumin and increased AAG plasma concentrations. Genentech scientists initiated the effort of building a database of the demographic and physiological characteristics of oncology patient to be used in the development of physiology-based pharmacokinetic modeling (Cheeti et al. 2013). In this effort, they also characterized the distribution of plasma proteins in a population of oncology patients (N = 2597) from phase 1, 2, and 3 trials. The distribution of albumin plasma concentrations in their database had a peak at 38 and 41 g/L in male and female subjects, respectively (with values ranging from 3 to 50 g/L), while the typical value observed in healthy subjects is approximately 45 g/L. Likely due to potential inflammation processes in this patient population, the AAG concentrations in the oncology patients of their database were higher than in healthy subjects. The AAG plasma concentration distribution peaked at 1.3 g/L (range 0.3–4 g/L) compared to a typical value of 0.8 g/L reported in healthy subjects.

Conclusions

The effect of protein binding and of its alterations on the pharmacokinetics of drugs has been the matter of numerous controversies in the last decades. Based on the understanding of the pharmacokinetics of drugs, it seems likely that this is relevant only in a relatively restricted number of cases, for instance, for drugs with fraction unbound <0.1 and characterized by high clearance. However, a scientific approach should always be adopted, which requires: (i) to assess the pharmacokinetics of total and unbound concentrations of a new chemical entity and, possibly, avoid reference to fu measurements obtained in separate (ex-vivo) experiments, (ii) consider unbound concentrations at the receptor site to assess the potential effect of a change of unbound fraction on the drug effects, and (iii) always exert clinical judgment, especially for the effects that can be observed in specific group of patients requiring special attentions, such as subjects with severe organ impairment, infants, frail elderly, critically ill subjects, etc. (Roberts et al. 2013; Schmidt et al. 2010).

References and Further Reading

Ascenzi P, Fanali G, Fasano M, Pallottini V, Trezza V (2014) Clinical relevance of drug binding to plasma proteins. J Mol Struct 1077:4–13

Benet LZ (2002) Changes in plasma protein binding have little clinical relevance. Clin Pharmacol Ther 71(3): 115–121

Cheeti S, Budha NR, Rajan S, Dresser MJ, Jin JY (2013) A physiologically based pharmacokinetic (PBPK) approach to evaluate pharmacokinetics in patients with cancer. Biopharm Drug Dispos 34:141–54

Chu CYT, Singla VP, Wang HP, Sweet B, Lai LTY (1981) Plasma α1-acid glycoprotein levels in pregnancy. Clin Chim Acta 112:235–240

DeVane L (2002) Clinical significance of drug binding, protei binding, and binding displacement drug interactions. Psychopharmacol Bull 36:5–21

Everson G, Lasseter KC, Anderson KE, Bauer LA, Carithens RL, Wilner KD, Johnson A, Anziano RJ, Smolarek TA, Turncliff RZ (2000) The pharmacokinetics of ziprasidone in subjects with normal and impaired hepatic function. Br J Clin Pharmacol 49:21–26

Grandison MK, Boudinot F (2000) D. Age-related changes in protein binding of drugs. Clin Pharmacokinet 38:271–290

Gugler R, Mueller G (1978) Plasma protein binding of valproic acid in healthy subjects and in patients with renal disease. Br J Clin Pharmacol 5:441–446

Hervé F, Urien S, Albengres E, Duché JC, Tillement JP (1994) Drug binding in plasma. Clin Pharmacokinet 26:44–58

Khosravan R, Grabowski BA, Mayer MD, Wu JT, Joseph-Ridge N, Vernillet L (2006) The effect of mild and moderate hepatic impairment on pharmacokinetics, pharmacodynamics, and safety of febuxostat, a novel nonpurine selective inhibitor of xanthine oxidase. J Clin Pharmacol 46:88–102

Kovarik JM, Sabia HD, Figueiredo J, Zimmermann H, Reynolds C, Dilzer SC, Lasseter K, Rordorf C (2001) Influence of hepatic impairment on everolimus pharmacokinetics: implications for dose adjustment. Clin Pharmacol Ther 70:425–30, 2001

Lau AH, Gustavson LE, Sperelakis R, Lam NP, El-Shourbagy T, Qian JX, Layden T (1997) Pharmacokinetics and safety of tiagabine in subjects with various degrees of hepatic function. Epilepsia 38:445–451

Loebstein R, Lalkin A, Koren G (1997) Pharmacokinetic changes during pregnancy and their clinical relevance. Clin Pharmacokinet 33:328–343

Lu H, Rosenbaum S (2014) Developmental pharmacokinetics in pediatric population. J Pediatr Pharmacol Ther 19:262–276

McNamara PJ, Alcorn J (2002) Protein binding predictions in infants. AAPS PharmSci 4:19–26

Mehvar R (2005) Role of protein binding in pharmacokinetics. Am J Pharm Educ 69:103

Muirhead GJ, Wilner K, Colburn W, Haug-Pihale G, Rouviex B (2002) The effects of age and renal and hepatic impairment on the pharmacokinetics of sildenafil. Br J Clin Pharmacol 53:21S–30S

Musteata FM (2012) Calculation of normalized drug concentrations in the presence of altered plasma protein binding. Clin Pharm 51:55–68

Notarianni LJ (1990) Plasma protein binding of drugs in pregnancy and in neonates. Clin Pharmacokinet 18:20–36

Pang KS, Rowland M (1977) Hepatic clearance of drugs. II. Experimental evidence for acceptance of the "well-stirred" model over the "parallel tube" model using lidocaine in the perfused rat liver in situ preparation. J Pharmacokinet Biopharm 5:655–680

Peng JZ, Pulido F, Causemaker SJ, Li J, Lorenzo A, Cepeda C, Cabanillas JA, DaSilva B, Brun SC, Arribas J (2006) Pharmacokinetics of lopinavir/ritonavir in HIV/hepatitis C virus–coinfected subjects with hepatic impairment. J Clin Pharmacol 46:265–274

Roberts JA, Pea F, Lipman J (2013) The clinical relevance of plasma protein binding changes. Clin Pharmacokinet 52:1–8

Rolan PE (1994) Plasma protein binding displacement interactions – why are they still regarded as clinically important? Br J Clin Pharmacol 37:125–128

Rowland M, Tozer TN, Derendorf H, Hochhaus G (2011a) Clinical pharmacokinetics and pharmacodynamics: concepts and applications, 4th edn. Wolters Kluwer Health/Lippincott William & Wilkins, Philadelphia, p 98

Rowland M, Tozer TN, Derendorf H, Hochhaus G (2011b) Clinical pharmacokinetics and pharmacodynamics: concepts and applications, 4th edn. Wolters Kluwer Health/Lippincott William & Wilkins, Philadelphia, p 75

Rowland M, Tozer TN, Derendorf H, Hochhaus G (2011c) Clinical pharmacokinetics and pharmacodynamics: concepts and applications, 4th edn. Wolters Kluwer Health/Lippincott William & Wilkins, Philadelphia, p 100

Rowland M, Tozer TN, Derendorf H, Hochhaus G (2011d) Clinical pharmacokinetics and pharmacodynamics: concepts and applications, 4th edn. Wolters Kluwer Health/Lippincott William & Wilkins, Philadelphia, p 388

Schmidt S, Gonzalez D, Derendorf H (2010) Significance of protein binding in pharmacokinetics and pharmacodynamics. J Pharm Sci 99:1107–1122

Stangier J, Stähle H, Rathgen K, Roth W, Shakeri-Nejad K (2008) Pharmacokinetics and pharmacodynamics of dabigatran etexilate, an oral direct thrombin inhibitor, are not affected by moderate hepatic impairment. J Clin Pharmacol 48:1411–1419

Verbeeck RK (2008) Pharmacokinetics and dosage adjustment in patients with hepatic dysfunction. In: European journal of clinical pharmacology, vol 64, pp 1147–1161

Viani A, Rizzo G, Carrai M, Pacifici GM (1992) The effect of ageing on plasma albumin and plasma protein binding of diazepam, salicylic acid and digitoxin in healthy subjects and patients with renal impairment. Br J Clin Pharmacol 33:299–304

Wallace SM, Verbeeck RK (1987) Plasma protein binding of drugs in the elderly. Clin Pharmacokinet 12:41–72

Wilkinson GR, Shand DG (1975) A physiological approach to hepatic drug clearance. Clin Pharmacol Ther 18:377–390

Zhang F, Xue J, Shao J, Jia L (2012) Compilation of 222 drugs' plasma prtein binding data and guidance for study design. Drug Discov Today 17:475–485

The Human ADME Study

43

Andrew McEwen

Contents

A. McEwen (✉)
Reverisco Solutions, Peterborough, UK
e-mail: andrew.mcewen@revirescosolutions.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_11

Abstract

The human mass balance study is a pivotal study in the drug development process. While a reasonable understanding of the absorption, distribution, metabolism, and excretion (ADME) properties of the candidate drug would have been determined using pre-clinical models, the ultimate validation is provided following administration to human volunteers. The human ADME (hADME) study provides the link between pre-clinical safety studies and the clinical observations. While described as a mass balance study, the key objective of the hADME study is the quantification, characterization, and identification of drug and drug metabolites present in systemic circulation. An assessment of the relative exposure between clinical subjects and the species used for pre-clinical safety studies enables a complete safety profile to be obtained.

Introduction

During the drug development process, considerable effort is expended in obtaining an understanding of the absorption, distribution, metabolism, and elimination (ADME) properties of drugs using, animal data, in vitro models and in silico models. While these models undergo constant revision and change to meet current regulatory requirements, the ultimate validation of these models comes with a pivotal study in the process – the human ADME study (hADME). In this study, all the predictions and measurements are tested using the target species, man. The study design usually covers a range of objectives but is essentially a metabolism study designed to investigate the routes and rates of excretion, pharmacokinetics in blood and plasma and confirm the identity of major circulating metabolites.

The regulatory question that the hADME study is intended to answer is often phrased as "Do you understand the metabolism of your compound in man?" but ultimately the question is better phrased "Do you understand the metabolism of your compound in man and how it relates to the metabolism in species selected for pre-clinical investigations?"

Before considering the modern human mass balance study, it is worth noting that the quantitative basis on which the study is reliant was established around 450 years ago by the work of Sanctorius. Sanctorius was a Venetian professor based at the University of Padua who is credited with numerous inventions including the thermometer. Over a period of 30 years he weighed himself, everything he ate or drank and also his urine and feces. He found that for every 3.6 kg of food he consumed only 1.4 kg of waste product were excreted. The difference was assigned to a process he described as "metabolism," and a new science was established. The current human ADME study follows a similar experimental process; the drug administered is essentially weighed and administered to volunteers. The excreta is collected and

weighed before the concentrations of drug are determined and a mass balance obtained.

The first recorded example of a human ADME study was provided Alexander Ure (1841). The study was inspired by the work of Woehler and Tiedemann (1824) that showed benzoic acid was converted to hippuric acid and eliminated in the urine following administration to dogs. Ure had a particular interest in the treatment of gout, and he proposed that the symptoms of gout may be relieved by reducing the synthesis of uric acid by administration of benzoic acid. He therefore conducted a human metabolism study and was successful in demonstrating the presence of hippuric acid in the urine collected following administration. The first successful human metabolism study may also be the first example of a failed therapeutic use as the administration of benzoic acid failed to reduce the concentrations of uric acid observed in the urine. The potential therapeutic use of benzoic acid in controlling gout was not discussed in the subsequent publication.

Strategies for quantifying drug metabolites has been debated widely, but despite the great advances in modern bioanalytical methods the use of radioisotopes is still an invaluable part of pharmaceutical development programs. The European Bioanalytical Forum presented a white paper outlining the issues associated with LC-MS/MS methods and providing some recommendations (Timmerman et al. 2011). The paper compared of responses obtained using LC-MS/MS with those obtained using conventional radiometric detection and highlighted the issues associated with differential ionization in the mass spectrometer. Examples were provided showing that in some cases there was a 30-fold difference in relative exposure between the two methods. The use of a radiolabeled test compound relies on no assumptions as to the identity of the drug-related compounds present in excreta or plasma and needs no internal standard and the response is structure independent allowing direct quantification.

Radioisotopes commonly used as tracers in metabolism experiments include [3H], [131I], [32P], and [35S] inserted as replacements for the hydrogen, iodine, phosphorus, or sulfur atoms commonly found in drugs; a summary of their radioactive properties is provided in Table 1. The isotope of choice, however, is [14C] due to long half-life, good detection efficiency, and the ability to place the radiolabel in a metabolically stable position within the molecule. Tritium is widely used in preclinical studies but suffers from low counting efficiency, short half-life, and the potential for isotope exchange or metabolic cleavage. Careful selection of the labelling site and adoption of targeted labelling strategies can help mitigate these liabilities (Lockley et al. 2012), but the use of tritium labels in hADME studies is still quite limited.

Table 1 Properties of radionuclides commonly used in hADME studies

Radionuclide	Emission	Half-life (years)	Specific activity
3H	Beta	12.3	28.8 Ci/mmol
14C	Beta	5730	62.4 mCi/mmol
32P	Beta	0.04	5118 mCi/mmol
35S	Beta	0.24	1494 mCi/mmol
133I	Beta	0.16	16382 Ci/mmol

Purpose and Rationale

Human metabolism studies such as those performed by Ure were initially focused on identification of transformation products in excreta (urine and feces), and it remained this way for over 100 years. Improvements in analytical methodologies meant that bioanalytical measurements could be performed using plasma to demonstrate that the preclinical species used in safety testing were exposed to drug material. Most bioanalytical measurements are currently performed using liquid chromatography-mass spectrometric techniques, and for the interested reader, the development of the discipline has been reviewed elsewhere (Hill 2009). As the sensitivity of modern instruments improved, the focus moved from identification of metabolites present in excreta to quantification and identification of the metabolites in circulation. The importance of plasma as the

primary focus for metabolite identification studies was reinforced by the Metabolites in Safety Testing (MIST) guidelines introduced by the FDA in 2008 and the ICH in 2009 in response to a perceived industry need for clarification of the regulatory position.

Determination of the mass balance is included in the study design, but dependent on the PK/ADME properties of the test compound a full mass balance may not be obtained. Of greater significance is an assessment of the relative routes and rates of excretion. Similarly a determination of pharmacokinetic parameters for total radioactivity, parent compound, and metabolites in systemic circulation is a key activity included as part of the study plan. The key measurement here is the comparison of AUC's obtained for total radioactivity and parent. This assesses the metabolite load and combined with radio-chromatographic analysis can validate the findings of the bioanalytical measurements performed in support of the clinical studies performed earlier in development.

The main purpose of the hADME study is therefore to provide a validation of the preclinical species used for safety testing, identify disproportionate or unique human metabolites, and "close the circle" between preclinical and clinical investigations.

Despite (or possibly because of) constant advances in analytical instrumentation with respect to sensitivity and resolution, it can be advantageous to position the timing of the hADME study as close as possible to the preclinical evaluations as possible. This enables preclinical and clinical samples to be analyzed on the same equipment at the same time. A direct comparison of retention times is possible and scouting for "unique" human metabolites in the toxicological species is possible.

The absence of direct regulatory guidance helped promote a strong debate as to which matrix was most important with regard to metabolite identification. In 2002 a multidisciplinary committee sponsored by the Pharmaceutical Research and Manufacturers of America published a report outlining best practice entitled "Drug Metabolites in Safety Testing" (Baillie et al. 2002). The primary trigger for further investigation was suggested as any component that accounted for 25% drug-related material in the systemic circulation. The report received a rapid response from the FDA who challenged the 25% trigger citing recommendations in the veterinary medicine guidelines where 10% of the total drug present was classified as a major metabolite. The draft regulatory "MIST" guidance document was published in 2005 and provided a trigger level of 10% drug-related material as the requirement for further investigation. The proposed guidance significantly increased the dialogue surrounding metabolite safety, and several further publications were generated reflecting opinion from within the industry (Smith and Obach 2005, 2006; Humphreys and Unger 2006; Prueksaritanont et al. 2006; Naito et al. 2007). Of particular interest was the proposition by Smith and Obach (2005) that absolute abundance and dose should be taken into consideration, noting that a metabolite representing 10% AUC obtained following a 1 g dose should be treated differently from a metabolite representing 10% AUC observed following a 1 mg dose due to the fact that the body burden would differ by a factor of 1000.

After considerable debate, the final FDA guidance was published in 2008 which contained a surprising change to the draft document where the trigger level for safety evaluation was set at 10% AUC of parent drug. For drugs that were extensively metabolized, this would mean that almost every component in circulation would require quantification and identification. The FDA position was undermined by the harmonized ICH document published in 2010 by the European Medicines Agency which re-iterated the 10% total drug-related material cut-off for metabolite investigations, and the FDA document was subsequently amended accordingly (FDA 2016). The revised guidance also makes reference to a document published by the EMEA in 2012 and notes this will represent the agencies current opinion on the subject. The definition of disproportionate metabolites was therefore harmonized, and the decision tree as now presented in the FDA guidance is reproduced in Fig. 1.

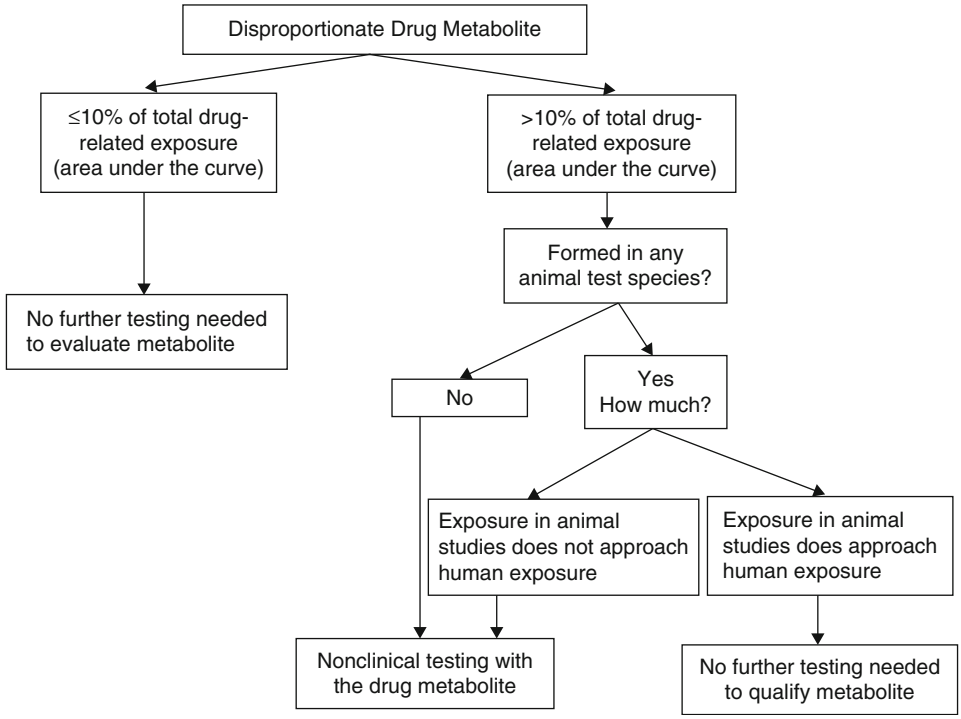


Fig. 1 FDA decision tree (Reproduced from US-FDA 2016)

Dosimetry

The hADME study is usually performed late in the drug development program, and there is generally a reasonable understanding of the chemical safety of the drug. The aim of the dosimetry study is therefore to assess the risk associated with administration of a radiolabeled test compound to human volunteers. The risk must be balanced with the objectives of the study, the aim being to administer sufficient radioactivity to achieve mass balance and fully characterize the proportions and identity of metabolites in systemic circulation.

The preclinical studies required to achieve this include a mass balance evaluation in the rat and a tissue distribution study to assess the relative contribution from each organ to the internal dose. The mass balance study provides information on relative rates and routes of elimination, and therefore, an assessment of the radioactive dose received by the gastrointestinal tract and bladder. The tissue distribution study can be performed using tissue excision and subsequent analysis of the tissues, but this method requires preselection of the tissues of interest prior to analysis and is heavily reliant on the skill of the technician performing the necropsy. The method of choice is now quantitative whole body autoradiography (Ullberg 1954), which relies on no preconceptions, provides information on the intraorgan distribution of radioactivity, and is nondestructive meaning that the tissue sections can be reanalyzed for radioactive content or taken for additional analyses such as tissue staining or MALDI experiments to determine the localization of both drug and metabolites (McEwen, Henson and Wood 2014).

Most tissue distribution studies are performed using the same strain of albino rats as chosen for the toxicology evaluations, thus enabling findings to be correlated with the presence of drug related material. The investigations also tend to run for up to 7 days reflecting the timeline for the mass balance studies. This design has two drawbacks when estimating the dose administered to man. There are no pigmented tissues in which to assess

the binding of the drug to melanin and there may still be tissues with radioactivity present at the last timepoint. From a chemical safety standpoint, there is no indication that melanin binding is in any way associated with toxicity (LeBlanc et al. 1998), but binding of a radioactive drug can provide a significant internal dose to the eye. While the eye is not one of the mandatory tissues for dosimetric calculation as listed by the ICRP, it is therefore prudent to include it in the risk assessment.

The second deficiency in the standard QWBA experiment when calculating the internal dose in man is that the presence of radioactivity in tissues at the final sampling point triggers the use of standard half-life values in the dosimetry calculations. These are generally taken as 100 h in the USA and a more conservative 100 days in the EU. Extending the experiment to 21, 28, or 35 days can demonstrate elimination of radioactivity from the tissues or alternatively provide sufficient data with which to calculate a biological half-life.

A description of the dosimetry calculation has been presented in depth in an earlier chapter in this series (Kuerzel et al. 2011) and so will not be discussed further here. The aim is to assess the risk to human volunteers following administration of a radioactive test substance, and these are related to guidance provided by the WHO (1977) and subsequent modification by the ICRP (1992).

The guidance takes into account the total risk associated with exposure, namely the probability of a fatal cancer, nonweighted probability of a nonfatal cancer, and probability of successive generations suffering serious hereditary disease as a result of administration of the radiolabeled test item. The risk categories are reproduced in Table 2.

The risks run from category I where the risk is one in a million to category III where the risk is one in a thousand. Exposure of volunteers to an effective dose between 0.1 and 1.0 mSv is generally considered acceptable for biomedical investigations.

When preparing the preclinical package, it is worth considering the clinical study when designing the mass balance and tissue distribution studies. The routes and rates of excretion can be taken from albino animals, and if a full balance is not obtained, then a ratio of urinary to fecal excretion can be employed. The data should be generated using the same sex animals and dose route as intended in the hADME study. A list of the tissues specified for dosimetric evaluation is provided in Table 3. It should be noted that the remainder is then provided as a separate list. Tissue distribution studies are generally performed using albino animals, the same strain used in toxicological safety studies, but for assessment of the risk to man it is advisable to use a pigmented strain for measurement of melanin binding. The additional tissues are provided in Table 4. It can be seen that some of

Table 2 Risk categories based on radioactive dose (ICRP)

Risk level	Risk category	Corresponding effective dose range in adults (mSv)	Level of societal benefit
Trivial	Category I	<0.1	Minor
Minor to intermediate	Category IIa	0.1–1.0	Intermediate to moderate
	Category IIb	1.0–10	
Moderate	Category III	>10	Substantial

Table 3 Tissues for dosimetry

Tissue or organ	Weighting factor W_R
Gonads	0.20
Bone marrow (red)	0.12
Colon	0.12
Lung	0.12
Stomach	0.12
Bladder	0.05
Breast	0.05
Liver	0.05
Esophagus	0.05
Thyroid	0.05
Skin	0.01
Bone surface	0.01
Remainder	0.05

Table 4 Additional tissues

Adrenals	Myocardium
Adipose tissue	Pancreas
Blood	Pituitary gland
Brain	Prostate
Eye (uveal tract)	Small intestine
Upper large intestine	Spleen
Kidney	Thymus
Muscle	Uterus

these tissues are quite small and the use of QWBA overcomes the potential contamination issues associated with traditional necropsy methods.

Radiolabeled Test Compound

A key factor in the successful completion of a hADME study is accurate quantification of the dose material. In 2001 the EU passed laws relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use and the application of good manufacturing practice (GMP) was extended to investigational products. The test compound should have a chemical and radiochemical purity greater than 98%, a well-characterized specific activity, and a stability trial should be performed to assure product integrity between synthesis and dosing.

Clinical specific activity (μCi/mg) = Amount of radioactivity (μCi)/Amount of drug (mg)

The drug material administered should be homogeneous and from a single batch. Experiments, for example, where cold material is administered in one capsule and the radioactive dose is administered in a separate capsule can fail due to differential absorption of different crystalline forms of the test material resulting in an absorbed dose with unknown specific activity.

The synthesis of radiolabeled compounds for use in clinical studies is covered in a separate chapter (Atztrodt) and will not be discussed further here. For investigational products administered intravenously, further work may be required prior to administration to assess binding to dosing cannula.

Clinical Study Design

Study Title

A general title for the human ADME study can be expressed as follows:

A phase 1, open label study investigating the absorption, metabolism, and elimination (AME) of 14C-ABC123 following a single oral dose to healthy male subjects.

It should be noted that "distribution" is missing in this title reflecting a difference of opinion between investigators as to whether distribution can be determined as part of the radiolabeled human mass balance study. The study is also open meaning that the identity of test compound administered is known to the investigators performing the analysis.

Objectives

When preparing a study protocol for the human ADME study, the objectives are generally categorized as primary and secondary. As an example the objectives could be listed as follows:

The primary objective of the study is to determine the pharmacokinetics of total radioactivity in plasma, whole blood, and red blood cells along with the pharmacokinetics of ABC123 following a single oral dose of 10 mg ABC123

The secondary objectives are:

- To determine the relative routes and rates of excretion and obtain a mass balance by measuring the urinary and fecal excretion of radioactivity.
- To estimate the protein binding of ABC123-related radioactivity in human plasma.
- To obtain samples for quantification and identification of components present in systemic circulation.
- To determine the safety and tolerability of a single dose of 10 mg 14C labelled ABC123 in healthy volunteers.

The primary objective is therefore to determine the systemic exposure of radioactivity following administration and relate that to the exposure to ABC123. The difference between total radioactivity and parent compound in systemic exposure is described as the metabolite load. If the exposure to total radioactivity and parent compound are the same (i.e., no metabolism), then there is no further work required. If the difference between total radioactivity and parent compound exposure is large, then further work is required to determine the proportions and identity of the radioactive components in addition to parent that are present. These investigations can sometimes be complex and time-consuming. For that reason, the metabolite identification activities are usually conducted and reported separately from the clinical investigations.

Study Design

The study design is generally presented as a single center, open label, single administration of 14C-ABC123 as an oral solution administered to fasted healthy male volunteers. As the study requires the use of a radioactive compound, female volunteers are generally not used in hADME studies. If the intended therapeutic use is for female conditions only, then the use of female volunteers would be required. In this case, the volunteers would be postmenopausal women.

Human Volunteers

The study would be conducted in healthy male subjects aged 30–55 years. Six subjects would be enrolled in the study with the expectation that at least four subjects would complete the study. Replacements are generally not required unless greater than two volunteers drop out. The process includes enrolment of eight subjects to present at clinic for screening on Day 1. Two subjects are eventually sent home once six subjects have successfully been dosed.

Dose Administration

A single oral dose of 10 mg 14C-labelled ABC123 will be given on the morning of Day 1 after an overnight fast. The radioactivity administered will correspond to an effective radiation dose of 0.1 to 1.0 mSv. The dose has been selected based on preclinical and clinical data and is considered safe and suitable for characterization of the pharmacokinetic properties of the drug.

The effective radiation dose is within the limits specified under the ICRP and WHO guidelines for administration of radioactivity to human volunteers.

Selection Criteria

The study plan generally lists specific criteria by which subjects can be enrolled, and these are split into inclusion and exclusion criteria.

Inclusion

Subjects who meet the following criteria can be admitted on the study:

1. The subject is able to read and understand the Volunteer Information sheet.
2. The subject has signed the study specific Informed Consent Form.
3. The subject is male.
4. The subject is between 40 and 55 years of age.
5. The subject has a minimum weight of 60 kg and a BMI between 19 and 29 kg/m^2. If we assume typical sample sizes of 1 g and 0.25 g for determination of radioactive concentrations in urine and feces, respectively, then the detection limits (expressed as % administered dose) are defined by the following graphs.
6. The subject has a resting pulse in the normal range of 51–100 bpm.
7. The subject has a resting systolic blood pressure within the normal range of 91–179 mmHg (supine) and a resting diastolic blood pressure of 51–100 mmHg. The subject should also have an orthostatic blood pressure less than 20 mmHg.

8. The subject is considered in good health based on the results of a prestudy physical investigation, medical history, vital signs, an electrocardiogram and laboratory investigations including blood biochemistry, hematology, serology, and urinalysis within acceptable range.
9. For some compounds, the subject must agree to the use of a suitable barrier method of contraception throughout the study and for 3 months following study completion.

Exclusion

Subjects who meet one or more of the following criteria cannot be admitted on the study:

1. Use of concomitant medicine, specifically any prescribed medicine within 2 weeks of the study, or over the counter (OTC) medication within 1 week prior to dosing. Patients who have taken nonprescribed or topical medication may still be allowed on the study if in the opinion of the investigator, the medication will have no effect on the outcome of the study.
2. Subjects who at the screening visit fail a urinary drugs of abuse test or alcohol breath test.
3. Significant history of drug or alcohol abuse within the last 6 months. For assessment of alcohol consumption, one unit is described as ½ a pint of beer/lager, 1 glass of wine, or 1 shot of spirits.
4. The subject has participated in a clinical trial within 3 months prior to dosing.
5. The subject has a known sensitivity to ABC123.
6. The subject has a history of severe drug allergy or hypersensitivity.
7. The subject has a serious illness, such as renal or liver dysfunction, or a cardiovascular, pulmonary, gastrointestinal, endocrine, neurological, infectious, neoplastic, or metabolic disorder.
8. A history of seizures.
9. The subject has been classified as either a "poor" or "fast" metabolizer for CYPD6.
10. The subject has tested positive for HIV.
11. Blood donation within 3 months prior to dosing.
12. A history of smoking or use of nicotine substitution therapy (patches, gum, or inhalers).
13. Positive test to a screen for drugs of abuse (amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, methadone, or opiates).
14. The presence of clinically relevant cardiovascular disease including signs of arrhythmia/tachycardia of QT prolongation.
15. Receipt of an X-ray (other than dental X-rays) or radiolabeled material within 12 months preceding the study.
16. Occupational exposure to radioactive substances.
17. An irregular defection pattern <1, >3 defecation per day.
18. An assessment that the subject may be unwilling to comply with the clinical study protocol.

Withdrawals and Replacements

Subjects may be withdrawn from a study for a variety of reasons; these include:

1. Withdrawal of consent.
2. The investigator decides that the subject should be withdrawn for safety reasons.
3. A serious adverse event (SAE) occurs.
4. The subject is lost on follow-up.

The date and reason for withdrawal should be noted and, where possible, subjects that withdraw should be seen for a final evaluation and completion of the records.

Subjects that withdraw prior to dose administration can usually be replaced, whereas subjects who withdraw following dose administration are generally not replaced.

Restrictions

The study protocol generally lists a set of restrictions that must be followed by the volunteers

during the conduct of the study. Standard exclusions include:

Alcohol: Subsects are usually asked to abstain from alcohol intake from 48 h prior to dosing until after the follow-up visit.

Caffeine- and/or xanthine-containing products: Subjects should refrain from consuming caffeine and/or xanthine from 48 h prior to dosing until they have completed the study. Typical xanthine-containing products include tea, coffee, cola, chocolate, and chewing gum.

Contraception: As the study involves the use of 14C-labelled material, the subjects will be asked to use a suitable method of contraception (such as a condom) throughout the study and for 3 months following study completion.

Exercise: Subjects are asked to refrain from strenuous exercise from 72 h prior to dosing until after the follow-up visit.

Grapefruit, grapefruit-containing products, marmalade, and Seville oranges: Subjects are asked to refrain from consuming grapefruit, grapefruit-containing products, marmalade, and Seville oranges from 48 h prior to dosing until after the follow-up visit.

Meals: Subjects are provided with a standard diet during their time in the clinic. Subjects usually fast from 10 h prior to dosing and should refrain from taking water from 2 h prior to dosing until 2 h postdose. After dosing, subjects may resume their usual rate of fluid consumption but may be limited to 2 l of water per day.

Smoking: Subjects are not permitted to smoke or use nicotine containing products throughout the study.

Discharge Procedures

The subjects will remain in the unit for a minimum of 120 h postdose (Day 5). After this period, collection of urine and feces will be discontinued and subjects will be released if the following criteria are met:

- The combined cumulative excretion of radioactivity in urine and feces exceeds 90% of the administered dose.
- The total radioactivity in two consecutive 24 h collections is below 1% of the administered dose.

If by Day 5 the release criteria are not met, the subjects can be kept in the unit until the release criteria are met or until Day 14 whichever comes first.

Duration of the Study

Subjects will remain in the clinic until Day 5 or until the release criteria are met. Further samples may be taken up to a maximum of 15 days.

Some study designs treat the subjects as a group and base release criteria upon the mean excretion from the group. This can lead to extended stays for volunteers if one of the groups has difficulty defecating or fails to comply with the collection procedures. An alternative strategy is to base the release criteria on individual excretion. This promotes compliance with the collection process as the subject can be released based on their personal criteria.

Pharmacokinetic Assessments

Blood Samples for Parent Drug Analysis and Total Radioactivity Determination

A series of samples will be taken for determination of total radioactivity and parent drug concentration throughout the study period. The exact sampling times will be recorded in the data.

Blood samples for determination of ABC123 will be drawn at predose and 1, 2, 4, 6, 7, 8, 9, 10, 12, 15, 24, 36, 48, 72, 96, and 120 h postdose.

Blood samples for measurement of radioactivity in plasma, whole blood, and red blood cells will be drawn at predose and 1, 2, 4, 6, 7, 8, 9, 10, 12, 15, 24, 36, 48, 72, 96, and 120 h postdose.

Overall 17 × 2 mL blood samples will be taken for determination of ABC123.

In addition, 17 × 6 mL blood samples will be taken for measurement of total radioactivity in plasma, whole blood, and red blood cells.

Subjects will be cannulated up to 24 h postdose to avoid repeated venipuncture. For each sample taken via cannula, the first 0.5 mL blood is discarded and following blood collection the cannula is flushed with saline.

Blood Samples for Protein Binding and Metabolite Investigations

A series of samples will be taken for determination of plasma protein binding and metabolite investigations throughout the study period.

Blood samples for determination of plasma protein binding will be drawn at predose and 4, 8, 12, and 24 h postdose.

Blood samples for metabolite investigations will be drawn at predose and 4, 8, 12, 24, 48, and 72 h postdose.

Overall 5 × 7 mL blood samples will be taken for determination of plasma protein binding.

In addition 7 × 10 mL blood samples will be taken for metabolite investigations.

Total Volume of Blood Drawn

Based on the sampling regime outlined above, the estimated volume of blood drawn from each subject would be around 280 mL over the course of the study.

The distribution of samples for each subject is as follows:

- 8.1 mL for hematology (3 × 2.7 mL)
- 9.0 mL for blood biochemistry and serology (1 × 9 mL)
- 14 mL for blood biochemistry (2 × 7 mL)
- 34 mL for drug assay (17 × 2 mL)
- 35 mL for determination of plasma protein binding
- 70 mL for metabolite investigations (7 × 10 mL)
- 6 mL for discard during sampling via cannula.

Urine Sampling for Total Radioactivity Determination and Metabolite Investigations

Urine samples will be collected at the following intervals: −24–0, 0–6, 6–12, 12–24, 24–48, 48–72, 72–96, 96–120, 120–144, and 144–168 h postdose. If required further samples can be taken covering the intervals 168–192, 192–216, and 216–240 h postdose. The actual start and end times for each collection will be recorded in the study data along with the weight.

Feces Sampling for Total Radioactivity Determination and Metabolite Investigations

Feces samples will be collected at the following intervals: −24–0, 0–24, 24–48, 48–72, 72–96, 96–120, 120–144, and 144–168 h postdose. If required further samples can be taken covering the intervals 168–192, 192–216, and 216–240 h postdose. The actual start and end times for each collection will be recorded in the study data along with the weight.

Analytical Instrumentation

Radioactivity Determination

Radioactivity determinations are performed using a liquid scintillation counter (e.g., TriCarb model 2300TR Perkin Elmer). For plasma and urine analysis, duplicate weighed aliquots are mixed with scintillation cocktail (Ultima Gold XR) and analyzed directly by liquid scintillation counting. Duplicate aliquots of fecal homogenates and blood samples are weighed and combusted using an automatic sample oxidizer (Model 307, Perkin Elmer). The resultant [$^{14}CO_2$] is trapped in CarboSorb (Perkin Elmer) in combination with PermaFluor and radioactive content determined using liquid scintillation counting. Detected counts per minute (cpm) are converted to disintegrations per minute (dpm) using quench correction. The quench curves were prepared using

standards purchased from Perkin Elmer Life and Analytical Sciences and are prepared from stock solutions that are calibrated against National Institute of Standards and Technology (NIST) Reference Materials. The validity of the curves is checked regularly throughout the study.

Measurement of ABC123 in Plasma

Plasma concentrations of ABC123 will be determined following protein precipitation using a liquid chromatography with tandem mass spectrometry (LC/MS/MS) assay validated for concentrations between 5.00 and 5000 ng/mL, with quality control (QC) samples prepared at 12.5, 750, and 3500 ng/mL. ABC123 and its deuterated analog D6-ABC123 (IS) will be extracted from plasma and concentrations of ABC123 determined.

Control blank human plasma containing K_2EDTA as an anticoagulant will be obtained and stored at nominally −20 °C when not in use. Stability of ABC123 in human plasma will be determined.

Chromatographic Analysis

Sample Preparation

Urine

Equal proportions (between 1 and 5%) of each total sample weight will be combined to produce two time points per subject (0–4 and 4–24 h). All subsamples will be stored at around −70 °C. Aliquots of each subsample (100 μL) will be transferred into an Eppendorf tube and deionized water (400 μL) will be added before vortex mixing for ca. 5 s. The Eppendorf tube will then be centrifuged at 12,000 rpm at 4 °C for 5 min to sediment out any particulate matter before supernatant (140 μL) is transferred to a glass autosampler vial; the sample is ready for analysis.

Plasma

Plasma subsamples obtained in the study will be pooled per subject (1–24 h). All subsamples will be stored ca −70 °C prior to analysis.

Samples for extraction (2 mL) will be diluted with an equal volume of 10 mM ammonium acetate (aq). The diluted sample will then be subjected to the following SPE procedure using a Phenomenex Strata C18-E SPE cartridge (200 mg/3 mL):-

Prime: methanol (2 mL) followed by 10 mM ammonium acetate (aq) (2 mL)
Load: diluted sample (2 × 2 mL aliquots)
Wash: 10% methanol in 10 mM ammonium acetate (aq) (2 × 1 mL aliquots)
Elute: methanol:water (9:1 v/v) (2 × 2 mL aliquots)

The volume of the eluate will be reduced to ca. 400 μL under a gentle stream of nitrogen. The proportion of radioactivity extracted will be measured.

An aliquot of each extract (230 μL) will then be transferred into a glass autosampler vial. The vial will then be centrifuged at 12,000 rpm at 4 °C for 5 min to sediment out any particulate matter before being capped for analysis.

Feces

Feces subsamples obtained will be analyzed at two time points per subject, generally 24–48 h and 48–96 h: All subsamples will be stored at ca −70 °C prior to analysis.

Samples for extraction (ca. 5 g) will be combined with 10 ml water pH 4.0/Acetonitrile (20/80 v/v). Samples will be vortex mixed for ca. 1 min and then sonicated for ca. 5 min. The extraction step will be repeated a further two times. Samples will then be centrifuged for 10 min at ca. 8500 rpm and the supernatant decanted into a clean tube. Samples will be extracted up to three times and the extraction efficiency determined.

An aliquot of each sample (200 μL) will be transferred into an Eppendorf tube and deionized water (600 μL) added before vortex mixing for ca. 5 s. The Eppendorf tube will then be centrifuged at 12,000 rpm at 4 °C for 5 min to sediment out any particulate matter before the supernatant (230 μL) is transferred to a glass autosampler vial; the sample is ready for analysis.

Chromatography

Analysis of [^{14}C]-ABC123 and its metabolites in the pooled urine, feces, and plasma samples will be performed using a gradient elution method. The equipment will consist of a Surveyor liquid chromatography system (Thermo Fisher Scientific, UK) equipped with an autosampler and a variable wavelength detector operating at 254°nm. Separations will be performed using reverse-phase chromatography on a YMC-Pack ODS-AQ 150 × 4.6 mm, 3μm, 12nm column (Crawford Scientific, UK) fitted with a C18 guard column (Phenomenex, UK). [^{14}C]-ABC123 and its metabolites will be eluted from the column using a gradient based on two solvent mixtures:

Mobile phase A:	10 mM Ammonium acetate (pH 4.0)
Mobile phase B:	100 mM Ammonium acetate: acetonitrile (1:9 v/v)

The gradient will be established from 90% A at zerotime and held for 2 min, then the proportion of A decreased to 60% in a linear manner up to 60 min. The column will then be washed with 100% B for 5 min, returned to the initial conditions (60% A) and conditioned for 4 min prior to further injections. The flow rate will be 1.0 mL/min with a 4:1 split between the fraction collector and the mass spectrometer giving an approximate 200 μL/min flow to the mass spectrometer. Fractions will be collected every 13 s into 96-well Scintiplates (Perkin Elmer) for the duration of the analytical run using a modified CTC HTX PAL fraction collector (Presearch, UK). The tray temperature for holding samples will be set at 10 °C, while the column is held at 30 °C. Injection volumes will typically be 50–200 μL dependent upon sample.

Metabolites will be characterized using a Finnigan TSQ Quantum Ultra AM mass spectrometer (Thermo Scientific, UK), equipped with an electrospray source. Identification was performed using an LTQ Orbitrap XL mass spectrometer (Thermo Scientific, UK) and confirmed using accurate mass. The mass spectrometers will operated in the positive and/or negative ionization mode as appropriate. The instrument settings and potentials will adjusted as necessary to provide optimal data.

Review of the Method

Mass Balance

Following the publication of the regulatory guidelines in 2006, the recoveries obtained in clinical studies performed pre- and postguidance were reviewed to see if there was any effect on the recoveries obtained (McEwen et al. 2012). Urine and sample weights obtained from the volunteers were collated and the following mean values obtained:

Daily urine excreted (mL)	2283 ± 413
Daily feces excreted (g)	182 ± 19.2

If we assume typical sample sizes of 1 g and 0.25 g for determination of radioactive concentrations in urine and feces, respectively, then the detection limits (expressed as % administered dose) are defined by the following graphs (Fig. 2).

Radioactive doses administered in the studies were in the range 0.48–11.1 MBq (13–300 μCi). Even after administration of the lowest dose, the detection limits in urine and feces enabled determination of less than 0.1% dose per subject per day, thus easily enabling the relative routes of elimination to be adequately characterized. Beyond the 100 μCi dose (a value commonly employed in hADME studies), the increase in sensitivity provides no further advantage with regards to obtaining a mass balance and cannot be justified on this basis.

The recoveries obtained pre- and postguideline are provided in Table 5. In all cases, radioanalysis was performed in real time with results available within 8 h of sampling. The mean recovery obtained preguideline was 90.9 ± 8.0%, while the recovery postguideline was 91.8 ± 8.0%; therefore, implementation of the guidance had no notable effect on the recoveries obtained. It is notable however that the proportion of studies with >95% recovery decreased from 40% to 30% after 2008, but this is attributable to the fact

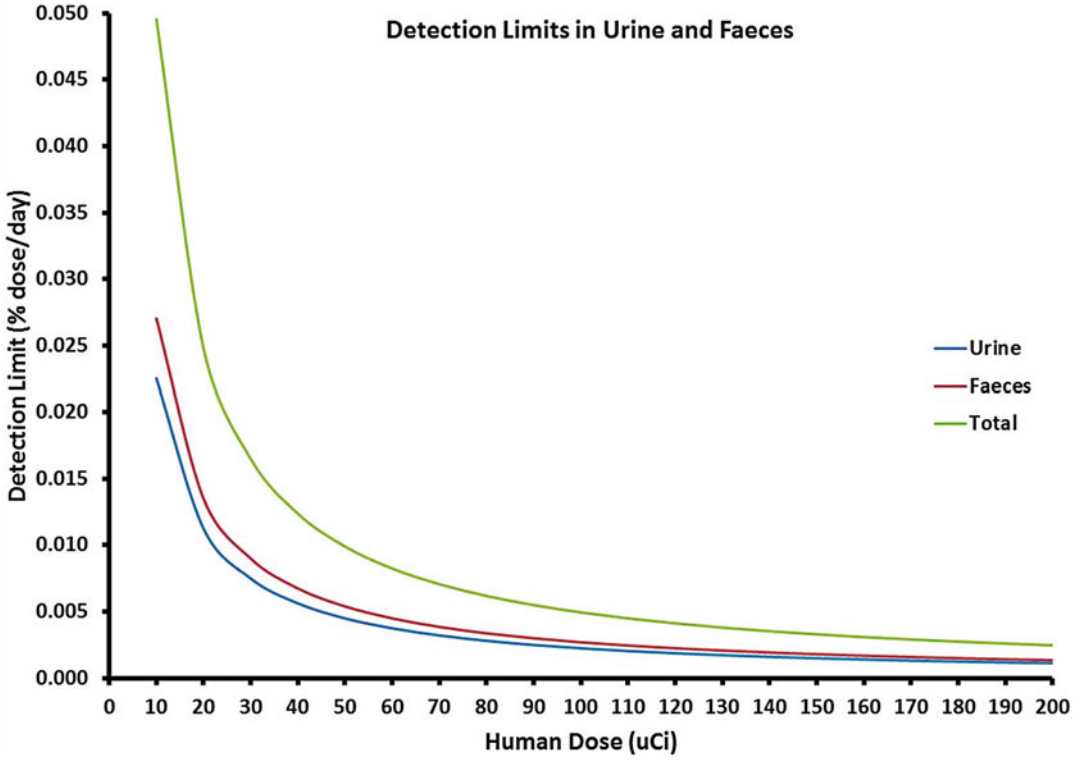


Fig. 2 Detection limits in urine and feces based on dose administered

Table 5 Recoveries pre- and post-MIST

Recoveries	Pre-Mist 2002–2008	Post-MIST 2008–2012	Overall
Mean 14C Recovery (%)	90.9 ± 8.0	91.8 ± 8.0	91.3 ± 7.9
Proportion of studies with			
>95% recovery	40.0	30.5	35.7
90–95% recovery	26.7	38.5	32.1
80–90% recovery	26.7	23.1	25.0
<80% recovery	6.7	7.7	7.1

that the release criteria were changed. The original release criteria were set at >95% recovery or <1% in 2 consecutive 24 h collections of excreta. This was amended to allow release at 90% recovery or 1% in two consecutive 24-h collections which explains the apparent fall in recoveries greater than 95%.

Recoveries obtained pre- and post-MIST are shown in Fig. 3. The recoveries were initially presented based on total recovery but were then re-evaluated based on the route of excretion. Three categories were selected, >75% radioactivity excreted in urine, >75% radioactivity excreted in feces, and "mixed" where the excretion was not primarily urinary or fecal. There was no clear relation between route of excretion and the recovery obtained.

Recoveries greater than 80% were achieved in 93.3% studies pre-MIST and 92.3% studies post-MIST. Overall 94.1% of studies performed over the review period achieved recoveries greater than 80%; 90% recovery was achieved in 66.7% studies pre-MIST and 69.2% studies post-MIST.

Key factors for obtaining a good mass balance were identified as:

- Accurate determination of the dose administered
- Complete collection of the urine/feces samples

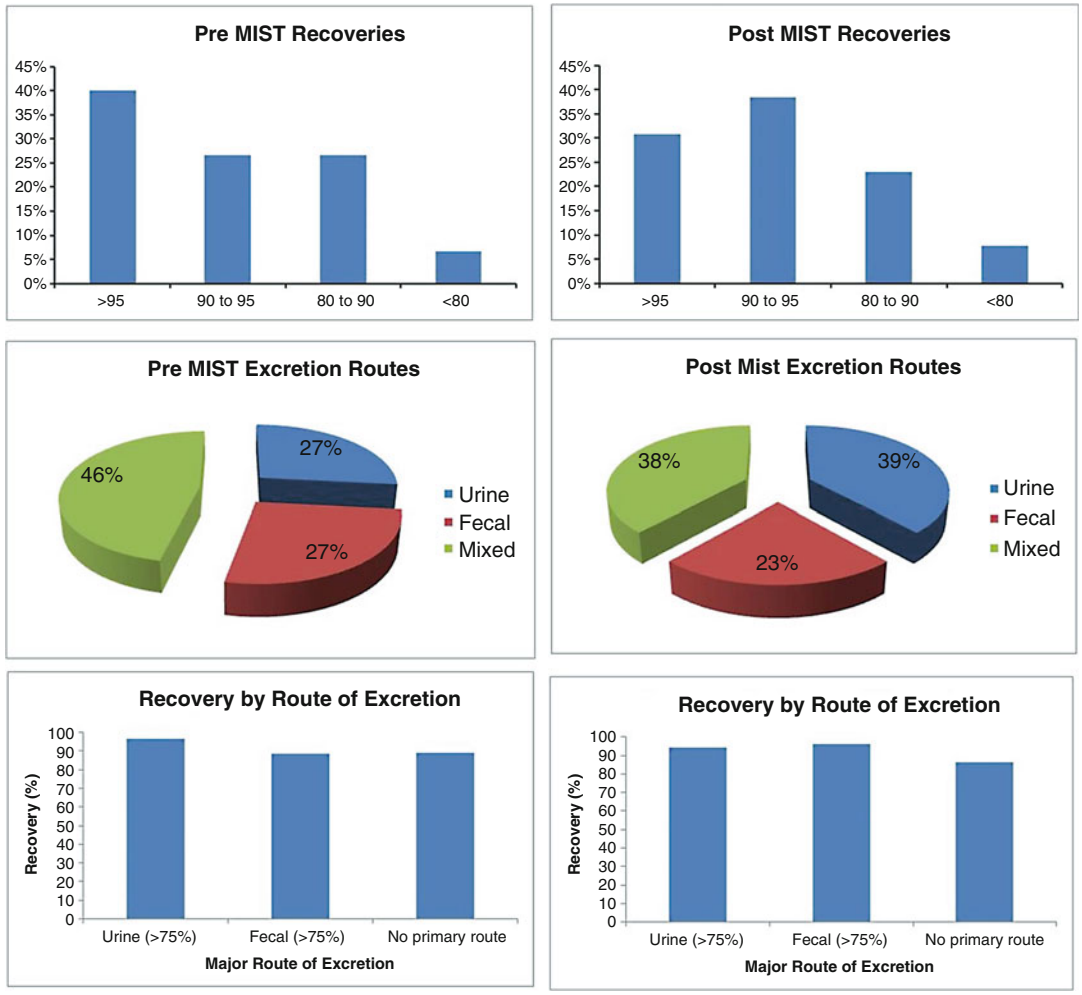


Fig. 3 Recoveries pre- and post- MIST

- The PK/ADME properties of the test compound

Metabolite Quantification and Identification

The hADME study provides valuable samples of urine, feces, and plasma which can provide useful information on the metabolic fate of the test compound in man. While the samples are initially assayed for radioactive content, with the aim of obtaining a mass balance, the main utility of the samples is in the quantification and identification of drug metabolites.

Due to the low quantities of plasma that can be taken during the hADME study, initial method development is usually performed using samples of urine and feces. This enables establishment of optimal chromatographic conditions prior to analyzing the precious plasma extracts. Excreta samples containing notable quantities of radioactivity (usually >90% excreted radioactivity) are selected for analysis and can be pooled by timepoint and subject as appropriate. The standard analytical method for metabolite investigations is liquid chromatography with radioactivity detection. The use of radioactivity makes no assumptions as to the fate of the drug entity and as the response is independent of structure relative

proportions of all components in the sample is achieved immediately.

The key objective of the hADME study is quantification and identification of the radioactive components in systemic circulation in order to meet the regulatory requirements of the ICH and FDA. As well as the limited amount of sample available, these investigations are also challenging due to fact that concentrations of radioactivity are generally much lower in plasma than in any other matrix. During the preclinical studies performed early in the drug development process, relatively high doses of radioactivity can be administered (50–100 μCi/kg) and the sample volumes are much lower. In the hADME study, the amount of radioactivity that can be administered is determined by the dosimetry assessment and for [^{14}C] is usually around 50–100 μCi. Where drugs exhibit poor bioavailability, extensive metabolism, or high volumes of distribution, this can result in extremely low concentrations of drug and thus radioactivity in circulation. The analytical challenge is compounded by the definition what constitutes a major radioactive component requiring identification. The initial regulatory guidance provided by the FDA indicated that a major radioactive component would be one that accounts for >10% AUC of parent AUC, the subsequent guideline produced by the ICH indicated that a major radioactive component was one that accounted for >10% total radioactivity AUC. For some time, this led to confusion and uncertainty with an unstated belief that the ICH guideline would be the standard that took precedence. The issue was resolved by the publication of an updated guidance from the FDA (FDA 2016) in which the definition of a major radioactive component was confirmed as one that accounted for >10% total radioactivity AUC.

In contrast to the determination of the mass balance of the test compound where measurements are expressed as % dose administered, the detection limits in plasma are dependent upon the specific activity of the test material. A standard bioanalytical assay developed for plasma will generally rely upon the amount of chemical present in the matrix, whereas for radioactive studies the detection limit increases as the chemical dose increases. This situation is shown graphically for different radioactive doses in Fig. 4. Taking 100 mg as the proposed human dose, the detection limit decreases as the dose given to the volunteers is increased. For a standard radioactive dose of 100 μCi, the detection limit increases as the proposed chemical dose increases. This should be borne in mind when reviewing the dosimetry data – will the proposed radioactive dose meet the objectives of the study – and can be used as a justification when moving from ICRP category I dose to a category IIa dose. It is also worth noting that low radioactive doses do not necessarily result in low specific activity material. Doses prepared for studies supported by accelerator mass spectrometry (AMS, discussed later in the chapter) may contain a low radioactive dose but also tend to be administered as a low chemical dose, thus resulting in a high specific activity. This has knock on effects on assessing the potential stability of the dose material.

Plasma samples can be prepared for quantification in a number of ways but are generally either across subjects at specific timepoints or more commonly by preparation of AUC pools using the methods published by Hamilton (Hamilton et al. 1981) or Hop (Hop et al. 1998). The AUC method provides information on the relative exposure of each metabolite compared to circulating total radioactivity, but can result in dilution of the radioactive response. In general, a mixed approach would be employed with an AUC pool supplemented by a couple of timepoint-specific analyses. AUC pools can be generated in a time- and volume-dependent manner across subject.

The gold standard approach to metabolite investigations using plasma samples from hADME studies has been the use of high performance liquid chromatography (HPLC) combined with radiodetection. In combination with modern high resolution mass spectrometers, the eluent from the column can be split (usually 10:90 or 20:80) to enable simultaneous quantification and identification of drug metabolites.

A major limitation to the use of radioflow detection methods is the poor sensitivity obtained. There are two main types of radioflow cell: those using a solid scintillant and those using a liquid

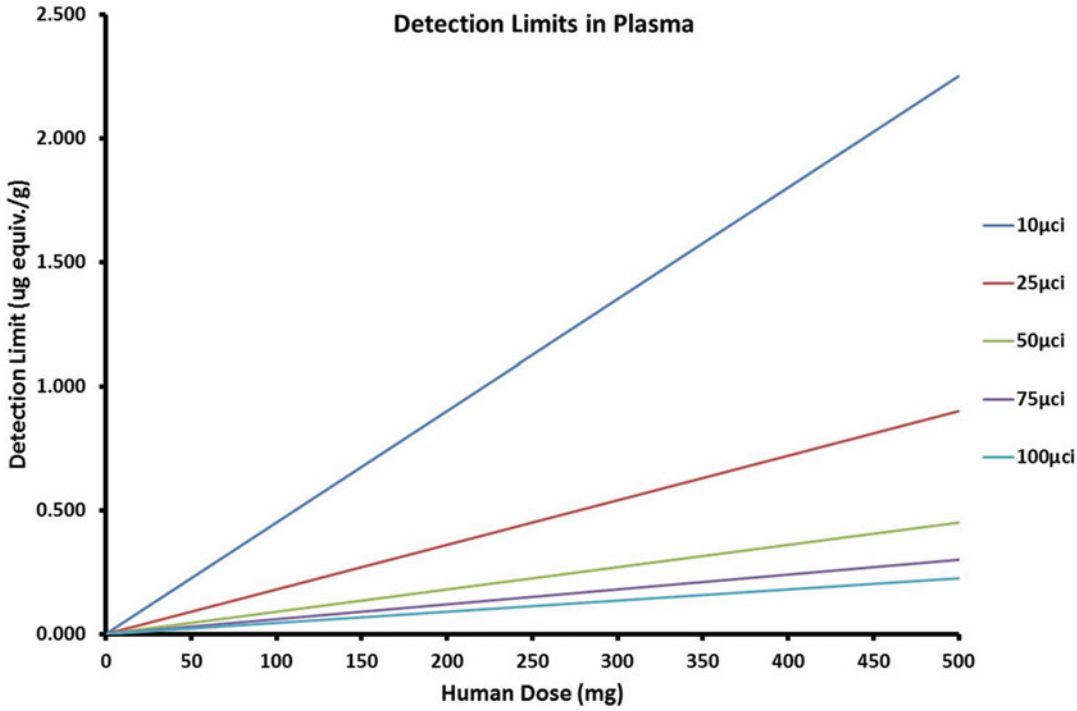


Fig. 4 Detection limits in plasma based on dose administered

flow cell where eluent is mixed with scintillator in the detector and counted. Of the two cell types, the flow through cell provides the best performance, but the key limiting factor for both detectors is the short residence time in the cell. Components in samples containing low concentrations of radioactivity can pass through the detector without providing sufficient counts for detection.

Recent detectors such as the BetaRam5 detector (LabLogic) use "active counting" (ACM^{TM}) to provide improvements in the signal observed. An evaluation of the technique was reported by Attwood et al. (2010) and indicated consistent retention times and compatibility with UPLC. The use of ACM^{TM} avoids the requirement for the sample to be fraction collected and counted off-line thereby eliminating the possible loss of volatile metabolites during sample processing.

Given the low concentrations of radioactivity in the critical plasma samples, the chances of success can be enhanced by employing good chromatographic practice. System refinements such as reducing dead volumes and shortening the distance between column and detector will increase the sensitivity, though it should be borne in mind that sharpening the chromatographic peak will shorten the residence time in the detector, potentially reducing the signal. Another consideration during the chromatographic analyses is the "quenching" effect that is often associated with radioactive measurements. Most investigators assume this to be constant throughout the chromatographic run although this is rarely checked. In addition to the radioactive components of interest, the samples will also contain a large number of co-eluting components present in the sample matrix. The presence or absence of quenching effects can be checked by running a blank sample and adding test material to the eluate. This can be achieved by a variety of methods such as direct infusion postcolumn, or postcolumn by collecting fractions and spiking.

During standard method establishment investigations, the suitability of the system is often established using parent material, but once a complex mixture is injected for analysis, the relative

recovery of components can vary. The use of radiolabeled materials in the hADME study allows column recovery to be determined easily. This can alleviate concerns that material has been retained on the column, the stainless steel tubing, or the radiodetector cell. By measuring the column recovery, the system performance can be optimized as part of the method establishment process.

In an attempt to overcome the limitations associated with the radioflow systems, alternative methods were investigated and these can be divided into two main categories: a) stop-flow and b) dynamic flow methods. The stop-flow methods are as the name suggests based on stopping the flow once a radioactive peak is detected, thus providing longer detection times and therefore higher detection efficiencies. The peak is held within the detector as opposed to passing through therefore increasing the signal to noise ratio. The technology was originally developed for investigations performed in the agrochemical industry where samples routinely contain low levels of radioactivity. The advantages of stop-flow technology have been discussed in the literature (Nassar et al. 2004), but as yet the stop-flow technique has not been widely adopted in the drug development process. The major drawbacks associated with stop-flow technology are that subsequent analyses can result in inconsistent retention times and the detector is incompatible with LC/MS, meaning the samples need to be analyzed twice. Stop-flow has a lower throughput than alternative methods as counting occurs during the chromatographic run and therefore extends the chromatographic run times.

The use of a modified "dynamic flow" radiodetector was described by Cuyckens et al. (2008). Improvements in sensitivity were achieved by a modification to the standard online radiochemical detection system introducing the capability to provide variable scintillation fluid flow. Further improvements were achieved by reducing the internal diameter of the tubing, resulting in better peak shape, increased sensitivity, and higher resolution. When compared to conventional radio-HPLC using [^{3}H]- and [^{14}C]-labelled compounds, the method was reported to have comparable sensitivity to conventional techniques, was compatible with UPLC (thus shortening the chromatographic run times), and was suitable for hyphenation with mass spectrometers.

The detection limits obtained using the radiochromatographic method can be significantly improved by performing the quantitation off-line. Eluent from the HPLC column can be fraction collected directly into scintillation vials, appropriate scintillation cocktail added, and the samples counted using standard liquid scintillation counting (LSC) methods. Each sample (fraction) can be counted for longer time periods, thus dramatically improving the sensitivity. On standard liquid scintillation counters, the samples are typically counted one vial at a time. If the chromatographic run is 30 min, with fractions collected at 15 s intervals, then counting each vial for 4 min would result in a total counting time of over 4 h.

The utility of off-line counting was significantly improved by the introduction of microplate scintillation counting (MSC) plates (Dear et al. 2006; Krauser et al. 2012). When performing chromatographic analysis with MSC counting, the eluent is collected directly into microplates, 96 or 384 well, using accurate fraction collectors. The technique has the advantage that several plates can be selected per run and the process can be automated thus improving the throughput. Two types of plate are commercially available, one with a solid scintillant base and another employing liquid scintillant. Use of solid scintillant plates results in a slightly lower sensitivity but allows recovery of notable fractions from the plate postcounting for further characterization using mass spectrometric techniques. Both types of plate give a notable improvement in sensitivity when directly compared to the results obtained using traditional fraction collection-liquid scintillation counting methods. Unlike traditional liquid scintillation counters, microplate scintillation counters are able to count multiple wells simultaneously (12–16 dependent on counter); therefore, the throughput is much higher. The technique should however be used with care as one of the key steps in sample analysis is evaporation of the eluent from the plate at which point there is

Table 6 Comparison radiodetector sensitivity

Radiodetection	Background (CPM)	Counting Efficiency (%)	Counting time (min)	Limit of Detection (DPM)	Limit of Quantification (DPM)
HPLC-RFD	15	70	5–10 s	250–500	750–1500
HPLC-LSC	25	90	10	10	31
HPLC-MSC	2	70	10	5	15
Stop-flow	15	70	1	25–50	75–150
HPLC-AMS				0.0001	

Taken from Zhu et al. 2005

the possibility to lose volatile components (parent drug or metabolites). It is therefore good practice to perform a system suitability check before committing the precious samples (especially plasma) for analysis. A good system suitability check would include a recovery check for radioactivity from plates spiked with parent compound and would ideally compare radioprofiles obtained from other biological matrices such as urine using both radioflow detection and microplate scintillation counting. An improved method for quantifying the radioactive content of microplate fractions has subsequently been reported (Dear et al. 2008). The method was essentially an imaging technique and was reported to shorten counting times required for analysis.

The counting methods detailed above were compared by Zhu et al. (2005) and the relative limits of detection discussed. The data are reproduced in Table 6 and show that of all the commonly used methods of radiometric detection the microplate scintillation counter provides the lowest limit of detection, with the exception of accelerator mass spectrometry. Whilst AMS has an extremely low detection, it is not a true radiometric method as quantitation is based on graphitization and measurement of [^{14}C] atoms.

The relative sensitivity of the common techniques used for quantification of drug metabolites (radioflow, standard fraction collection, and microplate scintillation counting) was assessed in a comparative study using the same sample (McEwen et al. 2014). In general good agreement was observed using all three counting techniques, but overall the microplate scintillation counting provided greater resolution of the chromatographic peaks and a lower background. This consideration assumes great importance when selecting the analytical method for the analysis of clinical plasma samples where concentrations of drug-related material, and therefore concentrations of radioactivity, are generally low.

Alternatives to Carbon-14

NMR

The development of high field NMR machines in the 1980s led researchers to explore the use of NMR for analysis of plasma and urine samples. The technique was suitable for the study of both endogenous compounds (biomarkers) and drug metabolites. The technique requires limited sample preparation and is nondestructive, meaning that the sample can be retained for use in additional experiments.

The use of NMR for quantification and identification of drug metabolites in clinical and preclinical samples was developed by Nicholson (Nicholson et al. 1983, 1984a, 1985; Bales et al. 1984a, b, 1985). NMR is an inherently insensitive technique but has the advantage that rapid multicomponent analyses can be performed with limited sample preparation and with no prior assumptions as to the sample identity. The technique also provides structural information sometimes absent from mass spectrometric data such as the specific site of hydroxylation. Analyte concentrations need to be >50 μM for effective detection and the molecule requires suitable proton groups such as CH3-, -CH2-, or CH. The early promise faded as researchers turned to mass spectrometric techniques, but recently NMR has been used to

generate quantitative data to answer regulatory questions (Dear et al. 2008; Caceres-Cortes and Reilly 2010). NMR has also been used to investigate likely routes of metabolism and provide an estimate of renal clearance (Nedderman et al. 2011) in the absence of traditional radioactive tracers.

Initial studies were conducted using proton NMR as most drugs contain hydrogen atoms, and this also provides structural information. For compounds containing fluorine, 19F NMR provides the opportunity to obtain quantitative data, although structural information obtained is limited. The low levels of fluorine in biological systems and the fact that fluorine 19 has a 100% isotopic abundance mean that all signals obtained can be related to drug or drug related material. This has led to many researchers using the technique in early drug metabolism (Dear et al 2000, Desmoulin et al. 2002, Ismail et al. 2002, Lenz et al. 2002, Malet-Martino et al. 2006). The technique continues to generate interest and recently two groups have published papers (James et al. 2017, Haitao et al. 2017) comparing the results obtained using F19 NMR with those obtained using the radiolabeled compounds.

Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

ICP-MS is a technique that can detect a wide range of elements (metals and nonmetals) at extremely low concentrations, in the order of on part in 10^{15} (ppq). While the technique generally covers elements not found in the pharmaceutical compounds, for drugs containing halogen atoms, especially bromine and iodine, ICP-MS offers an alternative method for detection and quantification of drug-related material. The advantage of the technique is that detection involves atomization and ionization of the compound, meaning that quantification is independent of chemical structure and can be performed without the requirement for synthetic standards. The technique can be used to determine several elements simultaneously and can be combined with both normal and reverse phase HPLC to separate and quantify drug metabolites in biological fluids (Axelsson et al. 2001; Duckett et al. 2002). While ADME studies are currently not conducted for in the development of "biologicals," it is interesting to note that the use of ICP-MS has recently been reported in the study of large molecule metabolism using a test compound labelled with iodine-127 (Lim et al 2014).

AMS

The concept of accelerator mass spectrometry (AMS) can be traced back to a cyclotron experiment conducted in the 1930s to measure 3H and 4He (Alvarez and Cornog 1939). The technique remained relatively underutilized until the publication of a paper by Richard A Muller in 1977. His paper in Science showed how particle accelerators (cyclotrons and linear) could be used for detection of tritium, carbon-14, and several other isotopes of scientific interest including beryllium-10, widely used in geology. He also published the first radiodating determination using tritium. Within month other groups (Nelson et al. 1977; Bennett et al. 1977) published further data using linear particle accelerators. These measurements were made using relatively large accelerators operating with terminal voltages of 7 and 8 MV. Since these early experiments, smaller instruments have been introduced and the trend toward smaller and smaller machines continues. Modern instruments operating at low accelerator voltages such as 200 kV for determination of 14C are being designed and built by companies such as ETH Zurich (Suter 2010).

Originally used in academic institutions as a radiocarbon dating technique, the ability to determine extremely low concentrations of radioactivity associated with carbon-14 and tritium was of great interest to scientists involved in the drug development process, especially those involved in the safety assessment of drug metabolites. The application of AMS in the drug development process has been the subject of several general reviews (Lappin and Garner 2005; Lappin and Stevens 2008). Despite being described in some quarters as a low level counting technique, AMS

does not directly count radioactive decay particles but instead provides C12:C14 isotope ratios for the sample. AMS was first applied to biomedical studies in the 1980s (Litherland 1980). Since then further reviews detailing the biomedical applications of AMS have appeared (Lappin et al. 2006; Young and Ellis 2007; Young et al. 2008; Vogel et al. 2010; Seymour 2011). Specific examples of AMS investigations in the drug development process were also reported by Swart et al. (2016) and Bloomer et al. (2016).

One of the common misconceptions surrounding AMS is that it is a nonradioactive technique. While the amount of radioactivity involved in AMS studies is extremely low, such that the samples can be treated as nonradioactive and the usual human dosimetry is not required, the material administered to the volunteers is still radioactive and will require radiosynthesis in the usual way. Terms such as "lightly labelled" have been used, but the drug material should be prepared under accepted quality standards with a defined specific activity and structural characterization to ensure the study fulfils the regulatory objectives. While the regulatory requirements associated with administration of radioactive material to human volunteers may be eliminated, it is still good practice in the drug development process to obtain good quality data on the ADME properties of the investigational product in the preclinical species used in the toxicology assessments. The gold standard for these investigations is still the use of radiolabeled material. If these investigations are conducted around the same time as the hADME study, then contemporaneous analysis can be performed to compare the nature of circulating metabolites in preclinical and clinical samples.

While AMS provides exceptional limits of detection, it is not a technique that enables structural elucidation of the drug metabolites under analysis. Prior to analysis using AMS, all samples must be converted to elemental carbon, a process known as graphitization (Young et al. 2008). Samples of urine, fecal extracts, and/or plasma must be separated using chromatography, fractionated, and each fraction graphitized for analysis by AMS, thus generating a radioprofile similar to HPLC-LSC (Young and Ellis 2007; Young et al. 2008). Currently this is an inherently slow process requiring manual processing, and while progress in designing a system capable of transforming liquid inlet to a gaseous CO_2 output has been reported (Liberman et al. 2004, 2007), this has yet to become commercially available.

Equally valuable information can be obtained using modern high resolution mass spectrometers using early clinical samples from the single ascending dose and multiple ascending dose (SAD/MAD) studies. In the hADME study, hyphenated LC-MS/MS-RAD analysis can help optimize the separation of components while obtaining good quality quantitative data.

It could be argued therefore that while AMS can undoubtedly provide a tool to solve specific problems associated with low bioavailability and/or extensive metabolism, adoption of AMS as the gold standard is still some way off and will ultimately depend upon further advances in AMS technology.

Summary

The hADME study is a pivotal study in the drug development process, providing a bridge between the preclinical toxicology findings and the clinical studies. The standard design involves administration of radiolabeled to human volunteers, collection of excreta and plasma, determination of a mass balance, and quantification and identification of the components in systemic circulation.

Examples

Example A: Dosimetry

One of the key activities in the planning of a hADME study is synthesis of the radiolabeled drug material, and this requires two pieces of information (1) the amount of radioactivity that can be administered to the volunteers and (2) the proposed chemical dose. The dosimetry study is therefore critical to providing this information forming the basis for the risk assessment in man.

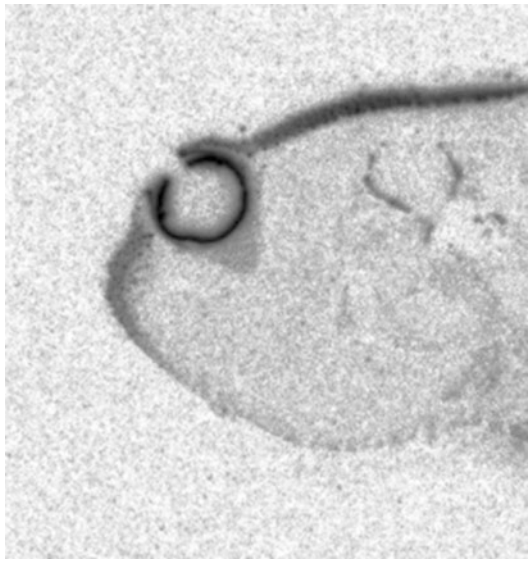


Fig. 5 Binding to melanin – uveal tract

Pigmented rats are used to assess binding of the drug material to melanin which is assessed from concentrations present in the uveal tract. Concentrations in pigmented skin can be measured, but there may be some confusion as to whether the observed radioactivity is bound to pigmented skin or the melanin present in the fur. An example image showing binding of test compound to the uveal tract is provided in Fig. 5.

Tissue distribution studies in support of dosimetry calculations are generally designed to provide information on the distribution and kinetics of the radiolabeled material following dosing. A typical set of images obtained is provided in Fig. 6. In this example, it can be seen that the radioactivity distributes throughout the animal by 2 h following administration and is then seen to be eliminated through the kidney and gastrointestinal tract so that at 168 h following dosing radioactivity is seen only in the liver, caecum, and uveal tract. Inclusion of an additional timepoint at 504 h shows that radioactivity has been almost completely eliminated from the animal with the exception of the uveal tract. In the example, provided tissues (with the exception of the uveal tract) were clear of radioactivity by 504 h. As a general rule the earlier complete elimination is observed, the higher will be the radioactive dose that can be administered to man.

To provide a realistic estimate of the risk associated with radioactivity present in the eye, the data can be analyzed to provide a biological half-life for modelling purposes. The tissue concentration data are provided Table 7 and the kinetic analysis is reproduced in Fig. 7. The data in this case indicate a half-life of around 213 h.

One further factor affecting the result will be the relative routes of elimination of radioactivity. In this case, radioactivity was eliminated by both fecal and urinary routes. Dosimetry assessment for compounds showing high fecal elimination usually result in lower values than those with high urinary elimination. This is due to the longer residence time of radioactivity that passes through the gastrointestinal tract which results in greater internal exposure for this organ.

Example B: Mass Balance

A typical hADME study usually runs for a fixed period (nominally 7 days) with radioactive content in urine and feces measured to assess the mass balance recovery. It should be stressed that the primary objective is to assess the routes and rates of excretion of radioactivity and while a full mass balance is welcomed it is not always achievable. Daily measurement of the radioactivity excreted in urine and feces allows the elimination to be followed in real time and can provide useful information on the rates of elimination as well as predicting the time to reach 90%. A useful method is to plot "body burden graphs" as shown in the examples below.

Compound A

The hADME study for compound A was designed on the basis that the metabolism was well understood and a fixed term of 9 days was assigned for residence in the clinic. Urine and feces were collected on a daily basis and radioactive content determined. The elimination was then characterized using a body burden calculation as described earlier. The data obtained are provided in Table 8 and the kinetic analysis is shown in Fig. 8.

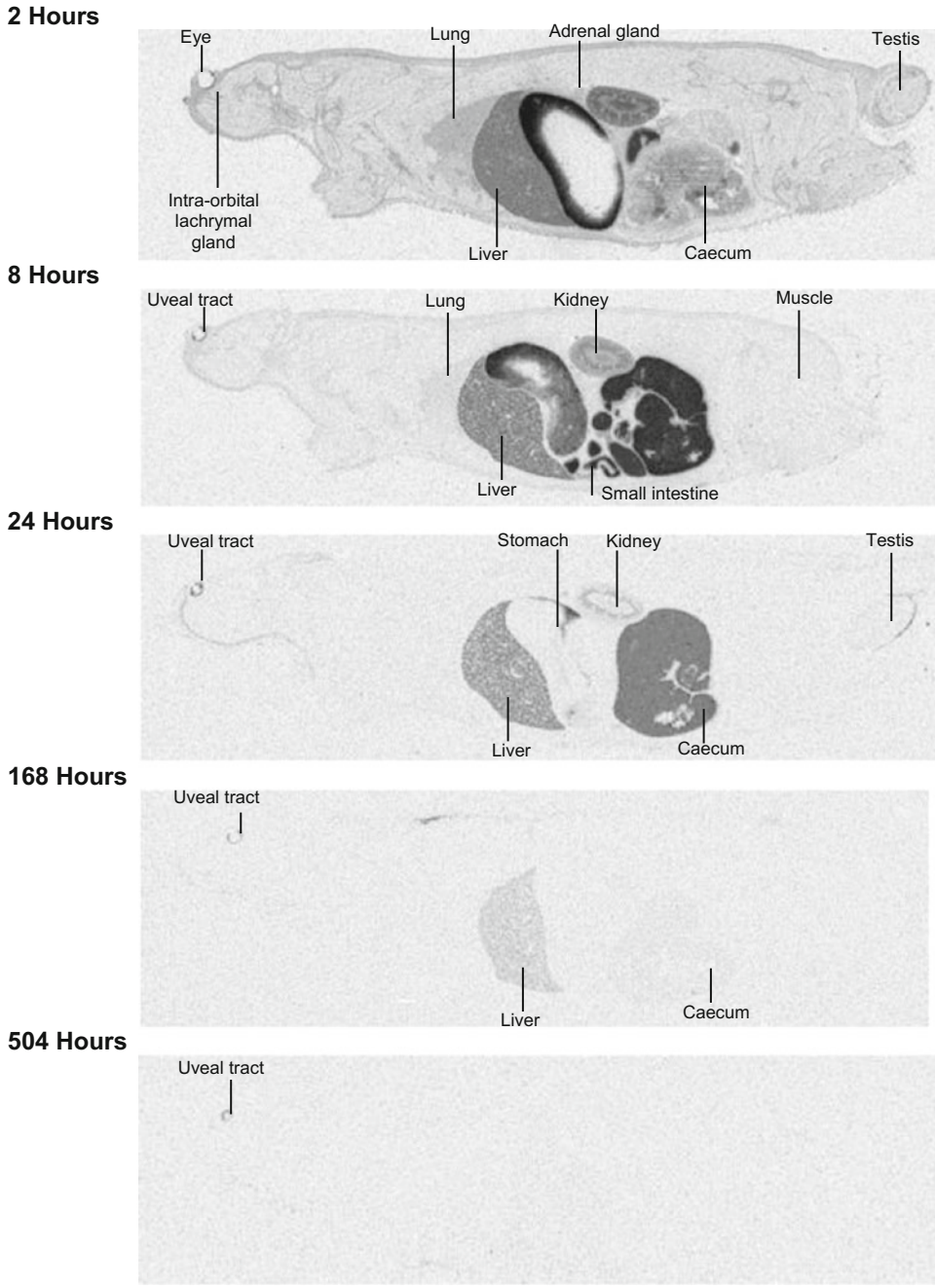


Fig. 6 Typical QWBA experiment for dosimetry assessment

Mean recovery obtained in this study was 87.5% (75.4% urinary, 12.1% fecal), thus below an arbitrary cut-off of 90% for a good recovery. The standard release criteria (<1% in two consecutive 24 h collections) were met in this study although there was no intention to extend collections beyond the original 9 days. The body burden analysis would have allowed a decision on whether to release the subjects and in this case the subjects would have been

released anyway. By 9 days the elimination half-life was estimated at around 425 h (17 days) meaning that if 90% recovery was the sole criterion for release, the subjects would be confined to the clinic for a further 10–14 days, while if 95% had been set as the release criteria, then the clinical phase would have been extended by around 3 weeks. It can be argued that the elimination routes and rates of radioactivity have been well characterized in this study and that any extension would have added little to the overall conclusions.

Table 7 Uveal tract data

Time (h)	Eye (uveal tract) concentration (μg equiv./g)
0	0
2	22.7
8	8.17
24	5.02
168	0.411
336	0.215
504	0.138

Compound B

The hADME study for compound A was designed on the basis that the elimination of radioactivity (predominantly fecal) would be protracted and a fixed term of 17 days was assigned for residence in the clinic. Urine and feces were collected on a daily basis and radioactive content determined. The elimination was then characterized using a body burden calculation as described earlier. The data obtained are provided in Table 9 and the kinetic analysis is shown in Fig. 9.

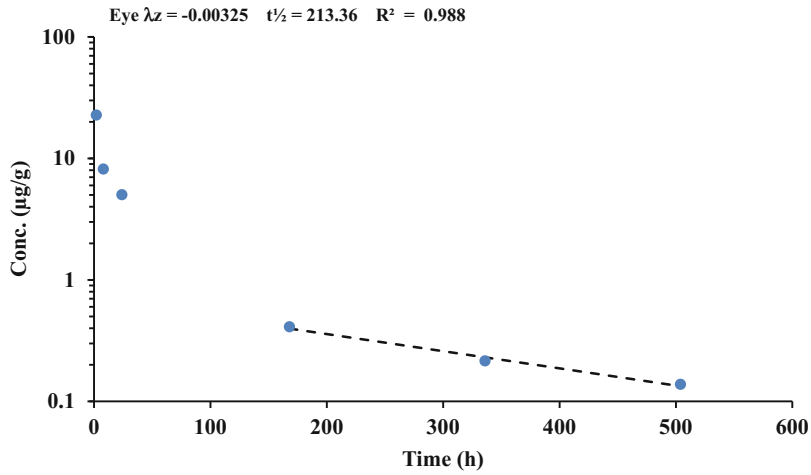


Fig. 7 Pharmacokinetic analysis of uveal tract data

Table 8 Excretion data obtained for compound A

Time (h)	Urine	Feces	Body Burden (100 – (U+F))
0	0	0	100
24	52.0	2.08	45.9
48	13.8	3.38	28.8
72	4.66	2.90	21.2
96	2.16	1.13	17.9
120	1.11	1.47	15.3
144	0.66	0.60	14.1
168	0.50	0.26	13.3
192	0.29	0.18	12.8
216	0.24	0.10	12.5
Total	75.4	12.1	–

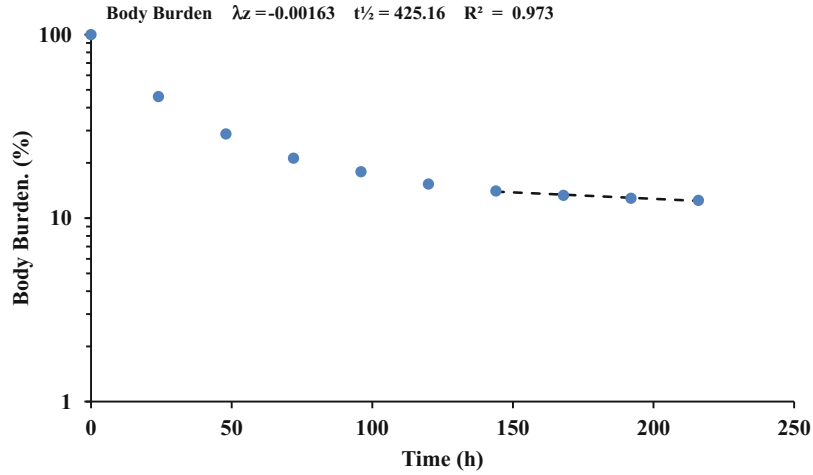


Fig. 8 Body burden compound A

Table 9 Excretion data obtained for compounds B and C

	Compound B			Compound C		
Time (h)	Urine	Feces	Body burden(100-(U+F)	Urine	Feces	Body burden(100-(U+F)
0	0	0	100	0	0	100
24	5.70	0.11	94.2	0.40	0.00	99.6
48	0.96	6.07	87.2	0.19	0.85	98.6
72	0.77	14.7	71.7	0.19	3.99	94.4
96	0.61	13.3	57.7	0.20	5.10	89.1
120	0.43	7.01	50.3	0.20	7.31	81.6
144	0.43	7.24	42.6	0.19	6.14	75.2
168	0.34	6.75	35.5	0.15	8.69	66.4
192	0.29	4.22	31.0	0.14	3.91	62.4
216	0.23	2.62	28.2	0.13	2.87	59.4
240	0.22	2.51	25.4	0.12	5.83	53.4
264	0.16	2.06	23.2	0.12	2.17	51.1
288	0.13	1.92	21.2	0.10	8.23	42.8
312	0.12	1.41	19.6	0.10	2.31	40.4
336	0.09	1.04	18.5	0.09	2.31	38.0
360	0.10	1.26	17.2	0.08	2.06	35.8
384	0.07	0.65	16.4	0.08	2.77	33.0
408	0.07	0.76	15.6	0.08	1.67	31.2
Total	10.7	73.7	–	2.56	66.2	–

Mean recovery obtained in this study was 84.4% (10.7% urinary, 73.7% fecal) thus below an arbitrary cut-off of 90% for a good recovery. The standard release criteria (<1% in two consecutive 24 h collections) were met in this study although there was no intention to extend collections beyond the original 17 days. The body burden analysis would have allowed a decision on whether to release the subjects and in this case the subjects would have been released anyway. By 17 days the elimination half-life was estimated at around 272 h (11 days), meaning that if 90% recovery was the sole criterion for release the subjects would be confined to the clinic for a further 7–10 days, while if 95% had been set as the release criteria then the clinical phase would have been extended by around 2 weeks. As for compound A it can be argued that the elimination

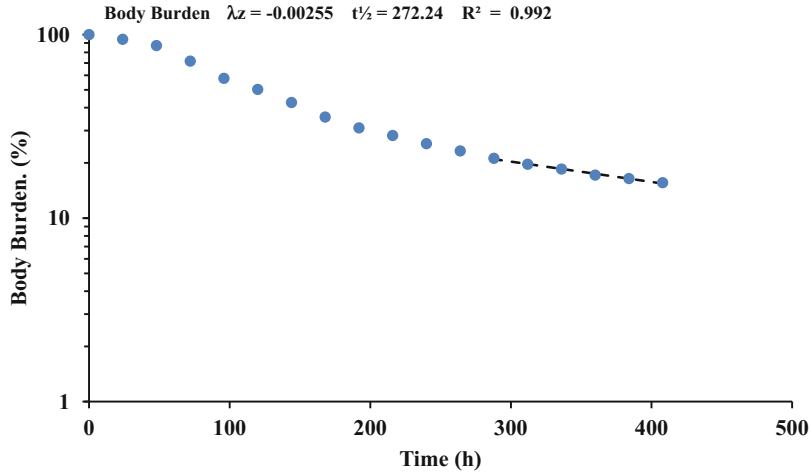


Fig. 9 Body burden compound B

routes and rates of radioactivity have been well characterized in this study and that any extension would have added little to the overall conclusions.

Compound C

The hADME study for compound C was again designed on the basis that the elimination of radioactivity (predominantly fecal) would be protracted and a fixed term of 17 days was assigned for residence in the clinic. Urine and feces were collected on a daily basis and radioactive content determined. The elimination was then characterized using a body burden calculation as described earlier. The data obtained are provided in Table 9 and the kinetic analysis is shown in Fig. 10.

Mean recovery obtained in this study was 68.8% (0.08% urinary, 66.2% fecal) thus below an arbitrary cut-off of 90% for a good recovery. The standard release criteria (<1% in two consecutive 24 h collections) were not met in this study although there was no intention to extend collections beyond the original 17 days. The body burden analysis would have allowed a decision on whether to release the subjects and in this case the subjects would have been confined to the clinic. By 17 days the elimination half-life was estimated at around 260 h (11 days), meaning that if 90% recovery was the sole criterion for release, the subjects would be confined to the clinic for a further 22 days, while if 95% had been set as the release criteria then the clinical phase would have been extended by around 4 weeks. As for compound A it can be argued that the elimination routes and rates of radioactivity have been well characterized in this study and that any extension would have added little to the overall conclusions.

Example C: Metabolite Quantification and Identification.

Once the radioactive measurements have been completed and a comparison of the AUC obtained from the bioanalytical measurements with the AUC obtained for total radioactivity performed, then the next key activity is the quantification and identification of notable metabolites.

The key to success at this stage is the preparation of a representative sample that can be analyzed using high performance liquid chromatography. Samples such as blood/plasma, bile, and urine are relatively easy to obtain and if they contain suitable quantities of radioactivity can be analyzed with minimal sample work up. Fecal samples require homogenization prior to analysis for radioactive content and can then be extracted prior to chromatography. The use of radiolabeled material is of benefit in this case as the efficiency of extraction can be determined and methods can be established beforehand using blank samples spiked with test compound. A limitation of the AMS technique is that the process is slower than if performed using standard radioactive doses. Fecal samples can be extracted, centrifuged and aliquots of the extract

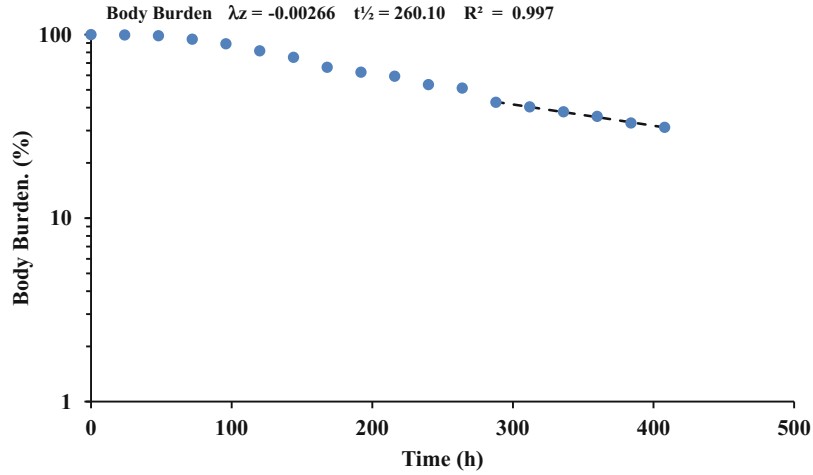


Fig. 10 Body burden graph compound C

counted in real time, thus allowing decisions to be made following each set of counts.

Standard preparation methods usually employ concentration as the final step prior to analysis, and while losses can occur at every step, the potential for loss of material at this stage is quite high and should be assessed as part of the system suitability evaluation. A wide range of sample preparation methods are available; the most commonly employed are liquid-liquid extraction (Pedersen-Bjergaard and Rasmussen 2005) and solid-phase extraction (Moriwaki et al. 2002). The aim of the work up is to separate the analyte from the endogenous material, thus minimizing matrix effects and improving sensitivity (Kruve et al. 2009; Marchi et al. 2009). Ideally the sample preparation method should be as simple as possible, the greater the number of steps employed, the more likely it is that losses can occur or artifacts arise due to instability of the test compound.

As discussed earlier in the chapter, the standard approach for separation and quantification of drug metabolites is liquid chromatography followed by microplate scintillation counting. The utility of the technique will now be discussed. Prior to analysis of the plasma samples, the technique was checked for quench and evaporation.

Quench was checked by analyzing a blank plasma extract and collecting fractions into microplates. By use of a t-valve, a solution of parent compound could be introduced into the eluant prior to evaporation. The chromatographic analysis was performed using a 90-min gradient method with collection of 14 s fractions therefore requiring the use of four 96-well microsintillation plates. The results are shown in Fig. 11 and indicate the response is relatively constant over the first three plates. A spike in the detector response is sometimes obtained during plate changeover and analysis of results should take this into account. It is good practice therefore to perform analysis in duplicate to check for potential artefacts. Towards the end of the collection, there is a notable decrease in the response obtained, which then recovers to around the normal response by the end of the collection period. The reduction in response is caused by quenching due to endogenous material being washed off the column at high organic solvent ratio.

The effect of evaporation can be checked using cold material spiked into the injection solvent and loaded into the microplate wells. The solvent is then evaporated and the contents of the well checked using standard MS analysis. An example is provided in Fig. 12. Here solutions of parent and metabolite (M1) were placed into the microplate wells and the solvent removed using the standard method. Multiple wells were prepared for parent and metabolite and at different times the contents of one well were checked at using MS. In this example, it can be seen that the evaporative process has no effect upon the concentration of parent observed, but that after a certain time the concentration of M1 in the cells decreases probably due to insufficient organic

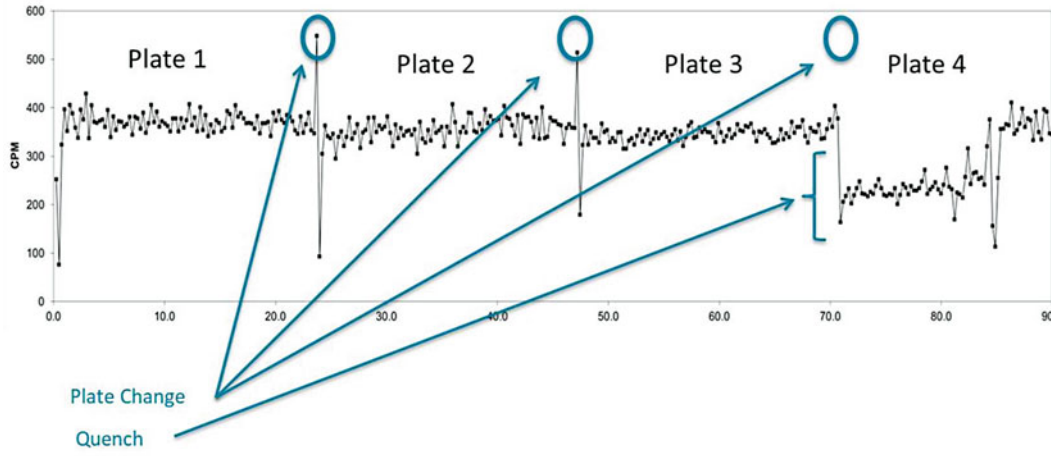
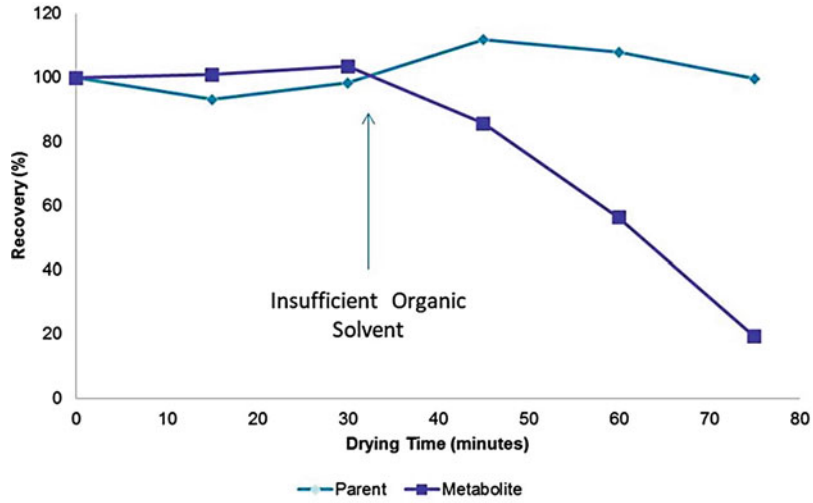


Fig. 11 Assessment of quench in the microplate counter

Fig. 12 Assessment of evaporation during sample processing



solvent remaining, thus allowing the component to volatilize.

The relative sensitivities of the radioflow detector and the microplate counting method were discussed by Zhu et al. (2005) and a graphical example is provided in Fig. 13. In the upper chromatogram, a radioflow detector is presented and shows the presence of parent and a very noisy baseline. In the lower chromatogram, the same sample is analyzed using microplate scintillation counting and quite clearly shows the presence of two minor metabolites in circulation. All three components were quantified and the relative exposure established.

An additional advantage of the microplate technique is that given the long half-life of carbon-14 the samples can be re-counted to improve the signal to noise ratio. Figure 14 shows the data that can be obtained in this way. The upper chromatogram was reconstructed from fractions counted for 4 min. While there is a suggestion of some metabolites present in the sample, the situation can be resolved by extending the counting time to 30 min. In this chromatogram (lower trace), the presence of parent and four metabolites (M1–M4) can be established. Counting times of 30 min are generally

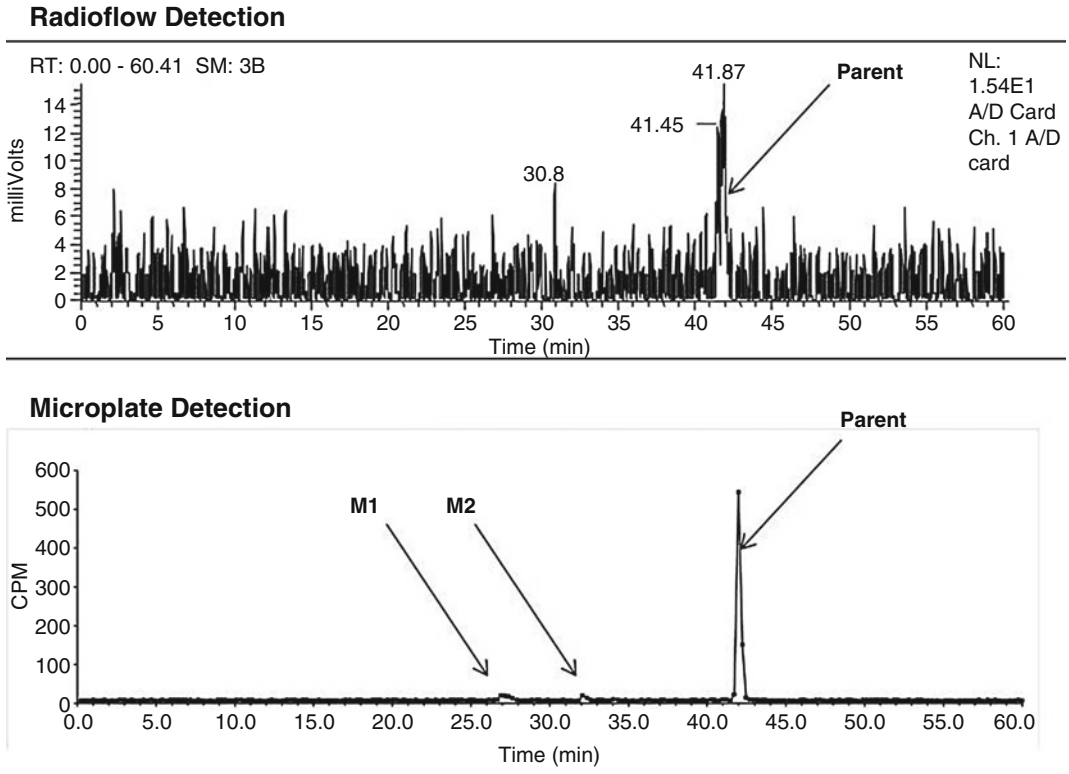


Fig. 13 Comparison radioflow and microplate detection methods

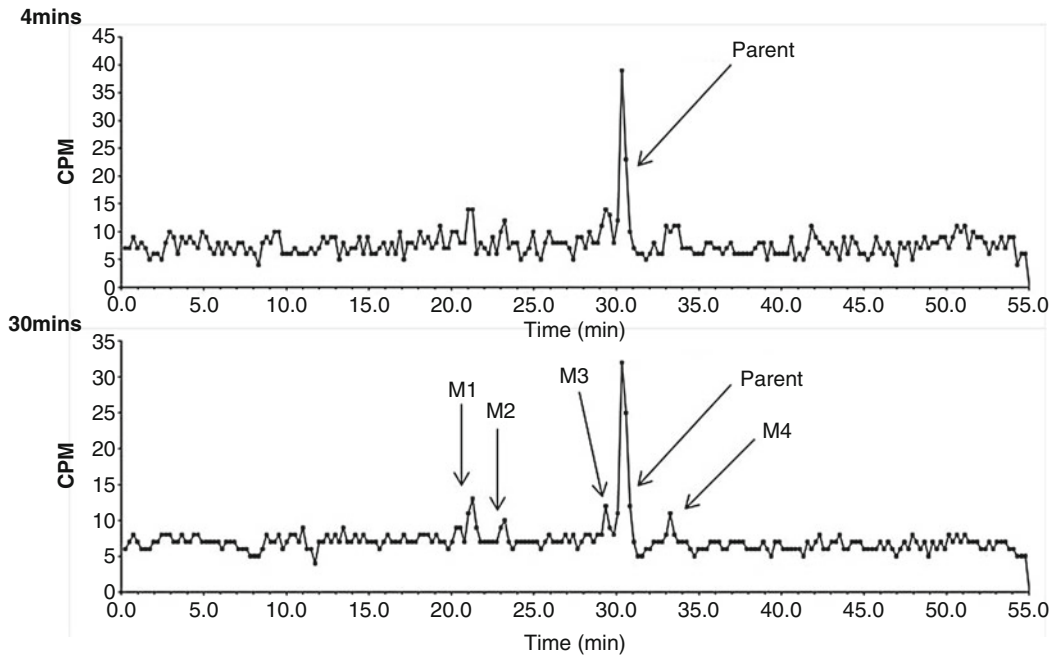


Fig. 14 Effect of extending counting time on microplate detector response

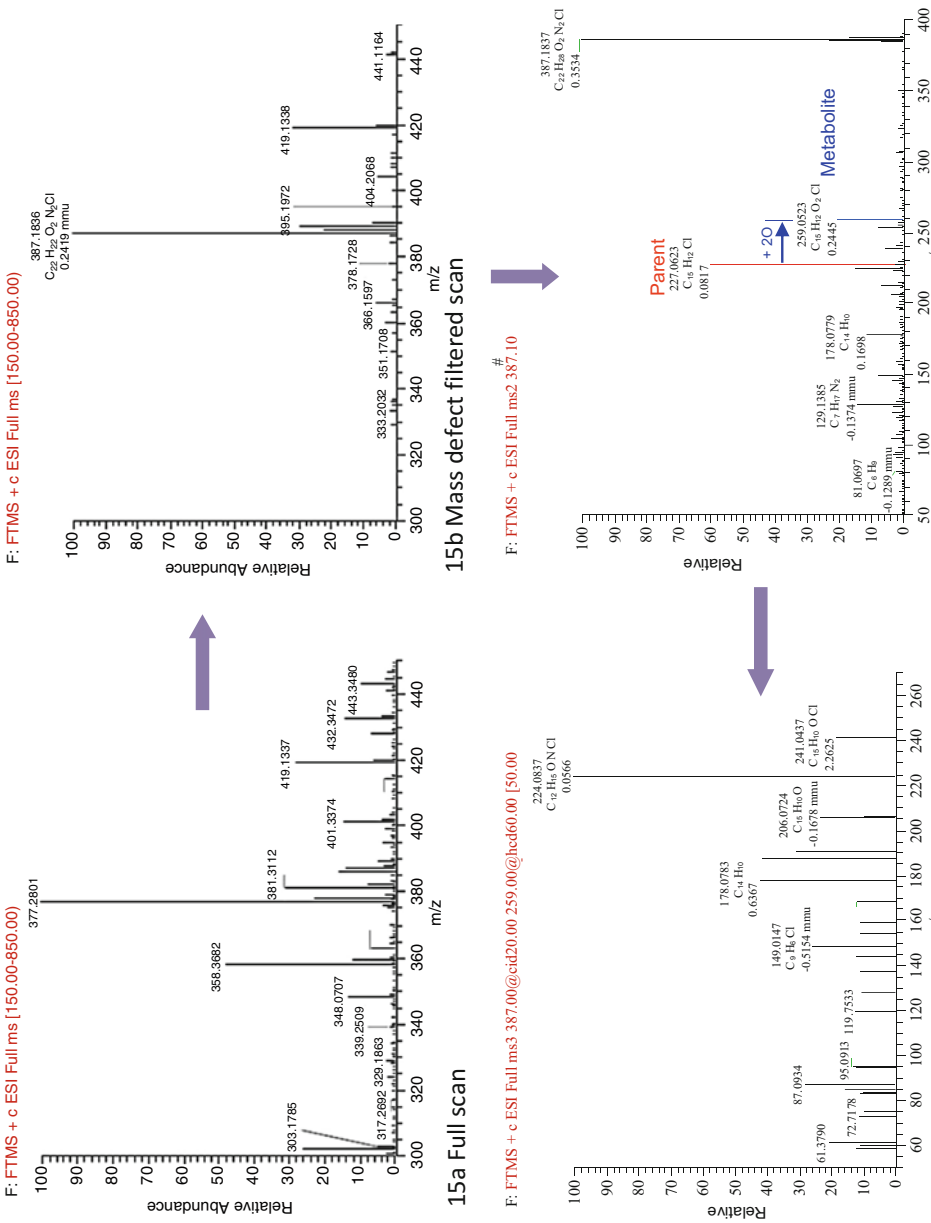


Fig. 15 Use of data processing in metabolite ID investigations

uncommon, but in some cases plates have been analyzed for up to 50 min.

Given the widespread use of mass spectrometric techniques in the preclinical and early clinical phases of drug development, it is unsurprising that LC-MS/MS is the analytical method of choice for the identification of drug metabolites in the hADME study. There has been some debate within the bioanalytical community regarding methods for quantitation of circulating components, and in 2010 the European Bioanalysis Forum produced a white paper offering recommendations and discussing the issues (Timmerman et al. 2011).

The drawback associated with relying on LC-MS/MS alone for the quantification is that in the absence of standards the LC-MS response can show a 30-fold difference in relative exposure when compared to radiometric methods. The analyte response is structure and matrix dependent and subject to unpredictable ion suppression effects. By contrast, quantification using radiometric techniques is structure independent and less subject to matrix effects.

Having optimized the sample preparation methods and the chromatographic separation obtained, the final key decision is the choice of mass spectrometer used for identification. The mass analyzers of greatest use in the analysis of human plasma samples tend to be based on either Q-TOF (Mamyrin 2001) or Orbitrap (Erve et al. 2009) technology in part due to their high resolving power but also due to the associated data processing techniques. Commonly used data processing methods rely on background subtraction, neutral loss filtering, isotopic pattern recognition (and the use of 14C can introduce an isotope pattern into the molecule), and mass defect filtering.

As an example, a full scan mass spectrum obtained on analysis of plasma from a hADME study is provided in Fig. 15a. The main peaks seen here are not compound related, arising instead from endogenous compounds within the matrix. By applying a mass defect filter to the chromatogram, only compounds that are related to the test compound are detected and the spectrum has therefore been simplified (Fig. 15b). Further processing can then be directed onto relevant components and MS/MS experiments conducted. In this case, the addition of two oxygen atoms onto the parent structure has been confirmed (Fig. 15c) and the product spectrum can then be obtained for full identification of the drug metabolite (Fig. 15d).

References and Further Reading

Alvarez L, Cornog R (1939) Helium and hydrogen of mass 3. Phys Rev 56:613

Attwood J, O'Brien B, Loaring HW, Williams A (2010) Evaluation of active counting mode option in LabLogic β-RAM model 5 and Laura software. Poster ISSX 2010. Istanbul, Turkey

Axelsson B-O, Jornten-Karlsson M, Michelsen P, Abou-Shakra F (2001) The potential of inductively coupled plasma mass spectrometry detection for high-performance liquid chromatography combined with accurate mass measurement of organic pharmaceutical compounds. Rapid Commun Mass Sprectrom 15:375–385

Baillie TA, Cayen MN, Fouda H, Gerson RJ, Green JD, Grossman SJ, Klunk LJ, LeBlanc B, Perkins DG, Shipley LA (2002) Drug metabolites in safety testing. Toxicol Appl Pharmacol 182:188–196

Bales JR, Higham DP, Howe I, Nicholson JK, Sadler PJ (1984a) Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. Clin Chem 30:426

Bales JR, Sadler PJ, Nicholson JK, Timbrell JA (1984b) Urinary excretion of acetaminophen and its metabolites as studied by proton NMR spectroscopy. Clin Chem 30:1631

Bales JR, Nicholson JK, Sadler PJ (1985) Two-dimensional proton nuclear magnetic resonance "maps" of acetaminophen metabolites in human urine. Clin Chem 31:757

Bennett CL, Beukens RP, Clover MR, Gove HE, Liebert RB, Litherland AE, Purser KH, Sondheim WE (1977) Radiocarbon dating using electrostatic accelerators: negative ions provide the key. Science 198:508–510

Bloomer JC, Beaumont C, Dear GJ, North S, Young G (2016) The value of metabolite identification and quantification in clinical studies. Some case studies enabling early assessment of safety in humans: GlaxoSmithKline. In: Metabolite safety in drug development. Wiley, Hoboken, NJ, pp 275–292

Caceres-Cortes J, Reilly MD (2010) NMR spectroscopy as a tool to close the gap on metabolite characterisation under MIST. Bioanalysis 2:1263–1276

Cuyckens F, Koppen V, Kembuegler R, Leclercq L (2008) Improved liquid chromatography-Online radioactivity detection for metabolite profiling. J Chromatogr A 1209:128–135

Dear GJ, James AD, Sarda S (2006) Ultra-performance liquid chromatography coupled to linear ion trap mass spectrometry for the identification of drug metabolites in biological samples. Rapid Commun Mass Spectrom 20:1351–1360

Dear GJ, Roberts AD, Beamont C, North SE (2008) Evaluation of preparative high performance liquid chromatography and cryoprobe nuclear magnetic resonance spectroscopy for the early quantitative estimation of drug metabolites in early clinical studies. J Chromatogr B 876:182–190

Desmoulin F, Gilard V, Malet-Martino M, Martino R (2002) Metabolism of capecitabine, an oral fluorouracil prodrug: 19F NMR studies in animal models and human urine. Drug Metab Dispos 30:1221–1229

Duckett CJ, Bailey NJC, Walker H, Abou-Shakra F, Wilson ID, Lindon JC, Nicholson JK (2002) Quantitation in gradient high performance liquid chromatography/inductively coupled mass spectrometry using diclofenac and chlorpromazine. Rapid Commun Mass Spectrom 16:245–247

EMEA (2012) ICH guideline M3 (R2) – questions and answers EMA/CHMP/ICH/507008/2011

Erve JCL, Gu M, Wang Y, DeMaio W, Talaat RE (2009) Spectral accuracy of molecular ions in an LTQ/Orbitrap mass spectrometer and implications for elemental composition determination. J Am Soc Mass Spectrom 20:2058–2069

European Clinical Trials Directive Communication from the Commission – Detailed guidance on the request to the competent authorities for authorisation of a clinical trial on a medicinal product for human use, the notification of substantial amendments and the declaration of the end of the trial (CT-1) (2010/C 82/01) Available at https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2010_c82_01/2010_c82_01_en.pdf Accessed 8 Aug 2017

Haitao Hu, Kishore Kumar Katyayan, Boris A. Czeskis, Everett J (2017) Perkins and Palaniappan Kulanthaivel Drug Metabolism and Disposition 45(4):399–408

Hamilton RA, Garnett WR, Kline BJ (1981) Determination of mean valproic acid serum level by assay of a single pooled sample. Clin Pharmacol Ther 29:408–413

Hill H (2009) Development of bioanalysis: a short history. Bioanalysis 1(1):3–7

Hop CECA, Wang Z, Chen Q, Kwei G (1998) Plasma-pooling methods to increase throughput for in vivo pharmacokinetic screening. J Pharm Sci 87(7): 901–903

Hu H, Katyayan KK, Czeskis BA, Kulanthaivel EJPP (2017) Comparison between radioanalysis and 19f nuclear magnetic resonance spectroscopy in the determination of mass balance , metabolism, and distribution of pefloxacin. Drug Metab Dispos 45(4):399–408

Humphreys WG, Unger SE (2006) Safety assessment of drug metabolites: characterisation of chemically stable metabolites. Chem Res Toxicol 19:1564–1569

ICH (2009) M3 (R2) non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals. Center for Drug Evaluation and Research (CBER), Silver Spring, MD, USA

ICRP (1992) International commision on radiological protection. Radiological protection in biomedical research. ICRP publication 62 Ann ICRP 22 No 3

Ismail IM, Dear GJ, Roberts AD, Plumb RS, Ayrton J, Sweatma BC, Bowers GD (2002) N-O glucuronidation: a major human metabolic pathway in the elimination of two novel anti-convulsant drug candidates. Xenobiotica 32:29–43

James AD, Marvalin C, Luneau A, Meissner A, Camenisch G (2017) Comparison of 19F NMR and 14C measurements for the assessment of ADME of BYL719 (Alpelisib) in humans. Drug Metab Dispos 45:900–907

Kuerzel U, Krone V, Zimmer M, Shackleton G (2011) The human ADME study. Drug discovery and evaluation: methods in clinical pharmacology, pp 73–103

Krauser J, Walles M, Wolf T, Graf D, Swart P (2012) A unique automation platform for measuring low level radioactivity in metabolite identification studies. PLoS One 7:e39070

Kruve A, Leito I, Herodes K (2009) Combating matrix effects in LC/ESI/MS: the extrapolative dilution approach. Anal Chim Acta 651(1):75–80

Lappin G, Garner RC (2005) The use of accelerator mass spectrometry to obtain early human ADME/PK data. Expert Opin Drug Metab Toxicol 1:23–31

Lappin G, Stevens L (2008) Biomedical accelerator mass spectrometry: recent applications in metabolism and pharmacokinetics. Expert Opin Drug Metab Toxicol 4:1021–1033

Lappin G, Rowland M, Garner RC (2006) The use of isotopes in the determination of absolute bioavailability of drugs in humans. Expert Opin Drug Metab Toxicol 2:419–427

Leblanc B, Jezequel S, Davies T, Hanton G, Taradach T (1998) Binding of drugs to eye melanin is not predictive of ocular toxicity. Regul Toxicol Pharmacol 28:124–132

Lenz EM, Wilson ID, Wright B, Partridge EA, Rodgers CT, Haycock PR, Lindon JC, Nicholson JK (2002) A comparison of quantitative NMR and radiolabelling studies of the metabolism and excretion of Statil (3-(4-bromo-2-fluorobenzyl)-4-oxo-3H-phthalazin-1-ylacetic acid) in the rat. J Pharm Biomed Anal 28:31–43

Liberman RG, Tannenbaum SR, Hughey BJ, Shefer RE, Klinkowstein RE, Prakash C, Harriman SP, Skipper PL (2004) An interface for direct analysis of 14C in non-volatile samples by accelerator mass spectrometry. Anal Chem 76:328–334

Liberman RG, Skipper PL, Prakash C, Shaffer CL, Flarakos J, Tannenbaum SR (2007) BEAMS Lab: novel approaches to finding a balance between throughput and sensitivity. Nucl Instr Meth Phys Res B 259:773–778

Lim H-K, Yuan C, Qiu X, Silva J, Evans DCA (2015) non-radioactive approach to investigate the metabolism of

therapeutic peptides by tagging with 127I and using ICP-MS analysis. Drug Metab Dispos 43(1):17–26. https://doi.org/10.1124//dmd.114.059774

Litherland AE (1980) Ultrasensitive mass spectrometry with accelerators. Am Rev Nucl Part Sci 30:437–473

Lockley WJS, McEwen A, Cooke R (2012) Tritium: a coming of age for drug discovery and development ADME studies. J Label Compd Radiopharm 55:235–257

Malet-Martino M, Gilard V, Desmoulin F, Martino R (2006) Fluorine nuclear magnetic resonance spectroscopy of human biofluids in the field of metabolic studies of anticancer and antifungal fluoropyrimidine drugs. Clin Chim Acta 366:61–73

Mamyrin BA (2001) Time-of-flight mass spectrometry (concepts, achievements, and prospects). Int J Mass Spectrom 206(3):251–266

Marchi I, Rudaz S, Veuthey JL (2009) Sample preparation development and matrix effects evaluation for multi-analyte determination in urine. J Pharm Biomed Anal 49(2):459–467

McEwen A (2011) Low level radioactive counting. In: Scientific presentation CropWorld London 31 October 2011

McEwen A, Wood SG, Wilson KH, Ford GR (2012) Human ADME studies: a post MIST analysis. In: Poster presentation 12th European ISSX Meeting 17–21 June 2012, the Netherlands

McEwen A, Henson CM, Wood SG (2014) Quantitative whole-body autoradiography, LC-MS/MS and MALDI for drug-distribution studies in biological samples: the ultimate matrix trilogy. Bioanalysis 6(3):377–391; Miller R (1977) Radioisotope dating with a cyclotron. Science 196(4289):489–494

Moriwaki H, Watanabe A, Arakawa R, Tsujimoto Y, Shimizu M, Noda T, Warashina M, Tanaka M (2002) Simultaneous determination of metabolites of trimethylbenzenes, dimethylbenzylmercapturicacid and dimethylhippuric acid, in human urine by solid-phase extraction followed by liquid chromatography tandem mass spectrometry. J Mass Spectrom 37:1152–1157

Naito S, Furuta S, Yoshida T, KitadaO O, Fueki T, Unno Y, Ohno H, Onodera N, Kaeamura M, Kurokawa F, Sagami K, Shinoda T, Nakazawa T, Yamazak T (2007) Current opinion: Safety evaluation of drug metabolites in development of pharmaceuticals. J Toxicol Sci 32:329–341

Nassar A-EF, Parmentier Y, Martinet M, Lee DY (2004) Liquid chromatography–accurate radioisotope counting and microplate scintillation counter technologies in drug metabolism studies. J Chromatogr Sci 42:348–353

Nedderman AN, Dear GJ, North S, Obach RS, Higton D (2011) From definition to implementation: a cross industry perspective of past, current and future MIST strategies. Xenobiotica 41:605–622

Nelson DE, Korteling RG, Stott WR (1977) Carbon-14: direct detection at natural concentrations. Science 198:507–508

Nicholson JK, Buckingham MJ, Sadler PJ (1983) High resolution 1H n.m.r. studies of vertebrate blood and plasma. Biochem J 211(3):605

Nicholson JK, Bales JR, Sadler PJ, Juul SM, Mcleod A, Sonksen PH (1984a) Monitoring metabolic disease by proton NMR of urine. The Lancet ii:751

Nicholson JK, Flynn MO, Sadler PJ, Macleod A, Juul SM, Sonkson PH (1984b) Proton-nuclear-magnetic-resonance studies of serum, plasma and urine from fasting normal and diabetic subjects. Biochem J 217:365

Nicholson JK, Timbrell J, Sadler PJ (1985) Two-dimensional proton nuclear magnetic resonance "maps" of acetaminophen metabolites in human urine. Mol Pharmacol 27:644

Pedersen-Bjergaard S, Rasmussen KE (2005) Bioanalysis of drugs by liquid microextraction coupled to separation techniques. J Chromatogr B 817:3–12

Prueksaritanont T, Lon JH, Baillie T (2006) Complicating factors in safety testing of drug metabolites: kinetic differences between generated and preformed metabolites. Toxicol Appl Pharmacol 217: 143–152

Seymour MA (2011) Accelerator MS: its role as a frontline bioanalytical technique. Bioanalysis 3:2817–2823

Smith DA, Obach RS (2005) Seeing through the MIST: abundance versus percentage. Commentary on metabolites in safety testing. Drug Metab Dispos 33:1409–1417

Smith DA, Obach RS (2006) Metabolites and safety: what are the concerns and how should we address them? Chem Res Toxicol 19:1570–1579

Suter M (2010) Challenging developments in three decades of accelerator mass spectrometry at ETH: from large particle accelerators to table size instruments. Eur J Mass Spectrometry 16(3):471–478. https://doi.org/10.1255/ejms.1078

Swart P, Lozac'h F, Zollinger M (2016) A mass balance study and metabolite profiling of sonidegib in healthy male subjects using the microtrace approach. In: Metabolite safety in drug development. Wiley, Hoboken, NJ, pp 243–260

Timmerman P, Kall MA, Gordon B, Laakso S, Freislehen A, Hucker R (2011) Bioanalysis 2:1185–1194

Ullberg S (1954) Studies on the distribution and fate of 35S-labelled benzylpenicillin in the body. Acta Radiol Suppl 118:1–110

Ure A (1841) Gouty concretions. Pharm J Transact 1:30–35

US FDA guidance for industry (2008) Safety testing of drug metabolites

US FDA guidance for industry (2016) Safety testing of drug metabolites

Vogel JS, Giacomo JA, Schulze-König T, Keck BD, Lohstroh P, Dueker S (2010) Accelerator mass spectrometry best practices for accuracy and precision in bioanalytical 14C measurements. Bioanalysis 2:455–468

Woehler F, Tiedemann F (1824) Z Physiol 1:14

World Health Organisation (1977) Use of ionising radiation and radionuclides on human beings for medical purposes. Technical report No 611, Geneva

Young GC, Ellis WJ (2007) AMS in drug development at GSK. Nucl Inst Methods Phys Res B 259:752–757

Young GC, Corless S, Felgate CC, Colthup PV (2008) Comparison of a 250 kV single-stage accelerator mass spectrometer with a 5 MV tandem accelerator mass spectrometer-fitness for purpose in bioanalysis. Rapid Commun Mass Spectrom 22:4035–4042

Zhu M, Zhao W, Vazquez N, Mitroka JG (2005) Analysis of low level radioactive metabolites in biological fluids using high-performance liquid chromatography with microplate scintillation counting: method validation and application. J Pharm Biomed Anal 39:233–245

Synthesis of Radiolabelled Compounds for Clinical Studies

44

Jens Atzrodt, Volker Derdau, and Claudia Loewe

Contents

Abstract

Regulatory requirements, quality-related measures, as well as key manufacture, control, and release aspects for the synthesis of radiolabelled drugs for administration to human volunteers as part of clinical human ADME

J. Atzrodt (✉) · V. Derdau · C. Loewe
Integrated Drug Discovery, Medicinal Chemistry, Isotope Chemistry and Metabolite Synthesis Departments (ICMS), Sanofi-Aventis Deutschland GmbH, Frankfurt a. M, Germany
e-mail: jens.atzrodt@sanofi.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_12

studies are discussed in detail. Additionally this review provides a general overview of synthetic, technical, and methodological aspects to be considered for the synthesis of Tritium- and 14C-labelled compounds. Chemical and biochemical methods and new trends for isotope labelling are discussed based on published examples.

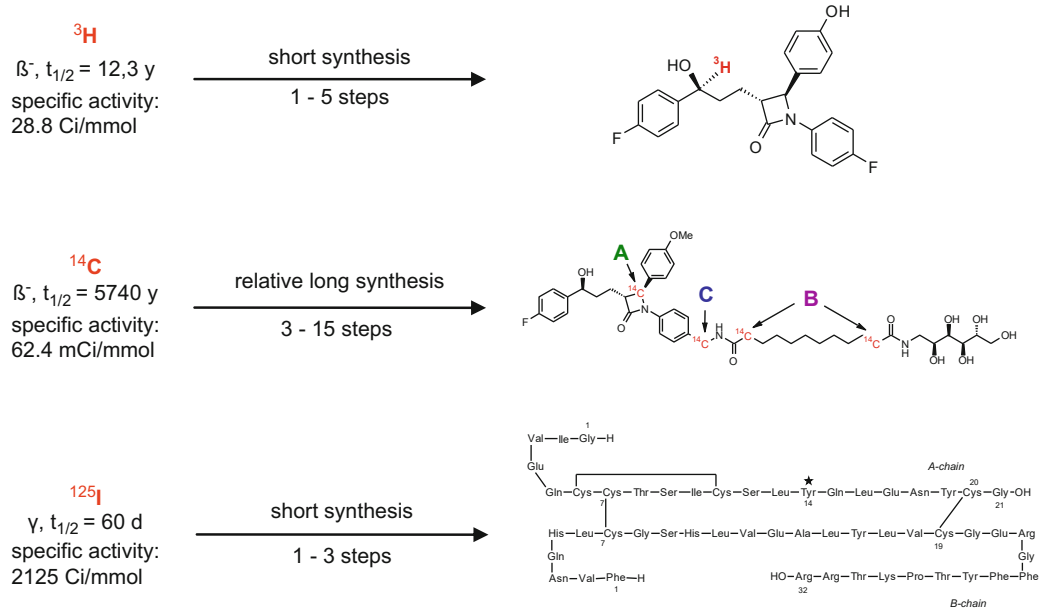
Introduction

During the development of new drugs, the candidate's pharmacokinetic (PK) properties and the absorption, distribution, metabolism, and elimination (ADME) characteristics have to be evaluated first in vitro, then in animals, and finally in humans (Caldwell et al. 1995; Roffey et al. 2007). The objectives of human ADME studies are to evaluate mass balance data and most importantly to confirm that the metabolism of the drug is similar to what was described in animal species (Deroubaix and Coquette 2004; Penner et al. 2009). In order to keep track of the drug molecules throughout the body and excreta even after their transformation into different metabolites, the administration of radiolabelled drugs is considered essential (Isin et al. 2012; Penner et al. 2012). Usually, ^{14}C is the label of choice for most drug candidates since it can be introduced into a metabolically stable position in the backbone of the compound, the detection is easy, and, in case of combustion of samples, the produced $^{14}CO_2$ can be nicely absorbed quantitatively (see Scheme 1) (Beumer et al. 2006). Generally ^{3}H-labelled drugs can be prepared more easily and more quickly than their ^{14}C counterparts. On the other hand, the ^{3}H-label is often less biologically stable, and it is more difficult to predict its metabolic stability, and therefore one always needs to bear in mind the potential risk of the in vivo formation of $^{3}H_2O$ (Lockley et al. 2012). The latter is highly toxic and can be distributed throughout the whole body which makes radioactivity measurement and quantification, even in animal studies, more difficult. Therefore, ^{3}H-labelled drug candidates are usually administered less frequently to humans and only if the specific activity of the ^{14}C compound is insufficient for the planned investigations, e.g., in case of high molecular weight and/or very low-dose drugs (Krauser 2013). For large complex biological molecules, like proteins or antibodies, ^{3}H- or ^{14}C-labelling via total synthesis may be extremely difficult or even impossible, and hence iodination with $^{125}I_2$- or a ^{125}I-precursor could be an alternative approach (Dewanjee 1992). However, the structural changes caused by an additional iodine atom in the molecule have to be considered and both materials (iodinated and non-iodinated) tested for bioequivalence. Other potential radioactive isotopes are ^{33}P and ^{35}S but compared to ^{14}C are much less frequently applied for labelling of drug candidates.

Short-lived radionuclides like ^{11}C and ^{18}F are used for PET (positron-emission tomography) (Cherry 2001) to study mostly in noninvasive experiments the drug passage over the blood-brain barrier, selective accumulation in critical organs, receptor occupancy, dose response or tumor metabolism, and proliferation rates (Rösch 2003). However, labelling syntheses using these short-lived isotopes require specific considerations which are not the subject of this chapter.

Microtracer Concept

Microdosing studies are commonly performed to collect pharmacokinetic data for drug substances following their in vivo administration at sub-pharmacological (trace) doses, a dose at which no pharmacological effect is predicted (Bae and Shon 2011). One method that has evolved significantly over the last few years for determining these data uses a labelled ^{14}C-tracer and a very sensitive detection instrument, the accelerated mass spectroscopy (AMS) technology (Vogel 2000; Vogel et al. 2007). AMS permits the option of decreasing the radioactive dose by a factor of 1,000 from around 50 μCi to about 50 nCi (Garner 2000). The very high sensitivity of AMS permits the evaluation of microdosing approaches including subtherapeutic doses (Lappin and Garner 2003). Smaller clinical doses and/or fewer radioactive administrations to humans may also change the regulatory view on the synthesis of ^{14}C-labelled compounds. The current downside of



Scheme 1 Radioactive isotopes used for labelling of drugs or drug candidates

AMS is that all samples need to be converted into solid graphite, which is an expensive process and only a few laboratories are offering this service. Consequently AMS is not yet used as a standard analytical technique. At the moment, pharmaceutical companies conducting ADME studies apply AMS only when it appears to be absolutely necessary, e.g., in case of high potency drugs, large natural products, or compounds with a very long biological half-life (Lappin et al. 2006). An additional disadvantage is that the same human samples can't be used for AMS quantification and metabolite identification.

In a recent investigation, researchers determined the absolute oral bioavailability of AZD5122 (Hickey et al. 2016). To assess these values, an intravenous (IV) and oral administration study is mandatory. In the past, absolute bioavailability studies have been conducted by both arms, and for insoluble drugs especially tremendously, efforts have been made to develop a suitable IV formulation, which itself needed further toxicological testing. By applying AMS technology into this microdosing study, the therapeutic oral dose was coadministered with a radioactive IV tracer. Due to the low dose, the IV formulation development was much simpler. By this method, very small differences between ^{12}C and ^{14}C can be quantified; however, the AMS analysis is still mainly outsourced to specialized CROs due to the high financial investment necessary. Nowadays, AMS technology is also used to study biomolecules such as insulins in human ADME studies (Salehpour et al. 2010).

In light of the recent FDA guidance "safety testing of drug metabolites," stable isotopically labelled (^{13}C, ^{15}N, ^{2}H) analogues can be used more frequently to obtain quantitative and qualitative information on drug metabolism in early human ADME studies (e.g., First in Man) even without specific studies administering radiolabelled drugs (Mutlib 2008). Modern LC/MS technologies and hyphenation of liquid chromatography with CRIMS (chemical reaction interface mass spectrometry) (Jorabchi et al. 2005) presents an opportunity to perform quantitative measurement of metabolites even in the absence of authentic standards or radiolabelled compounds (Abramson et al. 1996).

In these tracer studies, the stable isotopically labelled IV dose differs by mass from the parent drug. The plasma samples from these tracer studies are analyzed by standard HPLC-MS/MS

instruments and an isotopically labelled internal standard. By this method, the levels of the labelled IV and of the orally administered drug are determined, and the absolute oral bioavailability can be calculated (Cannady et al. 2016).

General Aspects to Be Considered for the Synthesis of ^{14}C-Labelled Compounds

Technical Considerations for ^{14}C-Labelling

Carbon-14 has many of the properties of an ideal tracer nuclide for human ADME studies (Catch 1961). Because of its very long half-life (5,730 years), it is unnecessary to correct for decay, and ^{14}C-labelled compounds can be prepared and stored for a long period if radiolytical decomposition can be minimized (see also under "Synthetic Considerations for ^{14}C-Labelling"). Carbon-14 decays to nitrogen-14 with the emission of a β-particle (maximum energy, 156 eV; average energy, 49 eV). This emission is sufficiently energetic to make measurement at moderate specific activity fairly simple but weak enough to make shielding unnecessary. The range of these soft β-particles is about 15–16 cm in air and 0–2 mm in a solid medium. This means even at very high specific activities, ^{14}C-labelled compounds can be safely handled in standard glass vessels and conventional lab equipment if reasonable radiation safety precautions are taken. However, in most countries, working with radioactive materials requires strict reporting, licenses, and/or authorizations from authorities. Local regulations may be different or stricter with respect to the handling of radioactives, to containment equipment and to lab facilities.

Synthetic Considerations for ^{14}C-Labelling

In general, any organic compound that can be synthesized can be labelled, but specific labelling of more complicated molecules may be difficult and expensive. Usually, a target directed total synthesis approach is required for ^{14}C-labelling of drug development candidates. For selecting the right labelling position, several aspects have to be carefully considered. The position of the label should be away from sites which are chemically unstable and away from sites of metabolic attack in order to ensure the label is kept in the main metabolic fragment. In cases where high specific activities are required for the planned study, a double ^{14}C-labelling or alternatively ^{3}H-labelling needs to be considered because the maximum achievable specific activity for single ^{14}C-labelled organic compounds is 62.4 mCi/mmol (2309 MBq/mmol) (see also Scheme 1 (Schulte 1966)).

As well, the development of a labelling synthesis is dependent on the availability of suitable precursors, the length and complexity of the synthetic pathway, the reliability of the process, as well as radiation safety aspects, e.g., avoidance of volatile reaction components if possible.

Planning of a ^{14}C-Synthesis

The reaction pathway developed for ^{14}C-synthesis should satisfy the following criteria (Raaen et al. 1968):

(a) Introduce the ^{14}C-label as near as possible to the last step of the synthesis.
(b) Introduce the ^{14}C-label after stereochemical resolution.
(c) Introduce the ^{14}C-label in a known position.
(d) Ensure adequate specific activity and high radiochemical purity of the product.
(e) Provide high radiochemical yields.
(f) Consider synthetic efforts for unlabelled precursor synthesis.

No special laboratory equipment is required, but preparation techniques may differ from conventional synthetic work especially when volatile radioactive precursors are handled or volatile intermediates and/or side products are expected. Further difficulties may increase as the scale of the reactions is reduced, e.g., yield, impurities, solvent content, crystallization.

A ^{14}C-labelled compound for which preparative methods are not well established should be synthesized as follows: First, the pathway is elaborated with non-radioactive material on the desired scale until a reliable process has been developed and the operator is adequately familiarized with the chemistry. Product purity should be checked by the usual chemical and physical methods. Then, the experiment is carried out at the tracer level to establish the nature of impurities and side products (e.g., volatility), to determine the yield of the desired product and to check for radio-accountability. Finally, the procedures developed are duplicated with limited amounts of the ^{14}C-labelled precursor.

Methods for ^{14}C-Syntheses

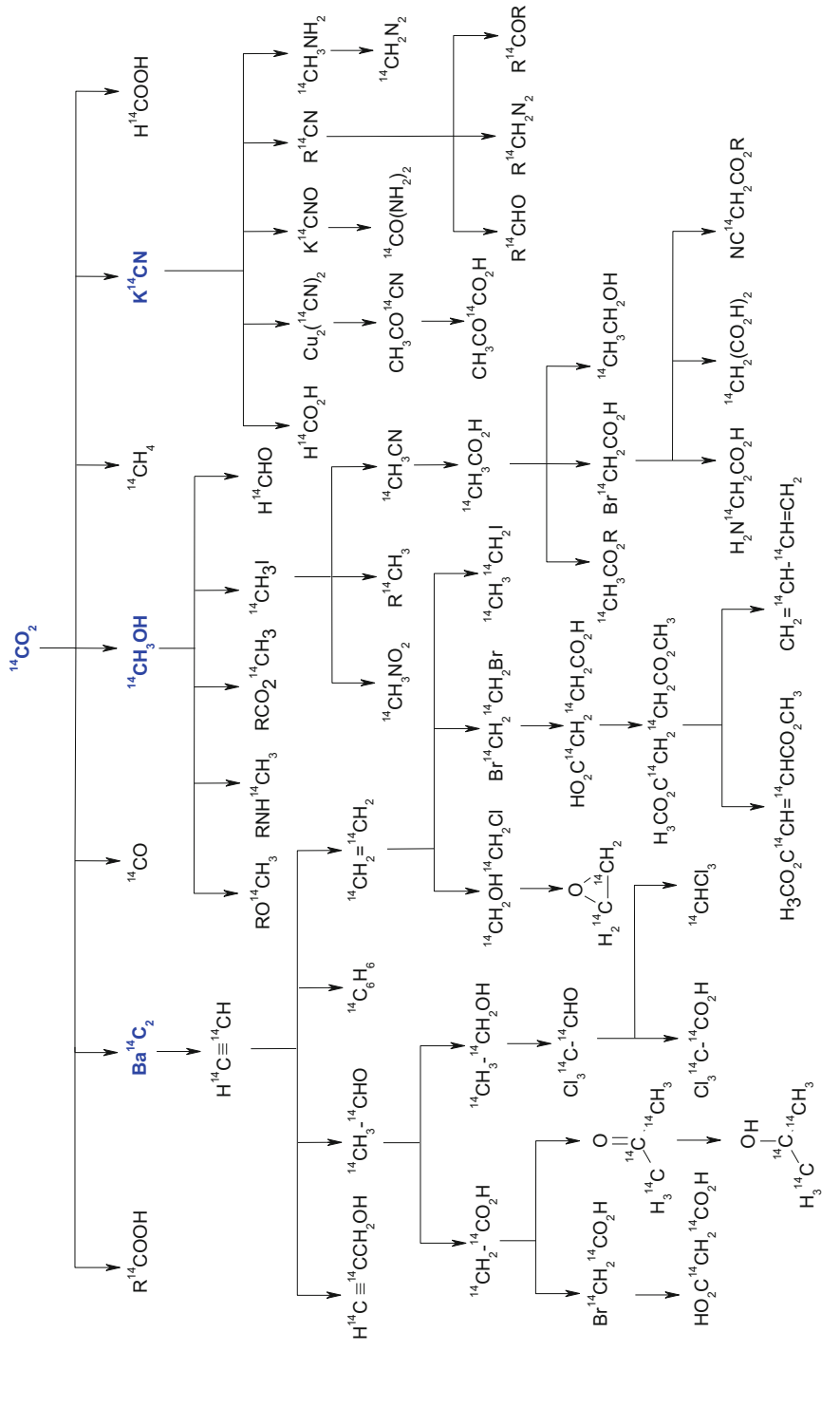
Reactor production of carbon-14 is achieved by neutron bombardment of solid beryllium nitride or solid aluminum nitride over a very long time and subsequent transformation of all ^{14}C-compounds formed into $BaCO_3$ (Wilson 1966). This means that the preparation of labelled compounds is limited to using $^{14}CO_2$ as the only practical starting material. Certain key intermediates can be used for the preparation of a great number of labelled compounds, e.g., [^{14}C]potassium cyanide, [^{14}C]barium carbide, and [^{14}C]methanol (see Scheme 2). In principle, both chemical and biological methods can be applied for converting [^{14}C]barium carbonate and its simple derivatives into more complex labelled compounds, with chemical synthesis the most generally used.

Chemical Methods

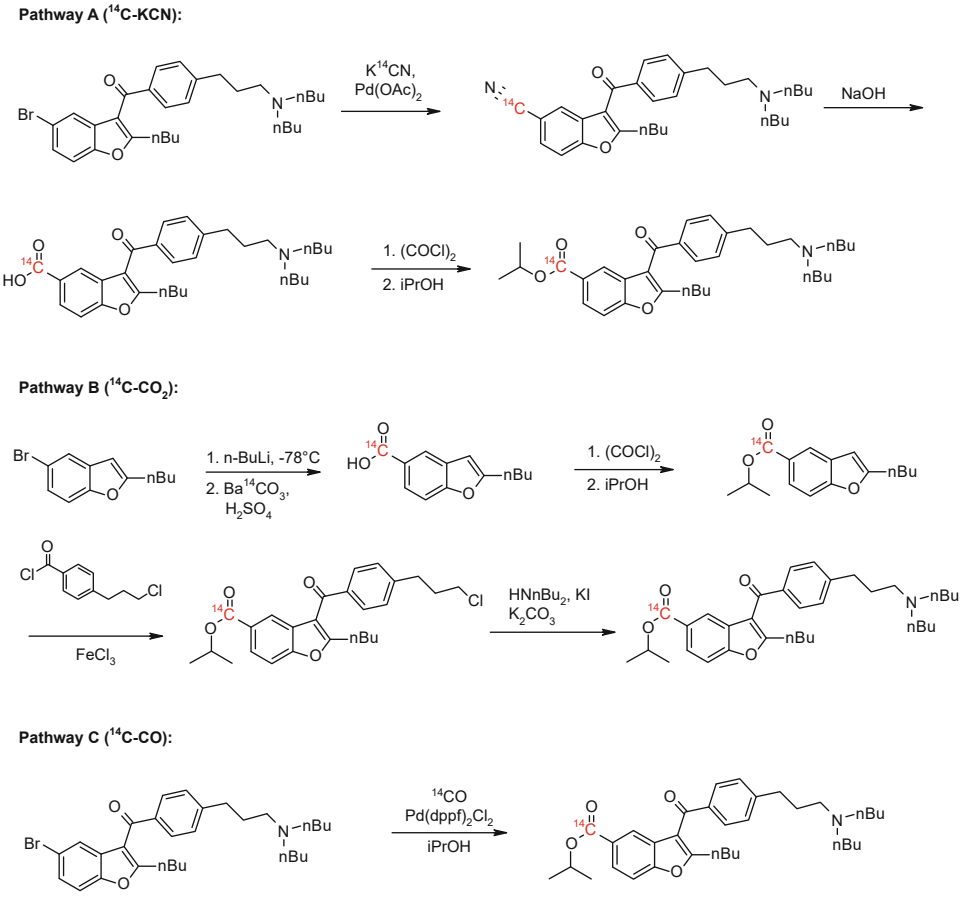
Many good standard procedures for the chemical preparation of the more common ^{14}C-labelled compounds have been summarized in literature reviews (McCarthy 2000) and in books or in book chapters (Murray and Williams 1958; Heys et al. 2009) (see also Scheme 2). Today, many specialized supply companies provide a variety of these basic ^{14}C-labelled precursors or even more complex molecules on a custom synthesis basis. Usually, the structural complexity of modern drug development candidates requires multistep labelling syntheses. Therefore, a common strategy in the pharmaceutical industry is to purchase or outsource the preparation of basic labelled organic compounds required as starting materials for an in-house synthesis of labelled drug development candidates. If applicable, a late-stage introduction of the ^{14}C-label into an unlabelled advanced intermediate can be highly efficient (see also "Planning of a ^{14}C-Synthesis," criteria (a)). For example, the Grignard reaction (Knochel 2005) of *organo*-magnesium halides with $^{14}CO_2$ or a transition metal catalyzed cyanation (Yu et al. 2017) with metal cyanide and subsequent saponification are convenient methods for introducing labelled carboxyl functionalities (Cao et al. 2007). Other small building blocks frequently used for ^{14}C incorporation into organic molecules are [^{14}C]formaldehyde, [^{14}C]methyl iodide, [^{14}C]thiocyanate, and ^{14}C-labelled acetic acid derivatives (McCarthy 2000; Heys et al. 2009).

A very good example of the advantage of a late-stage labelling strategy is the synthesis of ^{14}C-labelled celivarone. Initially, the synthesis was performed according to pathway **A** (see Scheme 3) starting with a cyanation of benzofuran derivative and subsequent synthetic construction of the molecule. However, as an alternative pathway, an approach applying lithiation and subsequent $^{14}CO_2$ addition was utilized, resulting in a longer synthesis but with improved overall yield (pathway B). Finally, by a late-stage ^{14}C-carbonylation of an unlabelled precursor (pathway **C**), the number of radioactive steps could be reduced from three to four to only a single step (Whitehead et al. 2013).

Often reducing the number of radioactive steps is a major objective of route development activities because the amount of radioactivity employed, the radioactive waste produced, and thus also the costs of a synthesis can be reduced dramatically. Besides the recognition of available labelled reagents, the accessibility of unlabelled precursors via reasonable synthetic efforts has to also be considered when planning a ^{14}C-labelling synthesis (see also "Planning of a ^{14}C-Synthesis," criteria (e)). One synthetic strategy applied to precursor synthesis is via a degradation reaction



Scheme 2 Synthetic routes to ^{14}C-labelled compounds (Raaen et al. 1968)



Scheme 3 Synthesis of [^{14}C] celivarone

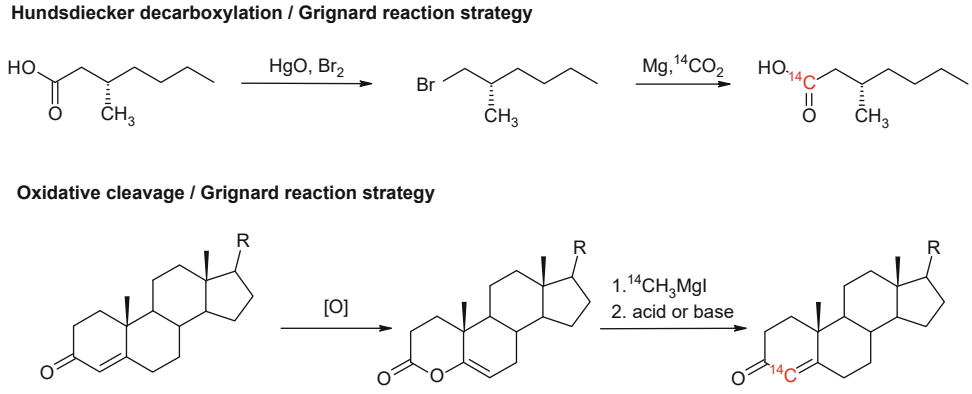
either by Hunsdiecker decarboxylation (Kurosowa et al. 1997) or oxidative cleavage (Shu and Heys 1994) as shown in Scheme 4. Subsequent labelling affords the ^{14}C-labelled version of the starting material using this protocol.

Other synthetic strategies for late-stage labelling include directed metalation or halogen-metal exchange by aromatic substitution and ^{14}C-introduction as demonstrated in the synthesis of [$^{14}C_2$] WIN-63394 ((Burgos et al. 1996); see also Scheme 5).

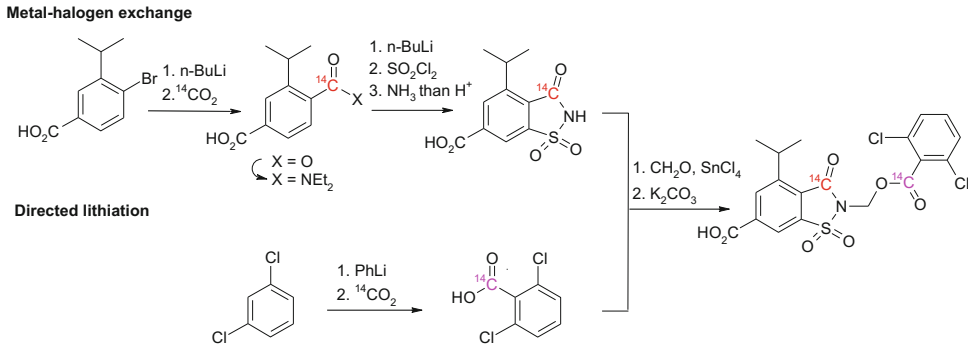
After successful introduction of the ^{14}C-label, all further chemical transformations can be performed applying classical or modern organic chemistry approaches as described for unlabelled compounds including asymmetric synthesis (Voges 2002) or cross-coupling reactions (Derdau et al. 2003).

Biochemical Methods

Besides chemical methods, biochemical synthesis offers the ability to obtain labelled natural products and organic compounds not always accessible by conventional synthesis (Evans 1981; Benakis 1994) Oligopeptides, proteins, antibodies, as well as a large number of pharmaceutically relevant compounds, e.g., antibiotics synthesized by fermentation with the aid of yeast, bacteria, or fungi, can be labelled using biochemical methods. Corresponding ^{14}C-labelled compounds can be prepared if the relevant ^{14}C-labelled starting materials, e.g., ^{14}C-labelled amino acids, are employed in the fermentation process. Yields are as important as in purely chemical syntheses, but unfortunately, only a few biochemical processes provide a single ^{14}C-labelled product in high yield (Wallace et al.



Scheme 4 Degradation/labelling strategy



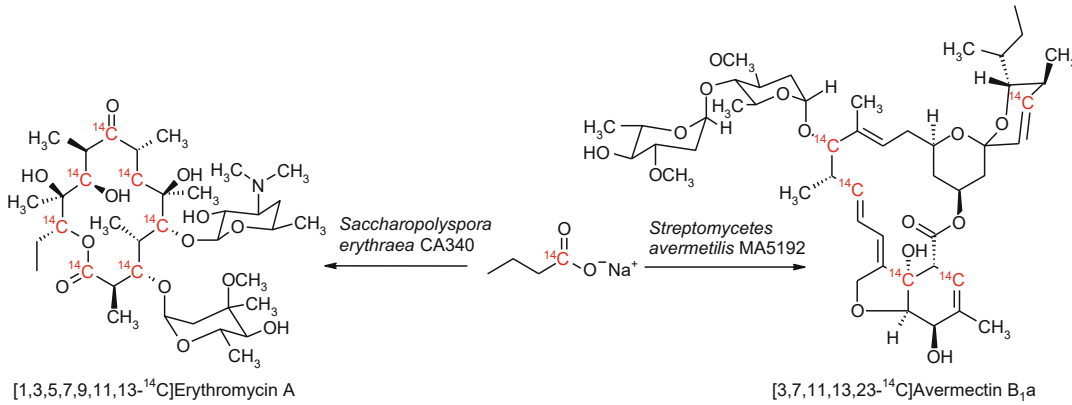
Scheme 5 Metal-halogen exchange and directed metalation strategy for ^{14}C-labelling of WIN-63394

1995). Decarboxylation and the production of large quantities of $^{14}CO_2$ might be another drawback that needs to be considered when planning fermentation for labelling purposes.

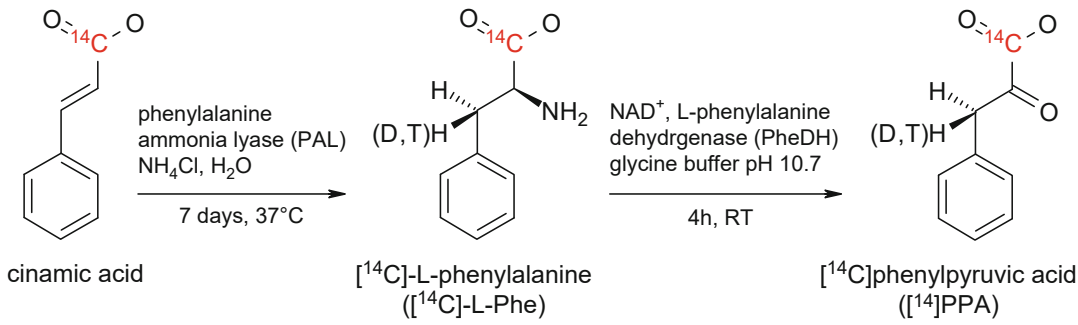
For example, [1,3,5,7,9,11,13-^{14}C]erythromycin A was produced in liquid fermentation broths of *Saccharopolyspora erythraea* CA340 in shake flasks after the administration of [1-^{14}C]sodium propionate. The labelled erythromycin A was separated by extraction of the fermentation broth and purified on Sephadex (see Scheme 6 (Walker et al. 1996)). A similar fermentation process with MA5192 (*Streptomyces avermitilis*) and [1-^{14}C] sodium propionate as precursor was also used for the synthesis of ^{14}C-labelled avermectin B_1a (Ku et al. 1984). For both compounds, a total synthesis approach for ^{14}C-label would have been very difficult, time and resource consuming, or even impossible.

On the other hand, enzymatic reactions may also be applied for specific chemical transformations, e.g., saponification, oxidation, and hydroxylation, as part of multistep conventional chemical synthesis pathways for the preparation of ^{14}C-labelled drug development candidates and corresponding ^{14}C-labelled relevant metabolites (Allen et al. 2007). As an example of biochemical labelling, the enzymatic synthesis of ^{14}C-labelled phenylpyruvic acid ([^{14}C]PPA) is shown in Scheme 7.

The addition of ammonia to [^{14}C]cinnamic acid catalyzed by the enzyme PAL (phenylalanine ammonia lyase) carried out in ammonia buffer at pH 9.8 led to the formation of [^{14}C]L-phenylalanine ([^{14}C]-L-Phe). Subsequently, this intermediate was converted to [^{14}C]PPA employing the enzyme PheDH (phenylalanine dehydrogenase) in the presence of NAD^+. As is normal for



Scheme 6 Biosynthetic labelling of erythromycin A and avermectin B_1a



Scheme 7 Biochemical ^{14}C-labelling of phenylpyruvic acid

enzymatic processes, the yield of this reaction strongly depends on the incubation time, concentration of L-Phe, buffer choice, pH, and enzyme quantity. Similarly, conducting the first step with fully deuterated or tritiated ammonia buffer gave [(*3S*)-^{2}H]-L-Phe and [(*3S*)-^{3}H]-L-Phe, respectively, which offered the opportunity to get access to the corresponding deuterium and tritium-labelled PPA as well (Skowera and Kanska 2008).

Another biochemical method for preparing labelled compounds is photosynthesis in which $^{14}CO_2$ is assimilated by, for example, algae such as *Chlorella vulgaris* (Godward 1960), cyanobacteria such as *Anacystis nidulans* (Tovey et al. 1974), or plants such as tobacco (*Nicotiana sp.)* or *Canna indica* (Putman and Hassid 1952), or by detached or full-sized plants grown in a $^{14}CO_2$ atmosphere in sealed greenhouses or plastic bags (Benkis et al. 1985). This technique called "isotope farming" can provide high specific activities and good yields of, for example, ^{14}C-labelled glucose, starch, nucleosides, amino acids, and lipids.

Stability of ^{14}C-Labelled Compounds

Compounds labelled with carbon-14 decay to nitrogen-14 with the emission of a β-particle. Since the energy of these β-rays by far exceeds bond energies of organic molecules, structural damage can occur. If the radiation energy is absorbed by the compound itself, the excited molecule may break up and/or react with other molecules. The activated decomposition fragments may also react in a sort of chain reaction with other molecules producing impurities (Bayly and Weigel 1960; Rochlin 1965). Typical reactions

resulting from irradiation are dehydrogenation, oxidation, decarboxylation, deamination, condensation, and polymerization, in many cases through radical reactions (Sheppard 1972). In some cases, the shelf life is reduced from years to weeks or even days.

The rate of decomposition depends on storage conditions, specific activity, and chemical structure. Though it is not yet possible to foresee exactly the behavior of each compound, a thorough stability study may result in storage conditions that provide reasonable shelf life of the compound and thus reduce purification efforts and secure rapid availability of stock material. The following storage rules have proved to be successful and can be applied to reduce decomposition (Evans 1976; Bayly and Evans 1966): optimize storage conditions as regards chemical stability, store at the lowest practical temperature, dilute the specific activity, avoid high amounts of activity, store labelled compounds in solution, add radical scavengers or other stabilizers (Fredenhagen 2002), and avoid unnecessary reopening of vials and warming/cooling cycles.

In spite of all potential precautionary measures, the shelf time of labelled compounds is always limited and requires repurification of the material more or less frequently (Bayly and Evans 1968).

Purification

The development of a suitable purification method for ^{14}C-labelled compounds can be crucial because typically a radiochemical purity of >98% is necessary for the planned studies (for specific applications up to 99.8% are required). Additionally, the limited shelf life of ^{14}C-labelled compounds may require repurifying the ^{14}C-labelled compound from time to time. Since synthetic impurities and degradation products can be completely structurally different, it might be necessary in some cases to develop different purification methods as well. Typically, ^{14}C-labelled compounds are purified by column chromatography, semi-prep reversed-phase HPLC or crystallization (Evans 1981).

Dilution

After purification, the ^{14}C-labelled drug development candidate often needs to be diluted to obtain the specific activity required for the planned studies. To this end, the highly radioactive compound is homogeneously mixed with unlabelled material of the same compound by dissolving both in a suitable solvent. Subsequently, the solvent is removed by evaporation, the product crystallized, precipitated, or lyophilized to afford the diluted ^{14}C-labelled drug development candidate ready for administration or further formulation to the drug product.

Analysis

After repurification and dilution of the radioactive drug substance, generally an intensive analytical release testing program according to specifications is followed to guarantee consistent product quality (Filer 1988; Dueker et al. 1998). Typically, the identity of the compound is confirmed by NMR (Schenk et al. 2015) and LC-MS; the radiochemical, chemical, and stereochemical purities are checked by analytical HPLC; and the specific activity is measured by LSC and/or high-resolution MS (Schenk et al. 2016). In addition, specific applications or specific compound-related properties may require supplementary analytical tests, e.g., ion chromatography, water content determination, polymorphism, or particle size investigations (Braun et al. 2004).

General Aspects to Be Considered for the Synthesis of Tritium-Labelled Compounds

Synthetic and Technical Considerations for ^{3}H-Labelling

Tritium-labelled compounds are much less frequently applied for human ADME studies (see also "Introduction") because an essential problem is the integrity of the carbon-tritium bond and the specificity of the labelling. There are however rare

examples for clinical trials with tritiated compounds (Straznicky et al. 2014). The high specific activity of 28.8 Ci/mmol of tritium which is approx. 500 times higher than achievable with a single carbon-14 label (62.4 mCi/mmol) is an advantage (Evans 1974). This property makes tritium irreplaceable for labelling of large molecules, particularly biomolecules such as peptides, proteins, oligonucleotides, and antibodies, as well as for early labelling strategies to support discovery purposes.

Tritium labelling is often much simpler than ^{14}C-synthesis and labelling can often be accomplished very late in the overall synthesis. The typical sources for the tritium label are tritium gas or specific tritiated reagents and sometimes tritium water. Modern stainless steel manifolds allow a safe handling and storage of tritium gas (Benakis 1994). However, as for ^{14}C-synthesis, a number of specialized supply companies offer custom tritiation services. In principle, similar points as mentioned under "Methods for ^{14}C-Syntheses" should be considered for the planning of a ^{3}H-synthesis. Compared to ^{14}C-synthesis, the scale of tritiations is even further reduced. Often, only a few milligrams of material is handled, which requires specific preparation techniques, operator training, and elaboration work. The shelf life of ^{3}H-labelled compounds is usually decreased compared to ^{14}C; however, it can be improved by storage in ethanol solution at low temperatures (<-80 °C). However, purification, dilution, and analysis aspects are similar to those already mentioned for ^{14}C-labelled materials (see also "Purification," "Dilution," and "Analysis").

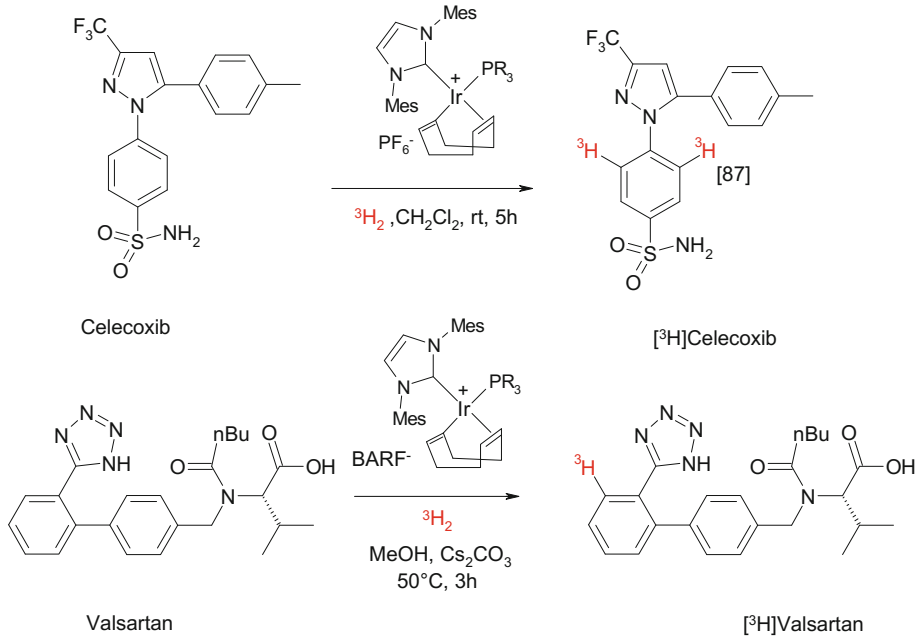
Chemical Methods for ^{3}H-Labelling

Comprehensive literature reviews summarizing synthetic techniques developed for tritium incorporation have been published recently (Heys et al. 2009; Saljoughian and Williams 2000), and therefore, only the main principles will be briefly discussed. There are two basic approaches for introducing tritium into organic molecules: exchange labelling (Atzrodt et al. 2007; Heys 2007; Lockley 2007) and synthetic tritiation methods (Evans 1981; Saljoughian 2002). While there have been several drawbacks applying exchange labelling methods in the past (e.g., lower tritium abundance or the isotope being widely distributed over the molecule), nowadays, newly developed catalysts have proven their unique efficiency for selective tritium introduction (Allen et al. 2010; Nilson and Kerr 2010). As exchange labelling can be highly cost and time efficient when carried out directly on the target molecule, it is a standard method in pharmaceutical laboratories today; however, a suitable catalyst directing group remains essential.

Excellent examples for the successful application of exchange labelling are the syntheses of tritiated celecoxib (Brown et al. 2014) and valsartan (Kerr et al. 2016). The tritium introduction was accomplished by a single, iridium-catalyzed hydrogen isotope exchange reaction with tritium gas starting directly from unlabelled celecoxib or valsartan, respectively (see Scheme 8).

In case of synthetic tritiations, the tritium is inserted into specific positions in the molecule resulting in high tritium abundances and mostly with high specific activities. Basically, beside H/T exchange reactions, four other chemical methods can be used to introduce tritium into the target molecule: (1) reduction of reducible functions (e.g., unsaturated CO, CN functions) with tritiated reagents, (2) metal catalyzed exchange of halogens by tritium, (3) hydrogenations of double or triple bonds with tritium gas, and (4) application of tritiated small molecule precursors such as methyl iodide in the labelling synthesis. Examples of the different labelling strategies are depicted in Scheme 9.

[^{3}H]Vardenafil was synthesized by reduction of a specially synthesized amide precursor with freshly prepared lithium aluminum tritide (Pleiss 2003). For the synthesis of [^{3}H]ragaglitazar at high specific activity, a suitable dihalogenated precursor was synthesized and then the tritium incorporated by a catalytic dehalogenation reaction in the presence of tritium gas (Kristensen et al. 2003). The tritiation step for the synthesis of [^{3}H]Mecillinam was performed by treating the corresponding dehydro-mecillinam, itself prepared via a six-step synthesis, with tritium gas in



Scheme 8 Synthesis of [^{3}H]celecoxib and [^{3}H]valsartan by iridium-catalyzed hydrogen isotope exchange labelling

the presence of 10% Pd/C as the catalyst (Frederiksen and Sörensen 2003). [^{3}H] Ansamitocin P-3 was prepared by alkylation of ansamitocin PDM3 with [$^3H_3$]methylnosylate under basic conditions. Ansamitocin P-3 is a precursor of LDM4 which is used to prepare drug-antibody conjugates (Sun et al. 2011). In all cases, specific unlabelled precursor molecules had to be synthesized, and therefore synthetic tritiations are often limited by the chemistry required both prior to and during the labelling process.

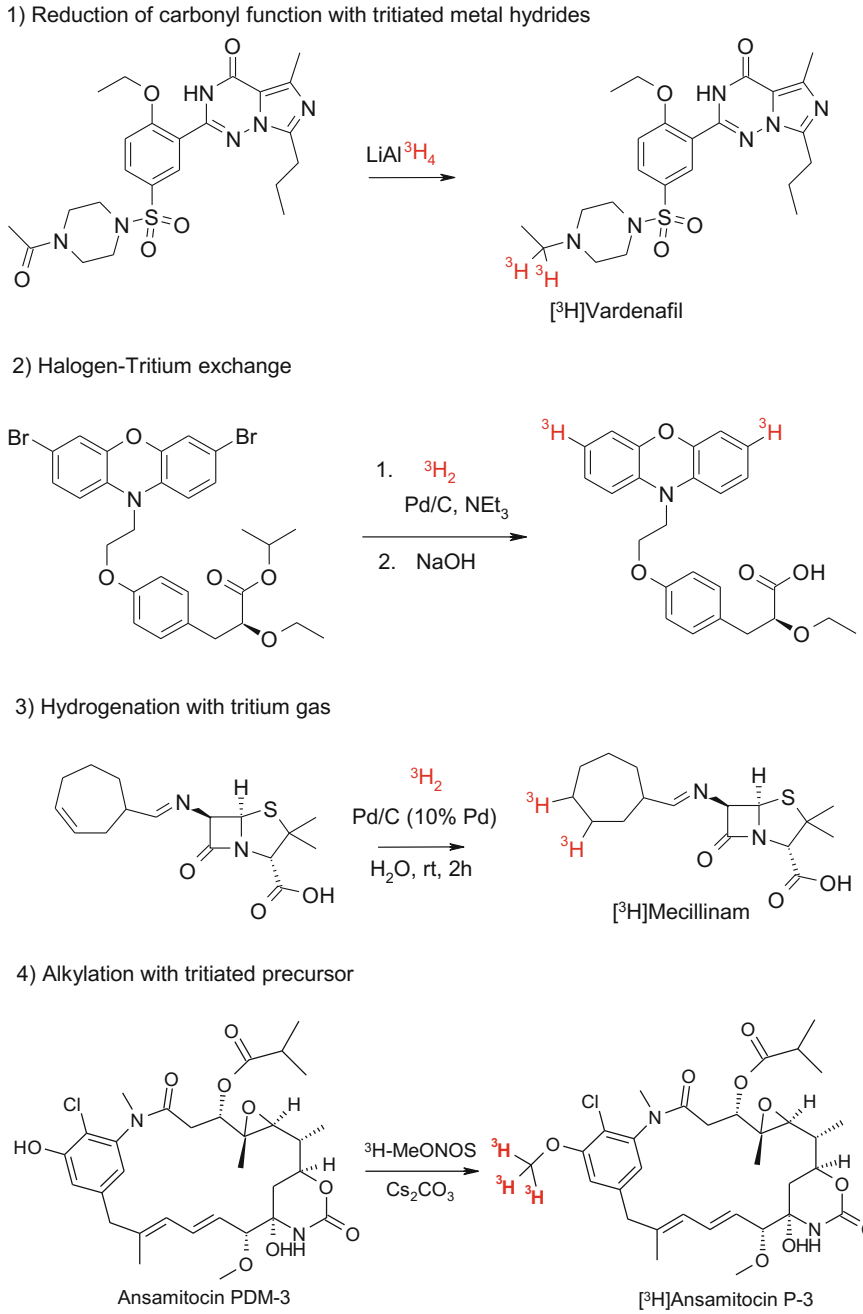
Regulatory Requirements for Application of Radiolabelled API to Humans

General Study Requirements

Administering radioactivity to human beings raises general ethical questions due to the well-established carcinogenic and/or teratogenic potential of radioactive compounds (Dain et al. 1994). Radiation exposure should be "as low as reasonably achievable" (ALARA concept), and even when administering small doses of radioactivity, given the residual risk, exposure should be minimized. Therefore, clinical studies involving the administration of radioactive drugs to humans need to be approved by a special ethical committee. Additionally, submission and approval of an Investigational New Drug Application (IND) for the USA or, if the study is carried out in the EU, an Investigational Medicinal Product Dossier (IMPD) is required. These documents include detailed information on the synthesis, medication, and analytical release, as well as a supportive stability study. The regulatory requirements for both the manufacturing process and study design are subject to various local and national authorities and strongly depend on the location and specifics of the study site and the country of the manufacturing facility.

Impurities

Based on the ICH Guidelines Q3A *Impurities in New Drug Substances* and Q6A *Specifications in New Drug Substance*, the limit for an unspecified



Scheme 9 Examples for synthetic tritiations: (1) [^{3}H]vardenafil, (2) [^{3}H]ragaglitazar, (3) [^{3}H]mecillinam, and (4) [^{3}H] ansamitocin P-3

impurity in drug substances for phase I/IIa/IIb is set at 0.1%. This threshold (see Scheme 10) is also valid for the diluted radioactive drug substance (rDS). For the highly radioactive drug substance (hrDS), the usual internal release criteria require a radiochemical purity of at least 98% or even higher. This material (hrDS) usually undergoes at least a tenfold dilution with GMP-produced

Maximum Daily Dose:	Reporting Threshold	Identification Threshold	Qualification Threshold
<= 2g/day	0.05%	0.10% or 1.0 mg per day*	0.15% or 1.0 mg per day*
		(* whichever is lower)	

Scheme 10 Thresholds of impurities in new drug substances (U.S. Department of Health and Human Services Food and Drug Administration Guidance for Industry Q3A Impurities in New Drug Substance 2008)

cold material to afford the rDS, and hence the radioactive impurity over all content is reduced below the threshold.

Regulatory Requirements

The GMP guide ICH Q7A *Good Manufacturing Practice for Active Pharmaceutical Ingredients* (*API*) does not apply to manufacturing/control aspects specific to radiolabelled compounds with long-lived isotopes (e.g., ^{14}C, ^{3}H). However, Chap. 19 contains guidance for the manufacture of APIs used in clinical trials (APIs for investigational use during early phases of development) in general. For EU countries, the EU GMP Guideline Part II *Basic Requirements for Active Substances used as Starting Materials* including Chap. 19 *APIs for Use in Clinical Trials* is recommended to be followed although not required by community legislation. During the manufacturing process, the highly radioactive drug substance is diluted with unlabelled API (manufactured according to GMP requirements) to achieve the specific radioactivity required for the planned study. Although typically only very small amounts of radioactive material are incorporated in the final drug substance, it could be considered a radioactive Active Pharmaceutical Ingredient (rAPI). Therefore, in some countries, the manufacture of radiolabelled APIs is covered by national drug laws and ordinances, whereas in others, this has not come within the scope of regulatory GMP inspections and does not require certification. These inconsistent interpretations may result in different levels of GMP (Good Manufacturing Practice) being requested for the synthesis of radiolabelled API's by different national health authorities.

Consequently, depending on national regulatory requirements, pharmaceutical companies have developed different creative approaches that combine quality principles, radiation safety aspects, and the other challenges inherent in radiosynthesis (Lloyd et al. 2003; Fontana et al. 2000). Approaches including full GMP compliance for the whole synthesis or only parts of it (e.g., only purification and dilution, including in some cases full environmental control), different GMP or GLP-like (Hong et al. 2008), or non-GMP classified processes (Bonacorsi et al. 2007) have been applied. Generally, the local authorities are allowed to inspect the manufacturing facilities to review compliance with the appropriate regulations.

General Quality-Related Measures that Should Be Applied for the Synthesis of Radioactive APIs

Independent of whether GMP is formally required or not, the synthesis of ^{14}C-labelled drug development candidates that are to be administered to humans should follow higher-quality standards than those for in vitro or animal study applications. Thereby, the stringency of quality standards should increase as the process proceeds from the early to final synthetic steps, purification, and dilution. Compared to orally administered compounds, the synthesis of radiolabelled drug substances that will be formulated for parenteral application should be subjected to more stringent control. At least the final steps, such as purification and dilution, or even the last covalent bond

modification step, should be performed as a kind of validation to confirm the correctness and quality of the applied methods. Usually all synthetic, manufacturing, analytical, and release activities are reviewed by an independent quality assurance organization to ensure the compliance of all process steps with regulatory and internal company quality requirements. SOPs (standard operation procedures) are established to define the details of all quality-related operations and processes.

Although in principle the labelled material could be any proportion of the administered drug substance (rDS), in practice, it is normally small, typically forming less than 5% of rDS.

Assuming that the hrDS forms less than 10% of the rDS, even in a worst-case scenario with a single 1% impurity in hrDS, after dilution this would end up in only 0.1% impurity of the final rDS (see also Scheme 11). Therefore, it might be appropriate to focus increased quality requirements on the purification and dilution steps only.

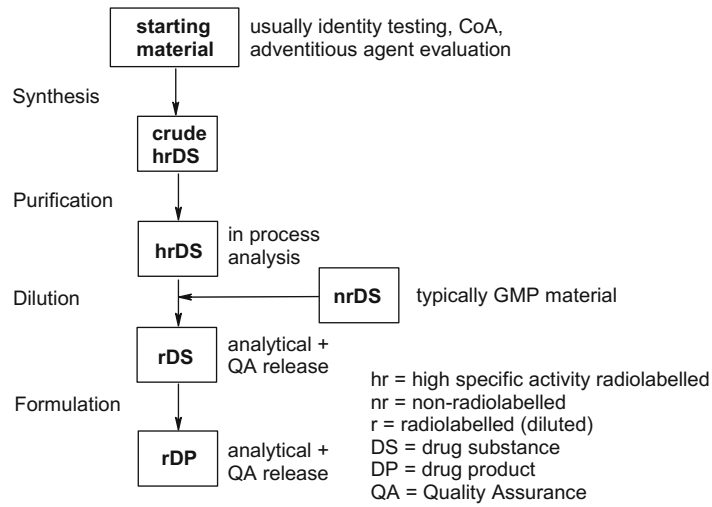
Adventitious agents evaluation and a complete review of the synthesis concerning BSE/TSE and viral safety is requested by several authorities. In addition, raw materials, intermediates, solvents and reagents, as well as materials that will come into direct contact with the radiolabelled drug (dry reagents, charcoal etc.) often have to be evaluated by testing or received with a supplier's Certificate of Analysis (CoA) and subjected to at least identity testing.

The retrievable and traceable recording of all process- and testing-related information (including signatures) in legible documents should be in place. In addition, several affiliations to combined manufacture units require written production instructions and records as well as a full quality assurance review of the synthesis documentation.

All equipment of the radiosynthesis laboratories that are critical to product quality (balances, preparative HPLCs, pH-meters) should be calibrated and serviced at appropriate intervals according to written procedures. Furthermore, the equipment used has to be qualified by a certain formal methodology in order to attest its proper technical status. This is supplemented by the regulations for computerized systems. Hence, for example, the hardware and software of an HPLC has to be evaluated regarding data security, robustness, validated operating procedures, etc.

Glassware and other equipment that will come into direct contact with the radiolabelled compound for clinical use should be new or cleaned according to standardized procedures. Stirrer bars, syringes, needles, and glass pipettes should be new.

The manufacturing facilities including fume hoods and HPLC cabinets should be dedicated and cleaned. Appropriate measures must be taken to prevent product contamination or cross-contamination. In general, a microbiological environmental monitoring of the work surfaces is



Scheme 11 Typical process steps for the synthesis and manufacture of radiolabelled drug product

required. Often due to the contradictory nature of radioprotection and GMP clean room regulations, establishing a GMP clean room facility is challenging (Loewe et al. 2016).

Another significant consideration in the manufacture of API for clinical use is the level of GMP expertise of all staff members. Their knowledge and awareness of GMP compliance needs to be kept to a high standard by continuous training on SOP and guideline compliance, in transparency and accurate documentation, and in hygiene or bio burden. For each employer, specific duties and responsibilities should be recorded in written job descriptions. Finally, a document management system has to be established including deviation and change control reporting, and last but not least, a regular interval of inspections should take place (Filer et al. 2016).

Conclusion and Outlook

Suitable quality measures should be in place in order to ensure that the manufacture, control, and release of radiolabelled compounds administered to human volunteers will satisfy quality requirements and guarantee both patient safety and the reliability of study results. In the last couple of years, the regulatory stringency for the synthesis of radiolabelled compounds for clinical use has increased, and certainly this will continue in the future. Whether radioactive tracers or the AMS technology will be used to study proteins, peptides, and other macromolecules on a frequent basis remains unclear. Up until now, there has been no requirement from the authorities to perform human ADME studies with biologicals in order to gain regulatory approval. However, this translational information might be valuable from a scientific point of view to help support milestones decisions in pharmaceutical drug development.

References and Further Reading

Abramson FP, Teffera Y, Kusmierz J, Steenwyck R, Pearson PG (1996) Replacing 14C with stable isotopes in drug metabolism studies. Drug Metab Dispos 24:697–701

Allen J, Brasseur DM, De Bruin B, Denoux M, Pérard S, Philippe N, Roy S (2007) The use of biocatalysis in the synthesis of labelled compounds, proceedings of the ninth international symposium on the synthesis and applications of isotopically labelled compounds, Edinburgh, 16–20 July, 2006. J Label Compd Radiopharm 50:342–346

Allen PH, Hickey MJ, Kingston LP, Wilkinson DJ (2010) Metal-catalysed isotopic exchange labelling: 30 years of experience in pharmaceutical R&D. J Label Compd Radiopharm 53:731–738

Atzrodt J, Derdau V, Fey T, Zimmermann J (2007) The renaissance of H/D exchange. Angew Chem Int Ed 46:7744–7765

Bae SK, Shon J-H (2011) Microdosing studies using accelerated mass spectrometry as exploratory investigational new drug trials. Arch Pharm Res 34:1789–1797

Bayly RJ, Evans EA (1966) Stability and storage of compounds labelled with radioisotopes. J Label Compd 2:1–34

Bayly RJ, Evans EA (1968) Storage and stability of compounds labelled with radioisotopes, Amersham Review No 7

Bayly RJ, Weigel H (1960) Self-decomposition of compounds labelled with radioisotopes. Nature 188:384–387

Benakis A (1994) The importance of labelling of bioactive compounds in the development of new drugs. Adv Drug Des Dev:123–136

Benkis A, Sugnaux FR, Collet FR, Kradolfer GF, Berney JP, Sion C, Necciari J, Cautreels W (1985) Carbon-14 photosynthesis labelling of natural compounds and drugs from plants. In: Synthesis and Applications of Isotopically Labeled Compounds, Proceeding of the Second International Symposium, p 219

Beumer JH, Beijnen JH, Schellens JHM (2006) Mass balance studies, with a focus on anticancer drugs. Clin Pharmacokinet 45(1):33–58

Bonacorsi SJ Jr, Burrell RC, Luke GM, Depue JS, Rinehart JK, Balasubramanian B, Christophers LJ, Iyer RA (2007) Synthesis of the anxiolytic agent [^{14}C] 6-hydroxy-buspirone for use in human ADME study. J Label Compd Radiopharm 50:65–71

Braun MP, Yao D, Wallace MA, Marques R, Jenkins HJ, Chang A, Jia X, Crocker LS, Dean DC, Melillo DG (2004) Ultrasonic particle size reduction of radiolabeled pharmaceuticals. J Label Compd Radiopharm 47:399–406

Brown JA, Cochrane AR, Irvine S, Kerr WJ, Mondal B, Parkinson JA, Paterson LC, Reid M, Tuttle T, Andersson S (2014) The synthesis of highly active iridium(I) complexes and their application in catalytic hydrogen isotope exchange. Adv Synth Catal 356:3551–3562

Burgos A, Duffin GR, Ellames GJ, Wedge KJ (1996) [^{14}C]-labelling of Benzisothiazolone based inhibitors of human leukocyte elastase. J Label Compd Radiopharm 38:193–201

Caldwell J, Gardner I, Swales N (1995) An introduction to drug disposition: the basic principle of absorption,

distribution, metabolism, and excretion. Tox Path 23(2):102–112

Cannady EA, Aburub A, Ward C, Hinds C, Czeskis B, Ruterbories K, Suico JG, Royalty J, Ortega D, Pack BW, Begum SL, Annes WF, Lin Q, Small DS (2016) Absolute bioavailability of evacetrapib in healthy subjects determined by simultaneous administration of oral evacetrapib and intravenous [$^{13}C_8$]-evacetrapib as a tracer. J Label Compd Radiopharm 59:238–244

Cao K, Bonacorsi SJ Jr, Balasubramanian B, Hanson RL, Manchand P, Godfrey JD Jr, Fox R, Christopher LJ, Su H, Iyer R (2007) Carbon-14 labeling of Saxaliptin (BMS-477118). J Label Compd Radiopharm 50:1224–1229

Catch JR (1961) Carbon-14 compounds. The Radiochemical Centre Amersham/Butterworths, Bucks/London

Cherry SR (2001) Fundamentals of positron emission tomography and applications in preclinical drug development. J Clin Pharmacol 41:482–491

Dain JG, Collins JM, Robinson WT (1994) A regulatory and industrial perspective of the use of carbon-14 and tritium isotopes in human ADME studies. Pharm Res 11(6):925–928

Derdau V, Oekonomopulos R, Schubert G (2003) [^{14}C]-labeled and large-scale synthesis of the angiotensin-(1-7)-receptor agonist AVE 0991 by cross-coupling reactions. J Org Chem 68:5168–5173

Deroubaix X, Coquette A (2004) The ins and outs of human ADME studies. Bus Brief: Pharmatech:1–4

Dewanjee MK (1992) Radioiodination; theory, practice, and biomedical application. Kluwer Academic, Boston

Dueker SR, Jones AD, Clifford AJ (1998) Protocol development for biological tracer studies. Adv Exp Med Biol 445:363–378

Evans EA (1974) Tritium and its compounds: second edition. Wiley, London

Evans EA (1976) Guide to the self-decomposition of radiochemicals, Amersham Review No 16, The Radiochemical Centre, Amersham; booklet Amersham International plc. 1992; booklet Amersham Biosciences 2002

Evans EA (1981) Synthesis of radiolabelled compounds. J Radional Chem 64(1–2):9–32

Filer CN (1988) The analysis of ^{14}C and ^{3}H-labelled compounds. Synth Apl Isot Labelled Cpd

Filer C, Maniscalco M, Thayer S (2016) GMP synthesis of carbon-14 labelled substances. J Label Compd Radiopharm 59:233–237

Fontana E, Dellavedova P, Gambini L (2000) Carbon-14 and tritium labelled compounds good manufacturing practices for studies in humans. In: Pleiss U, Voges R (eds) Proceedings of the seventh IIS symposium Dresden, Synth. Apl. Isot. Labelled Cpd. Wiley, Chichester/New York/Weinheim/Brisbane/Singapore/Toronto, pp 499–502

Fredenhagen A (2002) Radiochemical stability of ^{14}C-labelled compounds on storage: benefits of Thioethers. J Label Compd Radiopharm 45:211–220

Frederiksen SM, Sörensen GG (2003) Synthesis of tritium labelled mecillinam. J Label Compd Radiopharm 46:773–779

Garner RC (2000) Accelerated mass spectrometry in pharmaceutical research and development – a new analytical method for isotope measurement. Curr Drug Metab 1:205–213

Godward MBE (1960) Resistance of algae to radiation. Nature 185:706

Heys JR (2007) Organoiridium complexes for hydrogen isotope exchange labelling. J Label Compd Radiopharm 50:770–778

Heys JR, Voges R, Moenius T (2009) Preparation of compounds labelled with tritium and carbon-14, John Wiley & Sons, Inc. New York.

Hickey MJ, Allen PH, Kingston LP, Wilkinson DJ (2016) The synthesis of [^{14}C]AZD5122. Incorporation of an IV ^{14}C-microtracer dose into a first in human study to determine the absolute oral bioavailability of AZD5122. J Label Compd Radiopharm 59:245–249

Hong Y, Bonacorsi SJ Jr, Tian Y, Gong S, Zhang D, Humphreys WG, Balasubramanian B, Cheesman EH, Zhang Z, Caster JF, Crane PD (2008) Synthesis of [1,2-^{3}H]ethylamine hydrochloride and [^{3}H]-labelled Apadenoson for a human ADME study. J Label Compd Radiopharm 51:113–117

Isin EM, Elmore CS, Nilsson GN, Thompson RA, Weidolf L (2012) Use of radiolabeled compounds in drug metabolism and pharmacokinetic studies. Chem Res Toxicol 25:532–542

Jorabchi K, Kahen K, Lecchi P, Montaser A (2005) Chemical reaction interface mass spectroscopy with high efficiency nebulisation. Anal Chem 77:5402–5406

Kerr WJ, Lindsay DM, Reid M, Atzrodt J, Derdau V, Rojahn P, Weck R (2016) Iridium-catalysed *ortho*-H/D and -H/T exchange under basic conditions: C-H activation of unprotected Tetrazoles. Chem Commun 52:6669–6672

Knochel P (2005) Handbook of functionalized organometallics, applications and synthesis. Wiley-VCH, Weinheim.

Krauser JA (2013) A perspective on tritium versus carbon-14: ensuring optimal label selection in pharmaceutical research and development. J Label Compd Radiopharm 56:441–446

Kristensen JB, Johansen SK, Valsborg JS, Martiny L, Foged C (2003) Ragaglitazar[14C] and [3H]−labeling of ragaglitazar: a dual acting PPARα and PPARγ agonist with hypolipidemic and anti-diabetic activity. J Label Compd Radiopharm 46:475–488

Ku CC, Hwang SC, Kaplan L, Nallin MK, Jacob TA (1984) The preparation of carbon-14 labeled Avermectin B_1a. J Label Compd Radiopharm 22:451–459

Kurosowa M, Kanamuru N, Nishioka K (1997) [^{14}C]-labeling of novel prostacyclin I, derivative SM10902. J Label Compd Radiopharm 39:129–138

Lappin G, Garner RC (2003) Big physics, small, doses; the use of AMS and PET in human microdosing of development drugs. Nat Rev Drug Discov:233–240

Lappin G, Garner RC, Meyers T, Powell J, Varley P (2006) Novel use of accelerated mass spectrometry for the quantification of low levels of systemic therapeutic

recombinant protein. J Pharm Biomed Anal 41:1299–1302
Lloyd C, Potwin J, Wright C (2003) Radiosynthesis conducted under cGMP compliance. In: Dean DC, Filer CF, McCarthy KE (eds) Proceedings of the eight IIS symposium Boston, Synth. Apl. Isot. Labelled Cpd. Wiley, Chichester, pp 251–254
Lockley WJS (2007) 30 years with ortho-directed hydrogen isotope exchange labelling. J Label Compd Radiopharm 50:779–788
Lockley WJS, McEwen A, Cooke R (2012) Tritium: a coming of age for drug discovery and development ADME studies. J Label Compd Radiopharm 55:235–257
Loewe C, Atzrodt J, Reschke K (2016) Joe Schofield, conception, realization and qualification of a radioactive clean room lab facility dedicated to the synthesis of radiolabelled API for human ADME studies. J Label Compd Radiopharm 59:611–614
McCarthy KE (2000) Recent advances in the design and synthesis of carbon-14 labelled pharmaceuticals from small molecule precursors. Curr Pharm Dess 6:1057–1083
Murray A, Williams D (1958) Organic syntheses with isotopes. Interscience Publishers, New York. Muccino RR (1983) Organic syntheses with carbon-14. A Wiley-Interscience Publication
Mutlib AE (2008) Application of stable isotope-labelled compounds in metabolism and in metabolism-mediated toxicity studies. Chem Res Toxicol 21:1672–1689
Nilson GN, Kerr WJ (2010) The development and use of novel iridium complexes as catalysts for ortho-directed hydrogen isotope exchange reactions. J Label Compd Radiopharm 53:662–667
Penner N, Klunk LJ, Prakash C (2009) Human radiolabeled mass balance studies: objectives, utilities and limitations. Biopharm Drug Dispos 30:185–203
Penner N, Xu L, Prakash C (2012) Radiolabeled absorption, distribution, metabolism, and excretion studies in drug development: why, when, and how? Chem Res Toxicol 25:513–531
Pleiss U (2003) Synthesis of [^{3}H]vardenafil, Levitra®, using a new labelling technique. J Label Compd Radiopharm 46:1241–1247
Putman EW, Hassid WZ (1952) Isolation and purification of radioactive sugars by means of paper chrommatography. J Biol Chem 196:749–752
Raaen VF, Ropp GA, Raaen HP (1968) Carbon-14 in series in advanced chemistry. McGraw-Hill Book, New York/St. Louis/San Francisco/Toronto/London/Sydney
Rochat R, Lopez MJ, Tsurugi H, Mashima K (2016) Recent developments in homogeneous organomagnesium catalysis. Chem Cat Chem *8*(1):10–20
Rochlin P (1965) Self-decomposition of carbon-14-labelled organic compounds. Chem Rev 65:685–696
Roffey SJ, Obach RS, Gedge JI, Smith DA (2007) What is the objective of the mass balance study? A retrospective analysis of data in animal and human excretion studies employing radiolabelled drugs. Drug Metab Rev 39:17–43
Rösch F (2003) Radiochemistry and radiopharmaceutical chemistry in life sciences. In: Vertes A, Nagy S, Klencsar Z (eds) Handbook of nuclear chemistry, Vol 4, Kluwer Academic Publisher, Berlin.
Salehpour M, Ekblom J, Sabetsky V, Hakansson K, Possnert G (2010) Accelerator mass spectrometry offers new opportunities for microdosing of peptide and protein pharmaceuticals. Rapid Commun Mass Spectrom 24:1481–1489
Saljoughian M (2002) Synthetic tritium labeling: reagents and methodologies. Synthesis 13:1781–1801
Saljoughian M, Williams PG (2000) Resent developments in tritium incorporation for radiotracer studies. Curr Pharm Des 6:1029–1056
Schenk DJ, Dormer PG, Hesk D, Pollack SR, Flader Lavey C (2015) NMR-based approach to the analysis of radiopharmaceuticals: radiochemical purity, specific activity, and radioactive concentration values by proton and tritium NMR spectroscopy. J Label Compd Radiopharm 58:291–298
Schenk DJ, Lockley WJS, Elmore CS, Hesk D, Roberts D (2016) Determining the isotopic abundance of a labeled compound by mass spectrometry and how correcting for natural abundance distribution using analogous data from the unlabeled compound leads to a systematic error. J Label Compd Radiopharm 59:136–146
Schulte HR (1966) Synthesen ^{14}C-markierter organischer Verbindungen in Radioactive Isotope in der Organischen Chemie und Biochemie. Wiley-VCH, Weinheim.
Sheppard G (1972) The self-decomposition of radioactively labelled compounds. At Energy Rev 12:3–66
Shu AYL, Heys JR (1994) Synthesis of carbon-14 and tritiated steroidal Sa-reductase inhibitors. J Label Compd Radiopharm 34:578–596
Skowera K, Kanska M (2008) Enzymatic synthesis of phenylpyruvic acid labelled with deuterium, tritium, and carbon.14. J Label Compd Radiopharm 51:321–324
Straznicky NE, Grima MT, Sari CI, Eikelis N, Lambert GW, Nestel PJ, Karapanagiotidis S, Wong C, Richards K, Marusic P, Dixon JB, Schlaich MP, Lambert EA (2014) A randomized controlled trial of the effects of pioglitazone treatment on sympathetic nervous system activity and cardiovascular function in obese subjects with metabolic syndrome. J Clin Endocrinol Metab 99 (9):E1701–E1707
Sun X, Widdison W, Mayo M, Wilhelm S, Leece B, Chari R, Singh R, Erickson H (2011) Design of antibody-maytansinoid conjugates allows for efficient detoxification via liver metabolism. Bioconjug Chem 22:728–735
Tovey KC, Spiller GH, Oldham KG, Lucas N (1974) A new method for the preparation of uniformly [^{14}C]-labelled compounds by using *Anacystis nidulans*. Biochem J 142:47–56
U.S. Department of Health and Human Services Food and Drug Administration Guidance for Industry Q3A Impurities in New Drug Substance (2008)
Vogel JS (2000) Accelerated mass spectrometry for human biochemistry: the practice and the potential. Nucl Instr Meth Phys Res B172:884–891

Vogel JS, Palmblad NM, Ognibene T, Kabir MM, Buchholz BA, Bench G (2007) Biochemical paths in humans and cells: frontiers of AMS bioanalysis. Nucl Instr Meth Phys Res B259:745–751

Voges R (2002) From chiral bromo[13,14,Cn]acetyl sultams to complex molecules singly/multiply labelled with isotopic carbon. J Label Compd Radiopharm 45:867–897

Walker AE, Pothuluri JV, Heinze TM, Volmer D, Cerniglia CE (1996) Biosynthetic production of ^{13}C and ^{14}C Erythronolide labeled erythromycin A. J Label Compd Radiopharm 29:59–67

Wallace MA, Dean DC, Ellsworth RL, Marks T, Mellio DG, Maseurekar P (1995) Studies on the biosynthesis of Avermectins using carbon labelled 2-methyl butyric acids. In: Allen J, Voges R (eds) Synthesis and applications of isotopically labelled compounds 1994. Wiley, Chichester/New York. Paper 108, 605

Whitehead DM, Hartmann S, Ilyas T, Taylor KR, Kohler AD, Ellames GJ (2013) A convenient method to produce [^{14}C]carbon monoxide and its application to the radiosynthesis of [carboxyl- ^{14}C]celivarone, [carboxyl-14C]SSR149744. J Label Compd Radiopharm 56:36–41

Wilson RJ (1966) The radiochemical manual, 2nd edn. The Radiochemical Centre, Amersham

Yu J-T, Teng F, Cheng J (2017) The construction of X-CN (X=N, S, O) bonds. Adv Syn Cat 359(1):26–38

Drug–Drug Interaction Studies

45

Peter Stopfer

Contents

P. Stopfer (✉)
Boehringer Ingelheim Pharma GmbH & Co. KG,
Ingelheim, Germany
e-mail: peter.stopfer@boehringer-ingelheim.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_13

Abstract

Drug–drug interaction (DDI) can result when one drug alters the pharmacokinetics of another drug or its metabolites. The assessment of pharmacokinetic DDIs during clinical development is a part of the general clinical pharmacology and safety assessment of a new investigational compound. Market withdrawals of drugs were frequently caused by DDIs which underlines the importance of addressing these issues during drug development. This is also reflected by the latest DDI (DDI) guidelines from European Medicines Agency EMA (2012), Food and Drug Administration (FDA) (2012), and Pharmaceuticals and Medical Devices Agency (2014). The details of all aspects which have to be considered in the design of DDI studies are outlined in the respective guidelines from EMA (2012), FDA (2012), and PMDA (2014). This section is aiming to give a summary of the respective considerations of these guidelines for the design of DDI studies and also contains many aspects of the respective guidelines including the most relevant decision trees and tables.

This chapter:

- Describes how an evaluation of DDI is performed from in vitro to in vivo studies within clinical development
- Reflects recent recommendations by authorities as regards the design, conduct, and reporting of DDI studies
- Presents the requirement to assess the clinical significance of DDIs

Abbreviations

ABC	ATP-binding cassette
AhR	Aryl hydrocarbon receptor
AUC	Area under the plasma concentration-time curve
BCRP	Breast cancer resistance protein
BSEP	Bile salt export pump
CAR	Constitutive androstane receptor
CYP	Cytochrome P450
FMO	Flavin monooxygenase
MAO	Monoamine oxidase
MATE	Multidrug and toxin extrusion
MRP	Multidrug resistance-associated protein
NTR	Narrow therapeutic range
OAT	Organic anion transporter
OATP	Organic anion transporting polypeptide
OCT	Organic cation transporter
PBPK	Physiologically-based pharmacokinetic
PD	Pharmacodynamics
P-gp	P-glycoprotein
PK	Pharmacokinetics
PXR	Pregnane X receptor
SLC	Solute carrier
TDI	Time-dependent inhibition
UGT	Uridine diphosphate (UDP)-glucuronosyl transferase
XO	Xanthine oxidase

General DDI Considerations Inclusive Comparison of the DDI Guidelines EMA/FDA/PMDA

The main focus of this chapter is on pharmacokinetic DDIs. The understanding of the nature and magnitude of DDI is important for several reasons. Concomitant medications, dietary supplements, and some foods, such as grapefruit juice, may alter metabolism and/or drug transport abruptly in individuals who previously had been receiving and tolerating a particular dose of a drug. A respective alteration in metabolism or transport can change the known safety and efficacy of a drug. In a few cases, consequences of an interaction have led to the conclusion that the drug could not be marketed safely. Several drugs have been withdrawn from the market because of significant DDIs that led to, e.g., QT prolongation and Torsades de Pointes (TdP) arrhythmias, after warnings in drug labels did not adequately manage the risk of DDIs.

The overall objective of DDI studies for a new drug is to determine:

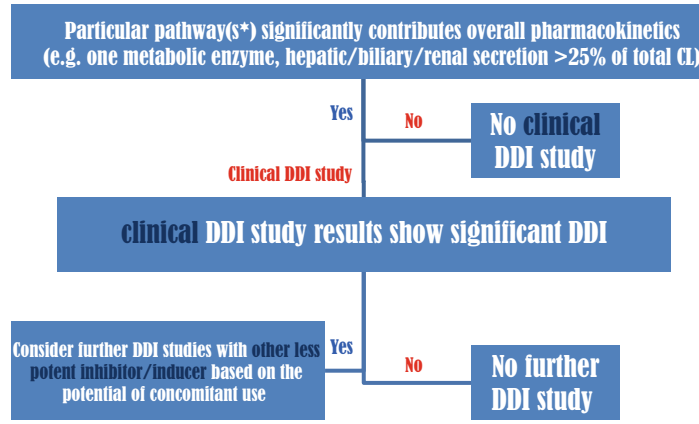


Fig. 1 General DDI decision tree for a drug as a victim for a DDI (*FDA requests the investigation on multiple enzymes)

- Whether any DDIs are sufficiently large to necessitate a dosage adjustment of the drug itself or of the drugs with which it might be used
- Whether any DDI calls for additional therapeutic monitoring
- Whether there should be a contraindication to concomitant use when lesser measures cannot mitigate risk

Therefore, the development of an investigational drug should include identification of the principal routes of elimination, quantitation of the contribution by enzymes and transporters to drug disposition, and characterization of the mechanism of DDIs.

The study of DDIs for a new drug generally begins with in vitro studies to determine whether a drug is a substrate, inhibitor, or inducer of metabolizing enzymes. The results of in vitro studies will inform the nature and extent of in vivo studies that may be required to assess potential interactions. Along with clinical pharmacokinetic data, results from in vitro studies may serve as a screening mechanism to rule out the need for additional in vivo studies, or provide a mechanistic basis for proper design of clinical studies.

Human clinical studies to assess DDIs are either designed for a dedicated administration of a perpetrator and a victim drug or may contain simultaneous administration of a mixture of substrates of multiple CYP enzymes and transporters in one study (i.e., a "cocktail approach") to evaluate a drug's inhibition or induction potential.

All three DDI Guidelines (EMA-CHMP 2012; US-FDA 2012; PMDA. Pharmaceuticals & Medical Device Agency-Japan 2014) are displaying general decision trees as to when a drug has to be considered for a DDI trial, when the drug is a **victim** for an enzyme or drug transporter (Fig. 1)

Additionally also decision trees can be found in all three DDI Guidelines (EMA-CHMP 2012; US-FDA 2012; PMDA. Pharmaceuticals & Medical Device Agency-Japan 2014) which display as to when a drug has to be considered for a DDI trial, when the drug is acting as a **perpetrator** for an enzyme or drug transporter DDI (Fig. 2).

Additionally, complex DDIs, which can occur in specific populations (e.g., patients with organ impairment, and pediatric and geriatric patients), should be considered on a case-by-case basis. Moreover also PK modeling approaches (if well verified for intended purposes) can be helpful to guide the determination of the need to conduct specific DDI studies or even to avoid respective DDI studies in special cases.

Nonclinical Assessment of DDI by In Vitro Investigations: Determining If a Drug Is a Victim or Perpetrator of a Potential DDI

The drug development process should include evaluation of a new drug's potential to affect the metabolism or transport of other drugs and the potential for the new drug's metabolism or

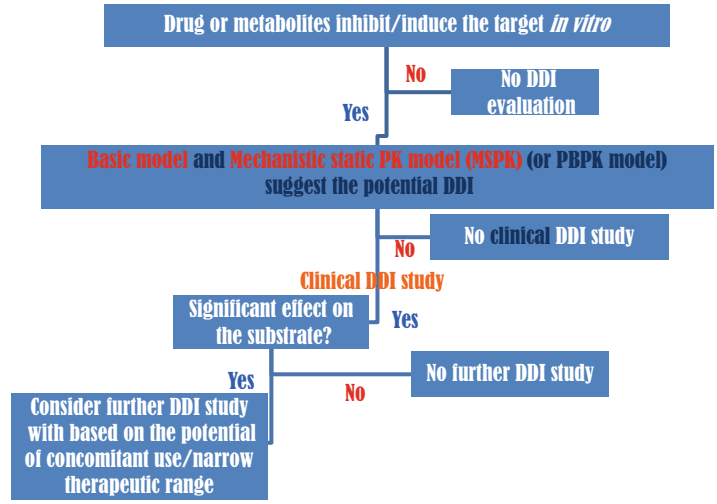


Fig. 2 General DDI decision tree for a drug acting as a perpetrator for a DDI

transport to be affected by other drugs. Use of in vitro tools to determine whether a drug is a substrate, inhibitor, or inducer of metabolizing enzymes or drug transporters, followed by in vivo interaction studies to assess potential interactions, has become an integral part of drug development and regulatory review. These results will be the basis for the determination of a clinical DDI study is needed or not. Authorities may consider in vitro data sufficient to exclude DDI liabilities if there is a clear indication from parameters outlined in regulatory documents (e.g., $[I]/K_i$ for CYP-based interactions in the absence of liver partitioning) of little or no DDI potential. These negative in vitro data can obviate the need for further in vivo clinical activities considering the pathways that were excluded from being clinically meaningful.

This section will separately discuss in vitro investigations at the levels of metabolizing enzymes and transporters. Also general considerations for situations when complex or multiple DDI mechanisms will be presented will be briefly described.

Metabolism-Based DDIs

Hepatic metabolism occurs primarily through the cytochrome P450 family (CYP) of enzymes located in the hepatic endoplasmic reticulum, but may also occur through non-CYP enzyme systems, such as glucuronosyl- and sulfo-transferases, which can, in general, inactivate a drug and increase its renal elimination. Some drug metabolizing enzymes are present in the gut wall and other extrahepatic tissues, in addition to the liver. Many metabolic routes of elimination can be inhibited or induced by concomitant drug treatment. Metabolic DDIs can cause substantial changes (an order of magnitude or more decrease or increase in the blood and tissue concentrations of a drug or metabolite) and can also significantly affect the extent to which toxic or active metabolites are formed. These large changes in exposure can alter the safety and efficacy profile of a drug and its active metabolites, regardless of whether the drug has a narrow therapeutic range (NTR).

Nonclinical in vitro experiments provide data on DDI potentially mediated by CYP enzymes (or others like (Uridyl diphosphate (UDP))-glucuronosyltransferases) using systems such as human liver microsomes, expressed human recombinant enzymes, and hepatocytes. These data are used in order to determine the substrate specificity of an NCE for a specific metabolic enzyme, the inhibitory potential (half maximal inhibitory concentration (IC_{50}), inhibition constant (K_i) for competitive inhibitors, and rate of enzyme inactivation (k_{inact}) for mechanism-based inhibitors) or induction in case of enzyme inducers (rate of metabolism). The evaluation of CYP enzyme induction may begin with studies of CYP1A2, CYP2B6, and

Table 1 Targets to be examined: whether the drug is a **substrate** or not

	FDA	EMA	PMDA
Metabolic enzymes			
CYPs [recommendation]	CYP1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 3A [2A5, 2 J12,4F2, 2E1]	CYP1A2, 2B5, 2C8, 2C9, 2C19, 2D6, 3A	CYP1A2, 2B5, 2C8, 2C9, 2C19, 2D6, 3A [2A5, 2 J12,4F2, 2E1]
UGTs	UGT1A1, 1A3, 1A4, 1A6, 1A9, 2B7, 2B15[b]	Not specified	UGT1A1, 1A3, 1A4, 1A5, 1A9, 2B7, 2B15[b]
Others [recommendation]	MAO, FMO, XO, ALDH, ADH	Not specified	MAO, FMO, XO, AO, ALDH, ADH
Transporters			
Gut and systemic	P-gp, BCRP	Not specified	P-gp, BCRP
Hepatic ($CL_H > 25\%$ of CL_{tot})		CATP1B1/1B3[a]	
Renal ($CL_{R,\ secretion} > 25\%$ of CL_{tot})	OCT2, OAT1/3	Not specified	OCT2, OAT1/3, MATE1/2-K

[a]In case the autoradiography (ARG) in animal shows significant accumulation in the liver (PMDA)

[b]Require the identification of UGT subtype

[c]In case the drug is not metabolized by major CYPs

Table 2 Targets to be examined: whether a drug is an **inhibitor** or not

	FDA	EMA	PMDA
Metabolic enzymes			
CYPs	CYP1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 3A	CYP1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 3A*	CYP1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 3A*
UGTs	UGT1A1, 1A3, 1A4, 1A6, 1A9, 2B7, 2B15	UGT1A1, 2B7	UGT1A1, 2B7
Transporters			
Gut and systemic	P-gp, BCRP		
Hepatic ($CL_H > 25\%$ of CL_{tot}) [recommendation]	OA1P1B1/1B3 [MRPs, BSEP]	OATP1B1/1B3, [BSEP, OCT1]	OATP1B1/1B3, [MRP2, BSEP, OCT1]
Renal ($CL_{R,\ secretion} > 25\%$ of CL_{tot}) [recommendation]	OC12, OAT1/3 [MRPs, MA1E1/2-K]	OCT2,OAT1/3 [MATE1/2-K]	OCT2,OAT1/3, MATE1/2-K, [MRP2, 4]

*For CYP3A4, investigation with multiple substrate with different binding site are required

CYP3A in vitro. If the in vitro induction results are positive according to predefined thresholds using basic models, the investigational drug is considered an enzyme inducer and further in vivo evaluation may be warranted

An overview as regards which metabolic enzymes to be examined based on FDA, EMA, and PMDA guideline, e.g., whether a drug is a substrate, inhibitor, or inducer or not is given in Tables 1, 2, and 3.

Transporter-Based DDIs

Although less well-recognized than metabolizing enzymes, membrane transporters can have

Table 3 Targets to be examined: whether the drug is an **inducer** or not

	FDA	BVIA	PMDA
Metabolic enzymes			
Enzyme (transcriptional factor)	CYP3A4/5 (PXR),[if positive, CYP2C8, 2C9, 2C19]CYP1A2 (AhR),CYP2B6 (CAR)	CYP3A4/5 (PXR), CYP1A2 (AhR), CYP2B6 (CAR)	CYP3A4/5 (PXR), [if positive, CYP2C9 et al] CYP1A2 (AhR), CYP2B6 (CAR)
Transporter			
	"Methods for in vitro evaluation are not well understood" "Should consult with FDA about studying induction in vivo"	P-gp (in case PXR and/or CAR mediated induction observed)	Not mentioned

important effects on pharmacokinetics and drug exposure. To date, most identified transporters belong to one of two superfamilies: ATP-binding cassette (ABC) and solute carrier (SLC). Transporters govern the transport of solutes (e.g., drugs and other xenobiotics) in and out of cells. In contrast to metabolizing enzymes, which are largely concentrated in the liver and intestine, transporters are present with varying abundance in all tissues in the body and play important roles in drug distribution, tissue-specific drug targeting, drug absorption, and elimination. Transporters can also work in concert with metabolizing enzymes (see also ▶ Chap. 31, "Pharmacodynamic Drug–Drug Interactions").

A number of transporter-based interactions have been documented in recent years. Analogous to drug interactions mediated by P450 enzymes, coadministration of a drug that is an inhibitor or an inducer of a drug transporter may affect the pharmacokinetics of a drug that is a substrate for that transporter. Transporters can affect the safety profile of a drug by affecting the concentration of a drug or its metabolites in various tissues. Transporter-based drug interactions and the potential effect of drug transporters on safety make it important to determine whether transporters affect the absorption and disposition of an investigational drug and whether the investigational drug can affect the absorption and disposition of other drugs through an effect on transporters.

The effect of a compound on drug transporter function will be investigated, e.g., in bidirectional transport experiments in tissue cultures. Results of these experiments show whether the investigated compound is a drug transporter substrate (by determining the net transport rate, efflux ratio, or Michaelis constant (Km)) or an inhibitor of the transporter (by IC_{50} or K_i values). Because of the lack of a validated in vitro system to study transporter induction, the definitive determination of induction potential of an investigational drug on transporters is based on in vivo induction studies.

An overview as regards which drugs transporters to be examined based on FDA, EMA and PMDA guideline, e.g., whether a drug is a substrate, inhibitor, or inducer or not is given in Tables 1, 2, and 3.

General Strategies for the Planning and Conduct of DDI Trials

The evaluation of a DDI potential for a compound is at first based on all in vitro data collected for a compound to whether the drug is a substrate, inhibitor, or inducer or a metabolic enzyme or drug transporter in relation to the (expected) in vivo plasma concentrations (e.g., maximum plasma concentrations). The respective cut offs, which have to be considered, are given based on a "basic model" for reversible, time-dependent inhibition and also induction. Respective overviews as regards specific recommendations (e.g., use of unbound or total drug concentrations) and cut off values (as to whether a clinical DDI trial needs to be performed) of the guidelines from different authorities are given in Tables 4, 5, and 6. Mechanistic static models and/or more comprehensive dynamic models (e.g., physiologically based PK

Table 4 Basic model: for reversible inhibition (R = 1 + [I]/Ki or IC50)

	FDA	EMA	PMDA
Metabolic enzyme: Reversible inhibition			
Systemic	[I] = total C_{max} R > 1.1	[I] = unbound C_{max} R ≥ 1.02	[I] = total C_{max} R > 1.1
Gut	$[I]_G$ = Dose/250 mL, R > (≥) 11		
Transporter: Reversible inhibition			
Systemic			
P-gp, BCRP	[I] = total C_{max} R ≥ 1.1	[I] = unbound C_{max} R ≥ 1.02	[I] = total C_{max} R ≥ 1.1
CAT1,3, OCT2 **MATE1, 2-K(PMDA)**	[I] = unbound C_{max} R ≥ 1.1	[I] = unbound C_{max} R ≥ 1.02	[I] = unbound C_{max} R ≥ 1.25
CA7P1B1,3	[I] = total C_{max} R ≥ 1.1 and [I] = unbound $C_{max,\ inlet}$[a] R ≥ 1.25	[I] = unbound $C_{max,\ inlet}$[a] R ≥ 1.04	[I] = unbound $C_{max,\ inlet}$[a] R ≥ 1.25
Gut			
P-gp, BCRP	$[I]_G$ = Dose/250 mL, R > (≥) 11		

[a]$C_{max,inlet}$ is calculated as $f_{u,b}$ x ($[I]_{max,b}$ + F_a x F_g x k_a x Dose/Q_h)

Table 5 Basic model: time-dependent inhibition (TDI)

	FDA	EMA	PMDA
Time-dependent inhibition (TDI)			
Systemic	[I] = total C_{max} R > 1.1	[I] = unbound C_{max} R ≥ 1.25	[I] = total C_{max} R > 1.1
Gut	$[I]_G$ = Dose/250 mL R > 11	$[I]_G$ = Dose/250 mL R ≥ 1.25	$[I]_G$ = Dose/250 mL R > 11

Table 6 Basic model: induction [R = 1 + Emax × [I]/(EC50 + [I])]

	FDA	EMA	PMDA
Metabolic enzyme induction			
mRNA change	> predefined threshold	>2-fold (concentration dependent increase) or 20% increase of the increase in positive control[a]	
Cut off value	[I] = total C_{max} R < 0.9	[I] = unbound C_{max} R: Not defined	[I] = total C_{max} R < 0.9
Transporter induction			
	Not mentioned		

[a]It is acceptable to use the enzyme activity as a measure, in case that the inhibition of enzyme can be clearly denied (PMDA)

(PBPK) models) may be used additionally, and specific recommendations can be found in the guidelines from FDA, EMA, and PMDA. It should be noted that currently only the FDA provides a dedicated flowchart how to explore DDI potential with PBPK models. The recommended cases to use PBPK (EMA/FDA) are to predict DDI's worst-case scenarios (additive "multiple DDI mechanisms" combined with, e.g., organ impairment), dose-dependent DDIs, the effect of a less potent inhibitor, or the impact of a DDI in subpopulations.

The initial approach to assessing clinical DDIs is the evaluation of underlying mechanisms with probe compounds. Examples of appropriate probe compounds that are considered to be specific and representative for a defined metabolic pathway or drug transporter are defined in current regulatory

Table 7 Overview as regards inhibitor, inducer, and substrate lists provided by authorities

	FDA	EMA	PMDA
Metabolic enzyme (in vitro/in vivo)			
Inhibitor	×/◎	○/◎	○/◎
Inducer	◎/◎	×/×	○/◎
Substrate	×/○	○/○	○/◎
Transporter (in vitro/in vivo)			
Inhibitor	×/○	×/×	◎/○
Inducer	×/○	×/×	◎/○
Substrate	×/○	×/×	◎/○

◎: listed with intensity for each drug (i.e., classification [weak, moderate, or strong] or Ki value), ○: listed, ×: not listed

guidelines (e.g., (1,2,3)). A brief overview, which inhibitors, inducers, and substrates are specified in the respective guidelines, can be found in Table 7.

If a clinical relevant DDI cannot be excluded through screening with a probe compound, further clinical DDI evaluations may become necessary (see Figs. 1 and 2). The demonstration of a relevant DDI in a study with a probe compound can result in the need of further studies with concomitant medications for the NME (Figs. 1 and 2) to determine:

1. Whether additional studies are needed to better quantify the effect and to examine the effects of weaker inhibitors (early studies usually examine strong inhibitors) on the investigational drugs as substrates and effects of investigational drugs (as inhibitors) on a range of substrates.
2. Whether dosage adjustments or other prescribing modifications (e.g., additional safety monitoring or contraindications) are needed based on the identified interaction(s) to avoid undesired consequences. Drug interaction information is used along with information about exposure-response relationships in the general population and specific populations, to help predict the clinical consequences of DDIs.

Should in vitro data show that a drug is an inhibitor or inducer of enzymes, it may be recommended to conduct a DDI study as early as possible in clinical development in order to exclude any possible liability of this interaction. For substrates of specific drug metabolizing or drug transporter pathways, the timing of DDI in vivo clinical studies depends on the safety range and on the frequency of co-medications that would act as inhibitor on the compound especially during Phase II or III.

Independently of underlying mechanisms for potential DDIs, it has to be evaluated whether compounds with narrow therapeutic windows (relevant co-medications) in the targeted therapeutic area need to be evaluated in addition to the before mentioned studies which use probe drugs. It should be noted that DDI studies with specific co-medications might be necessary in order to have a specific label of no clinically relevant DDI.

In general, the described principles of non-clinical and clinical assessment of DDI are also valid for oncological NCEs. However, studies in healthy volunteers might not be possible due to low tolerability of the compound. Furthermore, study design options might be limited due to a reduced clinical state of the patients. In general, all "combination trials" of standard chemotherapy together with a NME should be designed to investigate possible DDIs between the different compounds in light of this document (approaches to be discussed on a case-by-case basis).

Practical Considerations for DDI Trials

When testing an investigational drug for the possibility that its metabolism is inhibited or induced (i.e., as a substrate), selection of the interacting drugs should be based on in vitro or in vivo studies identifying the enzyme systems that

metabolize the investigational drug. The choice of the interacting drug can then be based on known, important inhibitors and inducers of the pathway under investigation. Strong inhibitors and inducers provide the most sensitive assessment and should generally be tested first

Study Design

In vivo DDI studies generally are designed to compare substrate concentrations with and without the interacting drug. Because a specific study can address a number of questions and clinical objectives, many study designs for investigating DDI can be considered. In general, crossover designs in which the same subjects receive substrate with and without the interacting drug are more efficient. A study can use a randomized crossover (e.g., Substrate (S) followed by S+ Inhibitor (I), S + I followed by S), one-sequence crossover (e.g., S followed by S + I), or a parallel (S in one group of subjects and S + I in another group) design, and there may be reasons to have another period when the I is removed to assess effect duration. The following possible dosing regimen combinations for a substrate and interacting drug can also be used: single dose/single dose, single dose/multiple dose, multiple dose/single dose, and multiple dose/multiple dose. Additional factors include consideration of the sequence of administration and the time interval between dosing of substrate and inhibitor/inducer. The selection of a study design depends on a number of factors for both the substrate and interacting drug, including:

1. Whether the substrate and/or interacting drug is used acutely or chronically
2. Safety considerations, including whether a substrate is a NTR drug (NTR drugs are defined as those drugs for which there is little separation between therapeutic and toxic doses or the associated blood or plasma concentrations) or non-NTR drug
3. Pharmacokinetic and pharmacodynamic characteristics of the substrate and interacting drugs
4. Whether there is a desire to assess induction as well as inhibition
5. Whether the inhibition is delayed
6. Whether there is a need to assess persistence of inhibition or induction after withdrawal of the interacting drug

The interacting drugs and the substrates should be dosed so that the exposures of both drugs are relevant to their clinical use, including the highest doses likely to be used in clinical practice, and plasma levels of both drugs should be obtained to show this. The following considerations may be useful:

- When attainment of steady state is important (especially for the drug being the perpetrator drug), and either the substrate or interacting drug or their metabolites have long half-lives, one or both periods of a crossover study should be long, but several other approaches can be considered, depending on pharmacokinetic characteristics of the drug and metabolites. For example, if the substrate has a long half-life, a loading dose could be used to reach steady-state concentrations earlier in a one-sequence crossover followed by an S + I period long enough to allow I to reach steady state (here too, using a loading dose could shorten that period).
- When it is important that a substrate and/or an interacting drug be studied at steady state for a long duration because the effect of an interacting drug is delayed, as is the case for inducers and time-dependent inhibition (TDI), documentation that near steady state has been attained for the pertinent substrate drug and metabolites as well as the interacting drug is critical, and both S and I should be present long enough to allow the full effect to be seen. This documentation can be accomplished by sampling over several days prior to the periods when test samples are collected. This information is important for metabolites and the parent drug, particularly when the half-life of the metabolite is longer than the parent. It is also important when the interacting drug and metabolites both are metabolic inhibitors (or

inducers). Finally, it is critical to evaluate the time it takes for the enzyme activities to return to normal when induction or TDI is involved so that a third crossover period in which the interacting drug (I) is removed will generally be recommended.

- Studies can usually be open label (unblinded), unless pharmacodynamic endpoints (e.g., adverse events that are subject to bias) are critical to the assessment of the interaction.
- For a rapidly reversible inhibitor, administration of the interacting drug either just before or simultaneously with the substrate on the test day might increase sensitivity by ensuring maximum exposure to the two drugs together. For a mechanism-based inhibitor (a drug that requires metabolism before it can inactivate the enzyme; an example is erythromycin), administration of the inhibitor prior to the administration of the substrate drug can maximize the effect. If the absorption of an interacting drug may be affected by other factors (e.g., the gastric pH), it may be appropriate to control the variables or confirm the absorption through plasma level measurements of the interacting drug.
- Timing of administration may be critical in situations of concurrent inhibition and induction. For example, if the investigational drug is a substrate for both enzymes and OATP, and rifampin is used as an enzyme inducer, the simultaneous administration of the drug with rifampin (an OATP inhibitor) may underestimate enzyme induction, so delayed administration of the substrate is recommended. The optimal delayed time should be determined. In addition, it is critical to evaluate the duration of the interaction effect after the interacting drug has been removed.
- When the effects of two drugs on one another are of interest, the potential for interactions can be evaluated in a single study or two separate studies. Some design options are randomized three-period crossover, parallel group, and one-sequence crossover.
- To avoid variable study results because of uncontrolled use of dietary/nutritional supplements, tobacco, alcohol, juices, or other foods that may affect various metabolizing enzymes and transporters during in vivo studies, it is important to exclude, when appropriate, subjects who used prescription or over-the-counter medications, dietary/nutritional supplements, tobacco, or alcohol within 1 week prior to enrollment. In addition, investigators should explain to subjects that for at least 1 week prior to the start of the study until its conclusion.
- Because interactions might differ in subgroups of different pharmacogenetic genotypes, genotyping for the enzymes and transporters involved in the interaction should be carried out when appropriate.
- Detailed information on the dose given and time of administration should be documented for the coadministered drugs.

Study Population

In most situations, clinical DDI studies can be performed using healthy volunteers, and findings in healthy volunteers will predict findings in the patient population for which the drug is intended. Safety considerations, however, may preclude the use of healthy subjects in studies of certain drugs. In addition, there are circumstances in which subjects drawn from the intended patient population offer advantages, including the opportunity to study pharmacodynamic endpoints not present in or relevant to healthy subjects. The extent of drug interactions (inhibition or induction) may be different depending on the subjects' genotype for the specific enzyme or transporter being evaluated. For example, subjects lacking the major polymorphic clearance pathway will show reduced total metabolism or transport. However, alternative pathways can become quantitatively more important in these subjects. In such cases, the alternative pathways should be understood and studied appropriately. Thus, phenotype or genotype determinations to identify genetically determined metabolic or transporter polymorphisms are important when evaluating effects on enzymes or transporters with polymorphisms, such as CYP2D6,

CYP2C19, CYP2C9, UGT1A1, and OATP1B1 (SLCO1B1). In addition, it is valuable to specify the need for stratifying the population based on genotype while conducting the DDI studies. Another alternative is to consider powering the study for the genotype status that is likely to have the highest potential for interaction.

Choice of Substrate and Interacting Drugs

CYP-Mediated Interactions

The Investigational Drug as a Substrate of CYP Enzymes – Effect of Other Drugs on Investigational Drugs

When testing an investigational drug for the possibility that its metabolism is inhibited or induced (i.e., being the victim drug of a DDI as a substrate), selection of the interacting drug can then be based on known, important inhibitors and inducers of the pathway under investigation. Strong inhibitors and inducers provide the most sensitive assessment and should generally be tested first. Consider, for example, an investigational drug metabolized by CYP3A with the contribution of this enzyme to the overall elimination of this drug that is either substantial ($\geq$ 25% of the clearance pathway) or unknown. In this case, the inhibitor and inducer can be itraconazole and rifampin, a strong inhibitor and a strong inducer, respectively. Respective strong inhibitors or inducers should be looked after in the respective sections of the guidelines from FDA, EMA, and PMDA or in the most current literature. If the study results are negative, then absence of a clinically important DDI for the metabolic pathway is demonstrated. If the clinical study of the strong inhibitor or inducer is positive, effects through in vivo studies of other less potent specific inhibitors or inducers may be needed to be evaluated. If the investigational drug is metabolized by CYP3A and its plasma AUC (Area under the plasma concentration time curve) is increased fivefold or higher by strong CYP3A inhibitors, it is considered a *sensitive substrate* of CYP3A. The labeling would indicate that the drug is a "sensitive CYP3A substrate" and that its use with strong or moderate inhibitors may call for caution, depending on the drug's exposure-response relationship.

For further information as regards the labeling of respective DDI effects, please look at the respective section of the DDI guidelines from EMA, FDA, and PMDA.

If a drug is metabolized by a polymorphic enzyme (such as CYP2D6, CYP2C9, 1327 CYP2C19, or UGT1A1), the comparison of pharmacokinetic parameters of this drug in poor metabolizers and extensive metabolizers may substitute for an interaction study for that particular pathway, as the PK in the poor metabolizers will indicate the effect of a strong inhibitor. When the study suggests the presence of a significant interaction with strong inhibitors or in poor metabolizers, further clinical DDI evaluation, e.g., with weaker inhibitors or intermediate metabolizers, may be recommended (additionally also mechanistic modeling approaches may be used supporting respective investigations).

The Investigational Drug as an Inhibitor or an Inducer of CYP Enzymes: Effect of Investigational Drugs on Other Drugs

When studying an investigational drug as the interacting drug (being the perpetrator drug), the choice of substrates (approved drugs) for initial in vivo studies depends on the P450 enzymes affected by the interacting drug. When testing inhibition, the substrate selected should generally be one whose pharmacokinetics are markedly altered by the coadministration of known specific inhibitors of the enzyme systems (sensitive substrates) to see the largest impact of the interacting investigational drug. Examples of such substrates include (refer also to the respective section in the EMA, FDA, and PMDA DDI guidelines and most recent literature):

1. Midazolam for CYP3A
2. Theophylline for CYP1A2
3. Bupropion for CYP2B6
4. Repaglinide for CYP2C8

5. Warfarin for CYP2C9 (with the evaluation of S-warfarin)
6. Omeprazole for CYP2C19
7. Desipramine for CYP2D6

If the initial study determines that an investigational drug either inhibits or induces metabolism of sensitive substrates, further studies using other substrates, representing a range of therapeutic classes, based on the likelihood of coadministration, may be useful. If the initial study with the most sensitive substrates is negative, it can be presumed that less sensitive substrates also will be unaffected. It should be noted that several of the substrates recommended for drug interaction studies are not specific because they are substrates for more than one CYP enzyme or may be substrates for drug transporters. While a given substrate may not be metabolized by a single enzyme (e.g., dextromethorphan elimination is carried out primarily by CYP2D6 but other enzymes also contribute in a minor way), its use in an interaction study is appropriate if the inhibitor (the investigational drug) to be evaluated is selective for the CYP enzyme of interest. If an investigational drug is a CYP inhibitor, it may be classified as a strong, moderate, or weak inhibitor based on its effect on a sensitive CYP substrate. For example, CYP3A inhibitors can be classified based on the magnitude of the change in plasma AUC of oral midazolam or other CYP3A substrates that are similar in characteristics (e.g., fm (%clearance contributed by CYP3A), half-life, not subject to transporter effect) as midazolam, when the substrate is given concomitantly with the inhibitor. If the investigational drug increases the AUC of oral midazolam or other CYP3A substrate by fivefold or higher ($\geq$ fivefold), it can be considered a *strong* CYP3A inhibitor. If the investigational drug, when given at its highest dose, increases the AUC of oral midazolam or other sensitive CYP3A substrates by between two- and fivefold ($\geq$ two- and <fivefold), it can be considered a *moderate* CYP3A inhibitor. If the investigational drug, when given at the highest dose and shortest dosing interval, increases the AUC of oral midazolam or other sensitive CYP3A substrates by between 1.25- and 2-fold ($\geq$1.25- and <2-fold), it can be considered a *weak* CYP3A inhibitor.

When an in vitro evaluation does not rule out the possibility that an investigational drug is an inducer of CYP3A, an in vivo evaluation can be conducted using the most sensitive substrate (e.g., oral midazolam). When midazolam is coadministered orally following the administration of multiple doses of the investigational drug, and there is no interaction, it can be concluded that the investigational drug is not an inducer of CYP3A (in addition to the conclusion that it is not an inhibitor of CYP3A). A caveat to this interpretation is that if the investigational drug is both an inducer and inhibitor of CYP3A, such as ritonavir, the net effect at any time it is introduced may vary. In this case, the net effect of the drug on CYP3A function may be time dependent. In vivo induction evaluations have often been conducted using oral contraceptives as the substrate. However, oral contraceptives are not the most sensitive substrates for CYP3A, so a negative result does not exclude the possibility that the investigational drug is an inducer of CYP3A. Some compounds listed as sensitive substrates for the other enzymes can also be used as substrates with the investigational drug as an inducer. For example, omeprazole and repaglinide are CYP2C19 and CYP2C8 substrates, respectively, but they are also metabolized by CYP3A. If omeprazole is used as a substrate to study CYP2C19 induction, measurement of its metabolites (CYP2C191397-mediated hydroxy-omeprazole and CYP3A4-mediated omeprazole sulfone) will be recommended for the interpretation of the study results.

Transporter-Mediated Interactions

Similar to CYP enzymes, transporters may be inhibited or induced. Inhibition of transporters by interacting drugs can lead to altered exposure of other drugs that are substrates of transporters. Therefore, the potential for an investigational drug as a substrate, inhibitor, or inducer for transporters should be evaluated during drug development.

In the most recent guidances from EMA, FDA, and PMDA, BCRP, OATP, OATs, and OCTs are considered important transporters (see also Tables 1, 2, and 3) in addition to P-gp and should be

routinely evaluated. Because the field of transporter pharmacology is rapidly evolving, other transporters (e.g., multidrug resistance-associated proteins (MRPs), multidrug and toxin extrusion (MATE) transporters, and bile salt export pump (BSEP) transporters) should be considered when appropriate.

The Investigational Drug as a Substrate of Transporters: The Effect of Other Drugs on an Investigational Drug

When testing an investigational drug for the possibility that its transport is inhibited or induced (i.e., as a substrate), selection of the interacting drugs should be based on in vitro or in vivo studies identifying the transporters that are involved in the absorption and disposition of the investigational drug (e.g., absorption and efflux in the gastrointestinal tract, uptake and secretion in the liver, and the secretion and reabsorption in the kidney). The choice of the interacting drug should be based on known, important inhibitors of the pathway under investigation. Strong inhibitors provide the most sensitive assessment and should generally be tested first. As there is overlapping selectivity in substrate and inhibitor among transporters, negative results from a study using a broad inhibitor may rule out the possibility for drug interaction mediated by multiple pathways. For example, it may be appropriate to use an inhibitor of many transporters (e.g., cyclosporine, which inhibits P-gp, OATP, and BCRP) to study its effect on a drug that may be a substrate for these transporters. A negative result rules out the involvement of these transporters in the drug's disposition. However, if the result is positive, it will be difficult to determine the relative contribution of each transporter to the disposition of the substrate drug. In contrast, if the goal of the study is to determine the role of a specific pathway in the PK of a substrate drug, then a selective and potent inhibitor for that transporter should be used. As an alternative, comparative PK of an investigational drug in subjects with different genotypes of specific transporters can be evaluated to determine the importance of a specific transporter in the clearance pathway for the drug. On the other hand, polymorphism data on P-gp is controversial and may not be used to determine the role of P-gp in the disposition of investigational drugs that are substrates of P-gp.

The Investigational Drug as an Inhibitor or an Inducer of Transporters: Effect of the Investigational Drugs on Other Drugs

When studying an investigational drug as the interacting drug, the choice of substrates for initial in vivo studies depends on the transport pathway that may be affected by the interacting drug. In general, when testing inhibition, the substrate selected should be one whose pharmacokinetics are markedly altered by coadministration of known specific inhibitors of the transporter pathway to see the largest impact of the interacting investigational drug. The choice of substrates can also be determined by the therapeutic area of the investigational drug and the probable coadministered drugs that are known substrates for transporters (respective lists of selected substrates for transporters can be found in the respective guidelines of the EMA, FDA and PMDA guideline and most recent literature). However, because many drugs are substrates of multiple transporters or enzymes, specific substrates for each transporter are not available. The observed clinical interactions may be a result of inhibition of multiple pathways if the investigational drug is also an inhibitor for the same multiple pathways.

Because of the lack of a validated in vitro system to study transporter induction, the definitive determination of induction potential of an investigator on transporters is based on in vivo induction studies. For example, because of similarities in the mechanisms of CYP3A and P-gp induction, information from the testing of CYP3A inducibility can inform decisions about P-gp. If an investigational drug is found not to induce CYP3A in vitro, no further tests of CYP3A and P-gp induction in vivo are necessary. If a study of the investigational drug's effect on CYP3A activity in vivo is indicated from a positive in vitro screen, but the drug is shown not to induce CYP3A in vivo, then no further test of P-gp induction in vivo is necessary. However, if the in vivo CYP3A induction test is positive, then an additional study of the investigational drug's

effect on a P-gp probe substrate is recommended. If the drug is also an inhibitor for P-gp, then the induction study can be conducted with the inhibitor study using a multiple-dose design.

EMA, FDA, and PMDA DDI guidelines contain valuable information regarding the classification of in vivo inhibitors or inducers for CYP enzymes, examples of sensitive in vivo CYP substrates and CYP substrates with narrow therapeutic ranges, examples of in vivo inhibitors and inducers of selected transporters, examples of in vivo substrates of selected transporters and examples of in vivo CYP3A and P-gp inhibitors and their relative potency.

Complex Drug Interactions

The above sections separately discussed DDIs related to effects on enzymes and transporters, but drug interactions for a specific drug may occur based on a combination of mechanism and have to be taken in consideration, when a clinical DDI trial needs to be designed.

Such "complex drug interaction" scenarios include, but are not limited to:

- Concurrent inhibition and induction of one enzyme or concurrent inhibition of enzyme and transporter by a drug
- Increased inhibition of drug elimination by the use of more than one inhibitor of the same enzyme that metabolizes the drug
- Increased inhibition of drug elimination by use of inhibitors of more than one enzyme that metabolizes the drug
- Inhibition by a drug and its metabolite or metabolites, both of which inhibit the enzyme that metabolizes the substrate drug
- Inhibition of an enzyme other than the genetic polymorphic enzyme in poor metabolizers taking substrate that is metabolized by both enzymes

Multiple CYP Inhibitors

There may be situations when an evaluation of the effect of multiple CYP inhibitors on the drug can be informative. For example, it may be appropriate to conduct a DDI study with more than one inhibitor simultaneously if all of the following conditions are met:

1. The drug exhibits blood concentration-dependent important safety concerns.
2. Multiple CYP enzymes are responsible for the metabolic clearance of the drug.
3. The predicted residual or noninhibitable drug clearance is low.

Under these conditions, the effect of multiple CYP-selective inhibitors on the investigational drug's blood AUC may be much greater than when the inhibitors are given individually with the drug, and more than the product of changes in AUC observed with each individual inhibitor. The magnitude of the combined effect will depend on the residual fractional clearance (the smaller the fraction, the greater the concern) and the relative fractional clearances of the inhibited pathways. Modeling and simulation approaches can help to project the magnitude of the effect based on single pair drug interaction studies. If results from a study with a single inhibitor have already triggered a major safety concern (i.e., a contraindication), multiple inhibitor studies are unlikely to add value.

Enzyme/Transporter Interplay

There is an overlap in enzyme and transporter specificity. For example, there is considerable overlap between CYP3A and P-gp inhibitors and inducers. Itraconazole inhibits CYP3A and P-gp and rifampin induces CYP3A and P-gp. However, dual inhibitors for CYP3A and P-gp do not necessarily have similar inhibition potency on CYP3A and P-gp. To assess the worst-case scenario for a dual CYP3A and P-gp substrate, inhibition should be studied using an inhibitor that shows strong inhibition for both P-gp and CYP3A, e.g., such as itraconazole. However, under this condition, if the result is positive, specific attribution of an AUC change to P-gp or CYP3A4 may not be possible. If the goal is to determine the specific contribution of CYP3A or P-gp on the AUC change, then a strong inhibitor for CYP3A only or a potent inhibitor for P-gp only should be selected to discern the effect of CYP3A versus P-gp.

In addition to the possibility that a drug is an inhibitor or inducer of multiple enzymes/transporters, a drug can be an inhibitor of one enzyme/transporter and inducer of another enzyme/transporter. For example, rifampin, an established inducer of multiple CYP enzymes and transporters, was recently found to be an inhibitor of the uptake transporter OATP1B1 and may inhibit the uptake of an investigational drug that is a substrate of OATP1B1. Accordingly, if a drug is a CYP enzyme substrate and an OATP1B1 substrate, an induction study with rifampin should be designed and interpreted carefully. The net steady-state effect may vary depending on the relative size of the individual effect on transporter and enzyme activities. Timing of administration may become critical in situations when both enzymes and transporters can be affected. These overlapping selectivities contribute to complex drug interactions and make the prediction of in vivo outcome based on in vitro evaluation challenging or impossible (Zhang et al. 2009a, b). The implications of simultaneous inhibition of a dominant CYP enzyme(s) and an uptake or efflux transporter that controls the availability of the drug to CYP enzymes can be just as profound as that of multiple CYP inhibition. For example, the large effect of coadministration of itraconazole and gemfibrozil on the systemic exposure (AUC) of repaglinide may be attributed to collective inhibitory effects on both the enzyme (CYP2C8) and transporters (OATP1B1) by itraconazole and gemfibrozil and their respective metabolites.

Effect of Organ Impairment

Another type of complex drug interaction is the coadministration of substrate and enzyme/transporter inhibitor in subjects with organ impairment. For example, if a substrate drug is eliminated through both hepatic metabolism and renal secretion/filtration, the use of an enzyme inhibitor in subjects with renal impairment may cause a more than projected increase in exposure of substrate drug based on individual effect alone. Unfortunately, current knowledge does not permit the presentation of specific guidance for studying some of these complex drug interaction scenarios because dedicated in vivo studies in humans may not be feasible or may raise ethical and practical considerations. Modeling and simulation approaches integrating prior in vitro and in vivo ADME and drug interaction data may be useful for evaluating complex drug interactions. For example, results from dedicated single pair drug interaction studies and separate pharmacokinetic evaluation in subjects with organ impairment may provide useful information to strengthen the model for the evaluation of complex drug interactions.

Pediatrics and Geriatrics

Age-related changes in physiological processes governing drug disposition and drug effect have been investigated. In some cases, disproportional alterations in binding proteins, drug metabolizing enzymes and/or transporters, and renal filtration/secretion caused by developmental changes have been known to result in different drug disposition characteristics in pediatric and geriatric populations. However, dedicated drug interaction studies in these populations may not be feasible. Simulations using system biology approaches such as PBPK models may be helpful to predict drug interaction potential when the model can be constructed based on sufficient in vitro and clinical pharmacology and drug interaction data and incorporates development changes.

Route of Administration

The route of administration chosen for a metabolic DDI study is important. For an investigational agent, the route of administration generally should be the one planned for clinical use. When multiple routes are being developed, the need for metabolic DDI studies by each route depends on the expected mechanisms of interaction and the similarity of corresponding concentration-time profiles for parent drug and metabolites. Sometimes certain routes of administration can reduce the utility of information from a study. For example, intravenous administration of a substrate drug may not reveal an interaction for substrate drugs where intestinal CYP3A activity markedly alters bioavailability

Dose Selection

The doses of the substrate and interacting drug used in studies should maximize the possibility of demonstrating an interaction. For this reason, the maximum planned or approved dose and shortest dosing interval of the interacting drug (as inhibitors or inducers) should be used. For example, when using itraconazole as an inhibitor of CYP3A, the decision whether to dose at 400 mg QD or 200 mg BID for multiple days can be determined based on the pharmacokinetic characteristics (e.g., the half-life) of the substrate drug (Zhao et al. 2009). When using rifampin as an inducer, dosing at 600 mg QD for multiple days would be preferable to lower doses. When there are safety concerns, doses lower than those used clinically may be recommended for substrates. In such instances, any limitations of the sensitivity of the study to detect the drug–drug interaction due to the use of lower doses should be carefully considered.

Endpoints

Changes in pharmacokinetic parameters generally are used to assess the clinical importance of DDIs. Interpretation of findings (i.e., deciding whether a given effect is clinically important) depends on a good understanding of dose/concentration and concentration/response relationships for both desirable and undesirable drug effects in the general population or in specific populations. In certain instances, reliance on pharmacodynamic endpoints in addition to pharmacokinetic measures and/or parameters may be useful. Examples include INR measurement (e.g., when studying warfarin interactions) or QT interval measurements.

Pharmacokinetic Endpoints

Substrate PK exposure measures such as AUC, C_{max}, time to C_{max} (T_{max}), and others as appropriate should be obtained in every study. Calculation of pharmacokinetic parameters such as clearance, volumes of distribution, and half-lives may help in the interpretation of the results of the trial. In some cases, obtaining these measures for the inhibitor or inducer may be of interest as well, notably where the study is intended to assess possible changes in the disposition of both study drugs. Additional measures may help in steady-state studies (e.g., trough concentration) to demonstrate that dosing strategies were adequate to achieve near steady state before and during the interaction. In certain instances, an understanding of the relationship between dose, plasma concentrations, and response may lead to a special interest in certain pharmacokinetic measures and/or parameters. For example, if a clinical outcome is most closely related to peak concentration (e.g., tachycardia with sympathomimetics), C_{max} or an early exposure measure may be most appropriate for evaluation. Conversely, if the clinical outcome is related more to extent of absorption, AUC would be preferred. The frequency of sampling should be adequate to allow accurate determination of the relevant measures and/or parameters for the parent molecule and metabolites. For the substrate, whether the investigational drug or the approved drug, determination of the pharmacokinetics of relevant metabolites is important. Also, measurement of these metabolites may be useful to differentiate the effect of inhibitor/inducer on pathways mediated by different CYP enzymes.

Statistical Considerations, Clinical Relevance, and Sample Size

The goal of a DDI study is to determine whether there is any increase or decrease in exposure to the substrate in the presence of the interacting drug. If there is, its implications should be assessed by an understanding of PK/PD relations both for C_{max} and AUC. Results of DDI studies should be reported as 90% confidence intervals about the geometric mean ratio of the observed pharmacokinetic measures with (S + I) and without the interacting drug (S alone). Confidence intervals provide an estimate of the distribution of the observed systemic exposure measure ratio of (S + I) versus (S alone) and convey a probability of the magnitude of the interaction. In contrast, tests of significance are not appropriate because

small, consistent systemic exposure differences can be statistically significant ($p < 0.05$), but not clinically relevant.

When a DDI of potential importance is clearly present, specific recommendations should be provided regarding the clinical significance of the interaction based on what is known about the dose-response and/or PK/PD relationship for the substrate drug used in the study. This information can form the basis for reporting study results and for making recommendations in the labeling. It should be recognized that dose-response and/or PK/PD information can sometimes be incomplete or unavailable, especially for an older approved drug used as a substrate. If the sponsor wishes to include a statement in the labeling that no known DDI of clinical significance exists, the sponsor should recommend specific *no effect* boundaries, or clinical equivalence intervals, for a DDI and should provide the scientific justification for the recommendations. No effect boundaries represent the interval within which a change in a systemic exposure measure is considered not clinically meaningful. These conclusions can be based on dose-response data or on PK/PD modeling.

There are two approaches to defining no effect boundaries:

Approach 1: No effect boundaries can be based on the population (group) average dose-related and/or individual concentration-response relationships derived from PK/PD models, and other available information for the substrate drug to define a degree of difference caused by the interaction that is of no clinical consequence. If the 90% confidence interval for the systemic exposure measurement change in the DDI study falls completely within these no effect boundaries, it can be concluded that no clinically significant drug–drug interaction is present.

Approach 2: In the absence of no effect boundaries defined in Approach 1, a default no effect boundary of 80–125% can be used for both the investigational drug and the approved drugs used in the study. When the 90% confidence intervals for systemic exposure ratios fall entirely within the equivalence range of 80–125%, standard practice is to conclude that no clinically significant differences are present. This is, however, a very conservative standard and a substantial number of subjects (sample size) would need to be studied to meet it.

The selection of the number of subjects for a given DDI study will depend on how small an effect is clinically important to detect or rule out the inter- and intra-subject variability in pharmacokinetic measurements, and possibly other factors or sources of variability not well recognized.

Cocktail Approaches

Simultaneous administration of a mixture of substrates of multiple CYP enzymes and transporters in one study (i.e., a "cocktail approach") in human volunteers is another way to evaluate a drug's inhibition or induction potential, provided that the study is designed properly and the following factors are present:

1. The substrates are specific for individual CYP enzymes or transporters.
2. There are no interactions among these substrates.
3. The study is conducted in a sufficient number of subjects.

Negative results from a well-conducted cocktail study can eliminate the need for further evaluation of particular CYP enzymes. However, positive results can indicate that further in vivo evaluation should be conducted to provide quantitative exposure changes (such as AUC, C_{max}), if the initial evaluation only assessed the changes in the urinary parent to metabolite ratios. The data generated from a cocktail study can supplement data from other in vitro and in vivo studies in assessing a drug's potential to inhibit or induce CYP enzymes and transporters

Negative results from a well-conducted cocktail study may eliminate the need for further evaluation of particular CYP enzymes and transporters. However, positive results may indicate that further in vivo evaluation should be conducted

Pharmacogenomic Considerations

When a DDI study uses a probe drug (e.g., omeprazole for CYP2C19) to evaluate the impact of the investigational drug on a polymorphic enzyme, individuals who have no functional enzyme activity would not be appropriate study subjects. Drug interaction studies that evaluate enzymes or transporters with known polymorphisms should include collection of genotype or phenotype information to allow appropriate interpretation of the study results. In some instances, an evaluation of the extent of drug interactions in subjects with various genotypes may be helpful. Moreover, DDIs can differ among individuals based on genetic variation of a polymorphic enzyme. For example, a strong CYP2D6 inhibitor (e.g., fluoxetine) will increase the plasma levels of a CYP2D6 substrate (e.g., atomoxetine) in subjects who are extensive metabolizers (EM) of CYP2D6, but will have minimal effect in subjects who are poor metabolizers (PM) of CYP2D6, because these individuals have no active enzyme to inhibit. It is noted that CYP2D6 PMs will already have greatly increased levels of atomoxetine if given usual doses. There are also situations where inhibition may have a greater effect in PMs than EMs. If a drug is metabolized by a minor pathway (nonpolymorphic enzyme) and a major pathway (polymorphic enzyme), inhibition of the minor pathway will usually have minimal effect on plasma concentrations in EMs. However, the minor pathway plays a greater role in clearance of the drug in PMs of the major pathway. Thus, inhibition of the minor pathway in PMs of the major pathway can have a significant effect on drug clearance and resulting drug concentrations. Therefore, studying the effect of interactions may be recommended in subjects with varied genotypes or phenotypes

DDI as Part of Pop PK in Phase II and Phase III Trials

Population pharmacokinetic (PopPK) analyses of data obtained from large-scale clinical studies that include sparse or intensive blood sampling can help characterize the clinical impact of known or newly identified interactions and determine recommendations for dosage modifications for the investigational drug as a substrate. The results of such analyses can be informative and sometimes conclusive when the clinical studies are adequately designed to detect significant changes in drug exposure due to DDIs. PopPK evaluations may also detect unsuspected DDIs, a particularly important possibility given the complexity of the potential interactions, not all of which are likely to have been anticipated and studied. PopPK evaluations can also provide further evidence of the absence of a DDI, when supported by prior evidence and mechanistic data. It is unlikely, however, that population analysis will persuasively show the absence of an interaction that is suggested by information from in vivo studies specifically designed to assess a DDI. To be optimally informative, PopPK studies should have carefully designed study procedures and sample collection protocols. Simulations (e.g., by population-based PBPK models) can provide valuable insight into optimizing the study design. Detailed information on the dose given and time of administration should be documented for the coadministered drugs. When relevant for the specific drug, the time of food consumption should be documented. Population analyses should focus on excluding a specific clinically meaningful PK change. Because exposure of coadministered drugs is not monitored in most PopPK studies, the PopPK approach may not be useful to assess the effect of the investigational drugs on other drugs.

DDI Considerations for NBEs

A comparison of the most recent DDI guidelines from EMA, FDA, and PMDA, only the guideline FDA provides a flowchart how to evaluate DDI for biologics (see respective part of the FDA DDI guideline). The respective EMA guideline does not mention about DDI of biologics. Whereas the FDA and PMDA recommend to conduct clinical DDI study, in case the biologics are cytokine or cytokine modulator. Additionally, FDA and

PMDA recommend to examine PK/PD, in case the biologics is used as combination therapy with other agents. Below, some more specific recommendations for DDI evaluations with NBEs are given based on the FDA DDI guideline:

Therapeutic proteins (TPs) typically do not undergo metabolism or transport as their clearance pathway, therefore the potential is limited for small molecule drugs to affect TPs through metabolism or transport pathways. However, a drug may affect the clearance of TPs through the drug's effect on immunogenicity (e.g., methotrexate reduces the clearance of infliximab, possibly due to methotrexate's effect on the antibodies formed against infliximab). In addition, TPs that are cytokines or cytokine modulators may modify the metabolism of drugs that are substrates for CYP P450 enzymes through their effects on the regulation pathways of CYP P450 enzymes. For example, cytokines such as IL-6 can produce concentration-dependent inhibition on various CYP isoforms at the transcription level or by alteration of CYP enzyme stability in patients with infection or inflammation and increase the plasma concentrations of specific CYP substrate drugs. In contrast, cytokine modulators such as tocilizumab (anti-IL-6 receptor antibody) may reverse the apparent "inhibition" effect of the cytokines on CYP substrates, resulting in a "normalization" of CYP activities.

Drug-TP interactions have been observed and information about these interactions is included in labeling and in the following some general considerations are given:

- If an investigational TP is a cytokine or cytokine modulator, studies should be conducted to determine the TP's effects on CYP enzymes or transporters (Huang et al. 2010; Le Vee et al. 2009). In vitro or animal studies have limited value in the qualitative and quantitative projection of clinical interactions because translation of in vitro to in vivo and animal to human results to date has been inconsistent, necessitating in vivo drug interaction studies. The in vivo evaluations of TPs in targeted patient populations can be conducted with individual substrates for specific CYP enzymes and transporters, or studies can be conducted using a "cocktail approach."
- For TPs that will be used in combination with other drug products (small molecule or TP) as a combination therapy, studies should evaluate the effect of each product on the other. The studies should assess effects on pharmacokinetics (PK) and, when appropriate, pharmacodynamics (PD) of either drug. This evaluation is particularly important when the drug used in combination has a NTR (e.g., chemotherapeutic agents).
- When there are known mechanisms or prior experience with certain PK or PD interactions, appropriate in vitro or in vivo assessments for possible interactions should be conducted. Some interactions between drugs and TPs are based on mechanisms other than CYP or transporter modulation. For example, methotrexate's immunosuppressive effect may alter the clearance of concomitantly administered TPs through the reduction of antibodies formed against TP.

References and Further Reading

EMA-CHMP (2012) Guideline on the investigation of drug interactions: final (CPMP/EWP/560/95/Rev. 1 corr. 2). http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf. Accessed 13 Mar 2017

Huang S-M, Zhao H, Lee J, Reynolds KS, Zhang L, Temple R, Lesko LJ (2010) Therapeutic protein drug interactions and impacts on drug development, Clin Pharmacol Ther 87:497–503

Le Vee M, Lecureur V, Stieger B, Fardel O (2009) Regulation of drug transporter expression in human hepatocytes exposed to the proinflammatory cytokines tumor necrosis factor-alpha or interleukin-6, Drug Metab Dispos 37:685–693

PMDA. Pharmaceuticals & Medical Device Agency-Japan (2014) Drug interaction guideline for drug development and labeling recommendations (draft for public comment) 2014. http://www.solvobiotech.com/documents/Japanese_DDI_guideline_(draft)_2014Jan.pdf. Accessed 13 Mar 2017

US-FDA (2012) Guidance for industry: drug interaction studies – study design, data analysis, implications for dosing, and labeling recommendations (draft guidance). http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm292362.pdf. Accessed 13 Mar 2017

Zhang L, Zhang Y, Zhao P, Huang S-M (2009a) Predicting drug-drug interactions: An FDA perspective, The AAPS Journal 11(2):300–306

Zhang L, Zhang Y, Huang S-M (2009b) Scientific and regulatory perspectives on metabolizing enzyme-transporter interplay and its role in drug interactions - challenges in predicting drug interaction, Molecular Pharmaceutics 6(6): 1766–1774

Zhao P, Ragueneau-Majlessi I, Zhang L, Strong J, Reynolds K, Levy R, Thummel K, Huang S-M (2009) Quantitative evaluation of pharmacokinetic inhibition of CYP3A substrates by ketoconazole – a simulation study, J Clin Pharmacol 49(3):351–359, 2143

In Vitro/In Vivo Correlation for Drug-Drug Interactions

46

Jan Wahlstrom and Larry Wienkers

Contents

Abstract

Characterizing the potential for drug-drug interactions is critical to underwriting patient safety as new chemical entities proceed through the drug discovery and development pipeline. In vitro experiments to characterize the type and extent of interaction have been developed to inform chemical modifications early in discovery and to estimate the magnitude of potential interactions as drugs progress into the clinic. Regulatory guidance provides flow schemes based on a comprehensive understanding of drug disposition to enable decision making as to whether particular clinical interaction studies need to be run and, if so, how they may be designed. Integration of information from in vitro, in vivo, and clinical sources provides the basis for drug labeling and the safe administration of drugs post-launch.

J. Wahlstrom
Amgen, Inc, Pharmacokinetics and Drug Metabolism, Thousand Oaks, CA, USA
e-mail: janw@amgen.com

L. Wienkers (✉)
Amgen, Inc, Pharmacokinetics and Drug Metabolism, South San Francisco, CA, USA
e-mail: wienkers@amgen.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_14

Introduction

Drug-drug interactions (DDIs) occur when the dosing of a drug influences the pharmacokinetics (PK) or pharmacodynamics (PD) of a second drug. It has been estimated that 1–5% of hospital admissions may be due to DDIs (Day et al. 2017). Drugs such as mibefradil, terfenadine, and nefazodone (QTc prolongation) or bromfenac, alosetron, and cerivastatin (toxicity) have been removed from the market due to a high potential for DDIs (Wienkers and Heath 2005). Drug labels may contain a black box warning if concern over DDI potential is great enough. Factors influencing the likelihood and severity of DDIs may include age (the very young and aged are more susceptible), disease state, genetics, and polypharmacy, where the risk of DDIs increases dramatically when a patient is taking four or more medications simultaneously (Jacubeit et al. 1990). Due to the importance of DDIs in the drug discovery and development process, the US Food and Drug Administration (in vitro and clinical FDA Guidance for Industry 2017a,b) and the European Medicines Agency (EMA Guideline on the investigation of drug interactions 2012) have released and continually update regulatory guidance on the design and performance of in vitro and clinical DDI studies, as well as decision trees to guide decision making as to whether particular DDI studies may be necessary as a new molecular entity (NME) proceeds through the drug discovery and development pipeline.

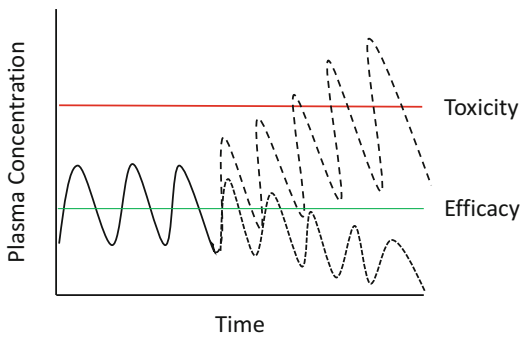


Fig. 1 Effects of inhibition or induction on a victim drug

As shown in Fig. 1, co-administration of a second drug to a drug dosed to steady state may increase or decrease the drug concentration of the first drug. Both the direction and magnitude of the DDI effect are important (Rowland and Martin 1973). Reduced drug concentrations, due to a process called induction, may fall below the efficacious threshold leading to loss of pharmacological effect. Increased drug concentrations, due to enzyme inhibition or inactivation, may exceed a toxicological threshold leading to adverse events. The range of drug concentrations and doses between the efficacious and toxicological threshold is called the therapeutic index. Drugs with a narrow therapeutic index are of high concern for DDIs, as small increases or decreases (less than two-fold) in drug concentration or doses may result in loss of efficacy or adverse drug reactions, respectively (Levy 1998). Drugs with a wide therapeutic index may exhibit large magnitude of DDIs without significant pharmacological or toxicological effect. The affected drug is called the victim, object, or substrate, whereas the affecting drug is called the perpetrator, precipitant, or inhibitor.

DDIs may be PK or PD based. PK is the study of what the body does to a drug. Typically, plasma or serum concentrations of a drug are measured as a surrogate for the concentration of drug at the site of pharmacological activity. This information is then visualized as a plasma concentration-time profile and quantified with PK parameters such as the area under the plasma concentration-time curve (AUC) or the maximum observed concentration of drug (C_{max}) as determinants of overall drug exposure. PK-based drug interactions typically involve the inhibition or inactivation of enzyme activity, or the enhancement of enzyme expression, leading increases or decreases in drug exposure, respectively. PK-based DDIs also include phenomena such as loss of exposure due to increased gastrointestinal pH observed upon co-administration with an acid-reducing agent (Chung et al. 2015). PD is the effect of a drug on the body. PD-based DDIs occur when drugs influence each other's pharmacological effects directly, such as the co-administration of two sedatives to potentiate activity.

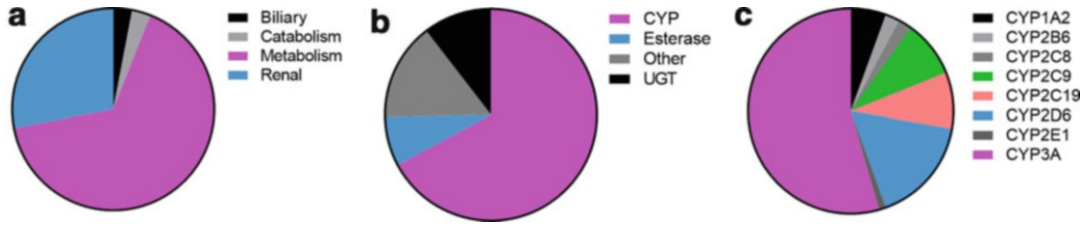


Fig. 2 Clearance mechanisms for the top 200 drugs of 2016

Small-molecule drugs are typically cleared through a combination of metabolism, renal excretion, and biliary excretion (Lin and Lu 1997). Those drugs undergoing metabolism may be cleared through phase I oxidation by enzymes such as the cytochrome P450s (CYPs), flavin monooxygenases (FMO), or aldehyde oxidase (AO). They may undergo phase II conjugation by enzymes such as the uridine 5′-diphosphoglucuronosyltransferase (UGTs), sulfotransferases (SULTs), or N-acetyltransferases (NATs). Drugs may also be cleared through transport by enzymes such as p-glycoprotein (P-gp), breast cancer resistance protein (BCRP), or the organic-peptide-transporting polypeptides (OATPs). Protein-based therapeutics, such as antibodies, have other routes of clearance such as catabolism. The relative contribution of clearance mechanisms for the 200 most prescribed drugs of 2016 is listed in Fig. 2a, where metabolism is the primary mechanism of clearance, followed by renal excretion. Of those drugs undergoing metabolism, approximately 70% are cleared by CYPs (Fig. 2b). The CYP3A enzymes are responsible for approximately 55% of the CYP-mediated metabolism (Fig. 2c). Thus, the CYPs, and particularly CYP3A, are the primary enzymes responsible for the metabolism of many of the highly prescribed drugs of 2016.

Methods and Assumptions

Brief Primer on In Vitro Enzyme Kinetics

A typical approach to characterizing enzyme kinetics in vitro is to examine product formation at multiple substrate concentrations. The results are then visualized in graph of substrate concentration versus reaction velocity (ν). Under Michaelis-Menten conditions, it is assumed that the enzyme (E), substrate (S), and enzyme-substrate complex (ES) are in rapid equilibrium and that product (P) formation is irreversible to give the following reaction scheme (Michaelis and Menten 1913).

$$E + S \leftrightarrow ES \rightarrow E + P$$

Mathematical representation of Michaelis-Menten kinetics provides the following equation:

$$\nu = \frac{(V_{\max}{}^{*}[S])}{(K_m + [S])} \quad (1)$$

where ν is the reaction velocity, $V_{\max}$ is the maximal reaction velocity, $[S]$ is the substrate concentration, and K_m is the substrate concentration at half-maximal velocity. Visualization (graph of ν versus $[S]$) of in vitro data under Michaelis-Menten conditions produces a hyperbolic curve (Fig. 3a) and a straight line in an Eadie-Hofstee visualization (Frere et al. 1983) of $\nu/[S]$ versus ν (Fig. 3e).

Atypical (non-Michaelis-Menten) kinetic profiles may be observed for some drug-metabolizing enzymes (DMEs). For experiments where substrate alone is present, apparent autoactivation or biphasic or substrate inhibition kinetics may be observed (Hutzler and Tracy 2002). It has been proposed that atypical kinetic phenomena occur because two or more substrates are present in the DME active site at the same time (Korzekwa et al. 1998). Atypical kinetics is more likely observed for DMEs with a large active site such as CYP2C9 and CYP3A4. Autoactivation occurs when a

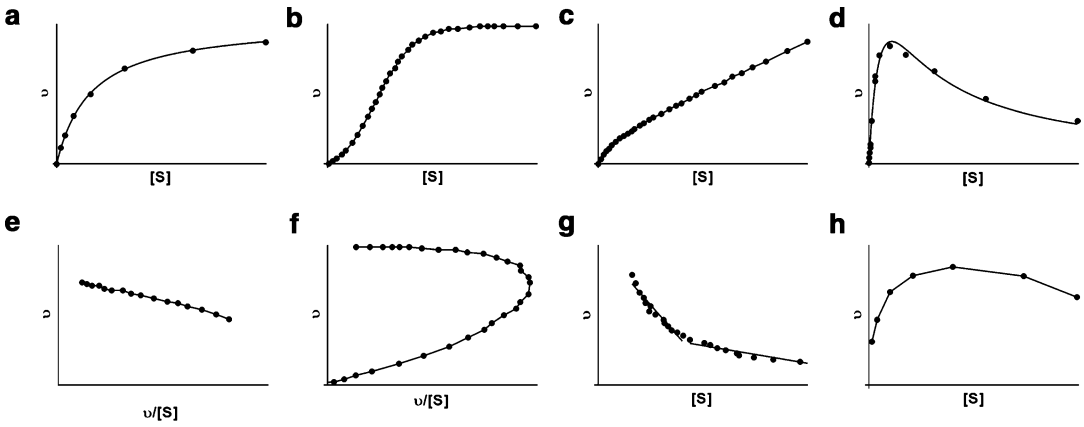


Fig. 3 Velocity versus substrate and Eadie-Hofstee graphs for hyperbolic kinetics (**a**, **e**), sigmoidal kinetics (**b**, **f**), biphasic kinetics (**c**, **g**), and apparent substrate inhibition kinetics (**d**, **h**)

substrate enhances its own metabolism as the substrate concentration increases. Autoactivation results in a sigmoidal visualization of ν versus $[S]$. Autoactivation often begins to occur at low substrate concentrations and may be difficult to observe in the ν versus $[S]$ visualization (Fig. 3b); a hook in the Eadie-Hofstee plot is a more readily observed diagnostic for the presence of autoactivation (Fig. 3f). The Hill equation may be used to estimate kinetic parameters:

$$\nu = \frac{V_{\max}{}^{*}[S]^{n}}{S_{50}{}^{n} + [S]^{n}} \tag{2}$$

where n is an exponent that indicates how far the observed kinetics are deviating from a standard value of 1 and S_{50} is the substrate concentration at which half-maximal reaction velocity has been achieved.

Biphasic conditions occur when enzyme architecture allows for both a low and a high affinity site for metabolism on the same enzyme for a single substrate; the reaction rate does not saturate and instead proceeds linearly as substrate concentrations are increased. Biphasic kinetic data may be modeled using the following equation (Kumar et al. 2006a):

$$\nu = \frac{(V_{\max 1}{}^{*}[S]) + \left(CL_{\text{int}}{}^{*}[S]^{2}\right)}{(K_{m1} + [S])} \tag{3}$$

where K_{m1} and $V_{\max 1}$ are the kinetic parameters for the high affinity contribution to the enzyme reaction and CL_{int} is used to represent the intrinsic clearance of the linear portion of the kinetics (Fig. 3c). Two distinct contributions to turnover are readily visible in the Eadie-Hofstee plot for this type of kinetics (Fig. 3g). Biphasic kinetics may also be observed in multienzyme systems when two or more enzymes are responsible for the turnover of a substrate that exhibit different K_m estimates.

Substrate inhibition kinetics occur when increasing substrate concentrations causes a decrease in product formation (Fig. 3d). Eadie-Hofstee visualization of this type of kinetics exhibits decreasing ν as $\nu/[S]$ increases (Fig. 3h) and may be represented by the equation (Lin et al. 2001):

$$\nu = \frac{V_{\max}}{1 + \frac{K_m}{[S]} + \frac{[S]}{K_i}} \tag{4}$$

where K_i represents the binding of the substrate to the inhibitory site.

Types of DDIs

There are three primary mechanisms of PK-based drug-drug interactions: reversible inhibition,

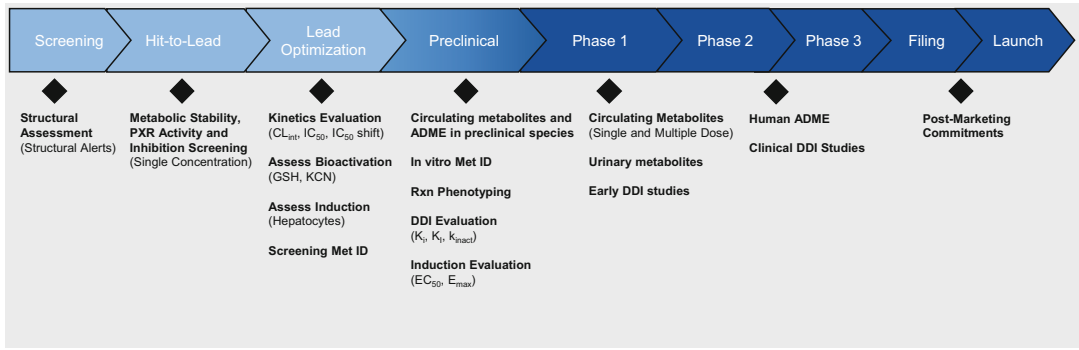


Fig. 4 Key drug information to enable characterization of DDI risk

time-dependent inhibition, and induction (Wahlstrom et al. 2006). Reversible inhibitors interact with an enzyme's architecture through non-covalent bonding, which may occur within or external to the enzyme active site. Time-dependent inhibitors are characterized by a change in observed inhibition potency over time, most often caused by metabolism-mediated enzyme inhibition (such as formation of an inhibitory metabolite) or inactivation. The primary concern with enzyme inhibition or inactivation is an increase in victim drug levels to a point where adverse effects may begin to occur. Induction occurs due to increases in protein expression levels generally caused by an increase in gene transcription. In contrast to inhibition, induction reduces victim drug levels, where the primary concern is loss of efficacy. Characterization of the potential for inhibition and induction occur throughout the drug discovery and development process, where initial single concentration screening experiments are followed up using higher content experiments that may be used to predict DDI effects in the clinical situation. A typical paradigm for information gathering relevant to predicting clinical DDIs is shown in Fig. 4. The higher content experiments include the IC_{50} estimation, where the effects of multiple concentrations of inhibitor are determined relative to a single concentration of substrate (at K_m); the K_i (inhibitory constant) estimation, where multiple substrate and inhibitor concentrations are tested to fully characterize the type of inhibition observed; and the EC_{50} and $E_{\max}$ experiment for induction, where multiple inducer concentrations are tested to determine their maximal effect of induction ($E_{\max}$) and the inducer concentration at half-maximal induction (EC_{50}). In addition to direct DDI characterization, an understanding of the mechanism(s) of clearance, the enzymes responsible for metabolism, the presence of metabolites (particularly circulating metabolites), and likely administered co-medications based on therapeutic area are key pieces of information to underwrite the potential risk of DDIs.

Reversible inhibition may be further differentiated by types, which include competitive inhibition, noncompetitive inhibition, uncompetitive inhibition, mixed inhibition, and atypical or multisite inhibition. A summary of inhibition kinetic characteristics is shown in Table 1. Competitive inhibitors compete with a substrate for binding to enzyme, where enzyme binding of the substrate or inhibitor is mutually exclusive. This competition is often for the enzyme active site. The presence of a competitive inhibitor raises the apparent K_m for the substrate, as increased substrate concentration is needed to outcompete the inhibitor for the active site. Nonlinear fitting of in vitro data (Eq. 5 for competitive inhibition) is now commonly used to estimate inhibition potency and assess inhibition mechanism; [I] is the concentration of the inhibitor in the equations. A Lineweaver-Burk visualization (graph of $1/[S]$ versus $1/\nu$, Fig. 5a) aids in determining inhibition type, as linear regression of the inhibition data at each substrate concentration will intersect at a common point

Table 1 Inhibition types and characterization

Inhibition type	Inhibition characteristics	Apparent effects	Lineweaver-Burk (1/[S] vs. 1/v)
Competitive	S and I compete for E binding	$K_m\uparrow$, $V_{max}\leftrightarrow$	Intersect on y axis
Noncompetitive	I binding alters E architecture, reducing activity	$K_m\leftrightarrow$, $V_{max}\downarrow$	Intersect on x axis
Uncompetitive	I binds only to the ES complex	$K_m\downarrow$, $V_{max}\downarrow$	Parallel lines
Mixed	I binds to E and ES	$K_m\downarrow$ or $\uparrow$, $V_{max}\downarrow$	Intersect with $x < 0$ and $y > 0$
Multisite or atypical	Multiple ES complexes possible	Situational	NA

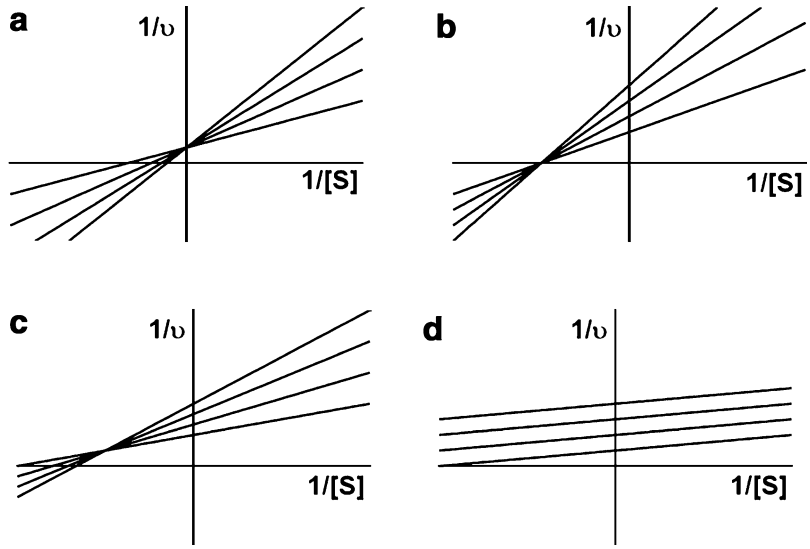


Fig. 5 Lineweaver-Burk plots for competitive (**a**), noncompetitive (**b**), mixed (**c**), and uncompetitive inhibition (**d**)

on the y-axis of the visualization (Lineweaver and Burk 1934).

$$v = \frac{V_{max} \bullet [S]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]} \tag{5}$$

Noncompetitive inhibition, in contrast, results when an inhibitor has equal affinity for the enzyme, regardless of whether or not substrate is bound, but reduces enzyme activity, typically through allosteric alteration of the enzyme architecture. Inhibition potency for noncompetitive inhibition may be determined using Eq. 6. In a Lineweaver-Burk visualization, linear regression of the inhibition data at each substrate concentration characteristically intersects on the x-axis (Fig. 5b).

$$v = \frac{V_{max} \bullet [S]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]\left(1 + \frac{[I]}{K_i}\right)} \tag{6}$$

Uncompetitive inhibition occurs when the inhibitor binds only the enzyme-substrate (ES) complex and not the enzyme itself. True uncompetitive inhibition is rarely observed in practice. Inhibition potency for an uncompetitive inhibitor may be estimated using Eq. 7. Lineweaver-Burk visualization results in diagnostic parallel lines that are parallel and do not intersect (Fig. 5c).

$$v = \frac{V_{max} \bullet [S]}{K_m + [S]\left(1 + \frac{[I]}{K_i}\right)} \tag{7}$$

Table 2 Mechanisms of CYP inactivation with diagnostic experiments

Inactivation mechanism	Loss of activity	Loss of CO binding	Loss of native heme	Other diagnostics	Prototypical example
Apoprotein adduction	Yes	Possibly	No	Formation of GSH, cysteine, lysine, or cyanide adducts	Raloxifene (CYP3A4)
Heme destruction	Yes	Yes	Yes	NA	Mibefradil (CYP3A4)
Heme adduction	Yes	Possibly	Yes	NA	Gemfibrozil glucuronide (CYP2C8)
MIC formation	Yes	Yes	No	Appearance of peak at 440–450 nm (UV–vis spectrum); $Fe(CN)_6$ restores activity	Verapamil (CYP3A4)

Mixed inhibition occurs when the inhibitor binds to the enzyme and the ES complex, but the binding is stronger to one of the reaction constituents. Inhibition potency may be estimated through nonlinear regression of Eq. 8. A Lineweaver-Burk representation of mixed inhibition has the intersection of linear regression lines in the upper left quadrant of the graph as a characteristic (Fig. 5d).

$$v = \frac{V_{\max} \bullet [S]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]\left(1 + \frac{[I]}{K_i'}\right)} \tag{8}$$

Analogous to the observations with substrates, enzymes with large or flexible active sites may exhibit atypical or two-site inhibition profiles, where the substrate and the inhibitor occupy the enzyme active site simultaneously. This simultaneous occupancy of the active site may result in partial inhibition, where an inhibitor incompletely inhibits the turnover of a substrate, even as the inhibitor level is increased. Two-site models may also be applied to in vitro inhibition data where fitting of competitive, noncompetitive, uncompetitive, and mixed inhibition models to observed data is poor. Two-site inhibition is one potential explanation for substrate-dependent inhibition, where the same inhibitor exhibits different apparent inhibition potency for the same enzyme when compared using different substrates (VandenBrink et al. 2012).

Time-dependent inhibition (TDI) refers to a general increase in inhibition potency over time. TDI may be caused by slow binding or access of the inhibitor to the enzyme or other, more common mechanisms including metabolism-dependent inhibition (MDI) or mechanism-based inactivation (MBI). MDI includes the formation of inhibitory or inactivating metabolites that are metabolism-dependent. MBI is characterized by irreversible or quasi-irreversible enzyme inactivation, which include mechanisms such as heme destruction, heme alkylation, metabolite intermediate complex (MIC) formation, or heme alkylation (Foti et al. 2011). Experiments to determine the mechanism of MBI and the corresponding diagnostic results for CYPs are outlined in Table 2. Seven criteria have been proposed by Silvermann to characterize a drug as a MBI, including time dependence, saturation of inactivation at high inactivator concentrations, protection of inactivation by the presence of substrate, irreversibility, 1:1 stoichiometry of inactivator to enzyme, and metabolism dependence; the inactivation must also occur before the reactive species exits the active site (Silverman 1995). A parameter used to characterize the efficiency of inactivation is the partition ratio, which is the sum of all metabolic events divided by the number of inactivation events (Kunze and Trager 1993). A low partition ratio indicates very efficient inactivation of an enzyme.

TDI may be assessed in vitro through the use of a preincubation, where an aliquot of preincubate is

added to initiate the probe substrate reaction after a predetermined period of time. For the single point experiment, activity loss is determined by the following equation:

$$\%\text{Loss} = 100 \bullet \left[\left(\frac{\text{Activity}_{\text{inactivator}}}{\text{Activity}_{\text{control}}} \right)_{\text{NoCofactor}} - \left(\frac{\text{Activity}_{\text{inactivator}}}{\text{Activity}_{\text{control}}} \right)_{\text{WithCofactor}} \right] \quad (9)$$

An IC_{50} shift experiment takes advantage of preparations needed for a reversible IC_{50} experiment but adds additional information through a preincubation step. The IC_{50} shift output (Fig. 6a) indicates TDI by a left shift in potency for the preincubated samples. For the high content TDI experiment, multiple concentrations of inhibitor and multiple time points are used to characterize TDI activity. The acquired data is transformed to natural log (ln) and linear regression of the loss of activity over time at each TDI concentration is calculated (Fig. 6b) to obtain the observed, pseudo-first-order rate constant of inactivation (k_{obs}). If the time-dependent inhibition is saturable with increasing concentrations of inhibitor, kinetics parameters can be estimated with an equation that is an analog of that used to obtain Michaelis-Menten kinetic parameters (Eq. 10, Fig. 6c):

$$k_{obs} = \frac{(k_{inact}{}^{*}[I])}{(K_I + [I])} \quad (10)$$

where k_{inact} is the maximal rate of inactivation and K_I is the inhibitor concentration at half-maximal inhibition.

Enzyme induction typically involves an increase in enzyme transcription (Lin 2006). For DMEs, the process often involves ligand binding to one of three nuclear receptors: the aryl hydrocarbon receptor (AHR), the constitutive androstane receptor (CAR), or the pregnane X receptor (PXR). CAR and PXR are primarily expressed in the gut and liver, sites that are co-localized with high expression levels of DMEs. PXR is the receptor of main concern, as it is a major transcriptional regulator of CYP3A4

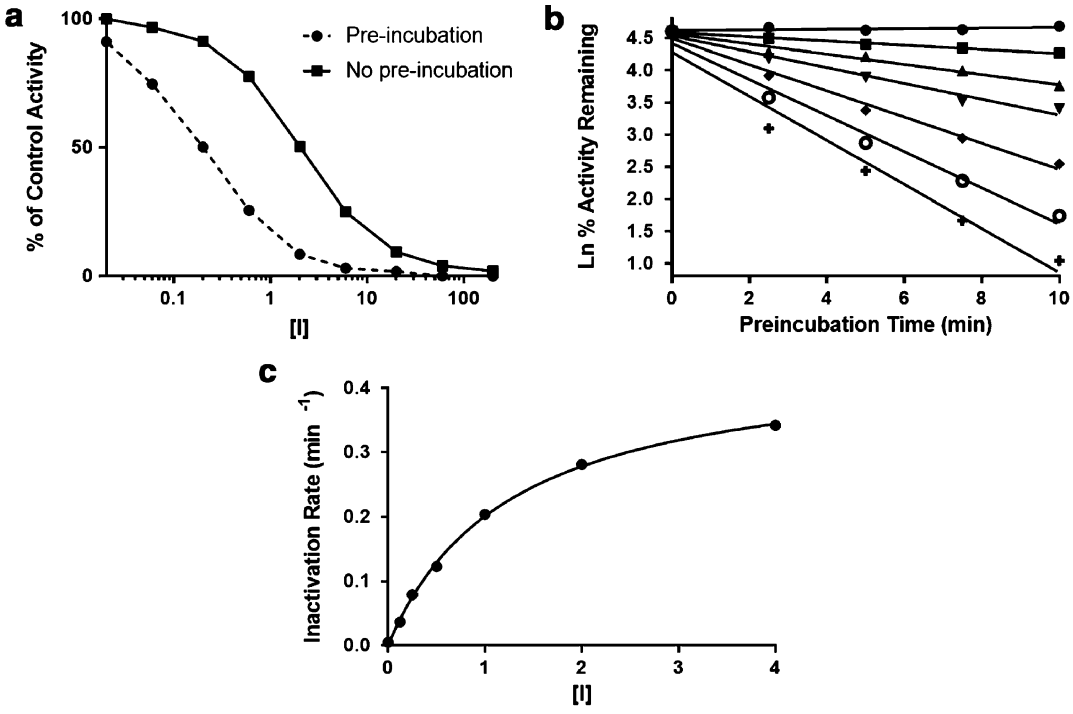


Fig. 6 Representative TDI experiments: IC50 shift (**a**), percent activity remaining (**b**), and K_I, k_{inact} (**c**)

and P-gp protein expression. Enzyme induction is a time-dependent process thought to involve ligand binding to the receptor and translocation of the bound receptor to the cell nucleus, where binding to DNA sequences in the promoter region causes the recruitment of coactivators, thus increasing protein transcription rates and the quantity of active enzyme available for drug metabolism. Early screening typically involves cell lines engineered to encode the PXR-binding domain. Interaction with PXR ligands activates the receptor, leading to an increase in luciferase production that may be used as a surrogate for functional PXR activity. The gold standards for characterizing DME induction are experiments using human hepatocytes. Increases in mRNA levels for individual DMEs are used to determine a maximal response (E_{max}), as well as the inducer concentration required for half-maximal activity (EC_{50}) using the following equation:

$$\text{Effect} = \frac{(E_{max}{}^{*}[\text{Inducer}])}{(EC_{50} + [\text{Inducer}])} \qquad (11)$$

These parameters may be used to quantitatively estimate induction potential in the clinical situation. As the induction assay requires fully functional hepatocytes, a commonly encountered confounding factor in the determination of induction parameters is cell toxicity; toxicity is commonly observed with increasing concentration for cytotoxic drugs and may prevent full characterization of induction potential.

The presence of inflammatory cytokines may lead to a process called downregulation, which has the opposite effect of induction. Downregulation is most commonly seen with inflammatory diseases, where IL-6 and other cytokines may lead to an apparent reduction in transcription and expression of DMEs, reducing the turnover of DME substrates and leading to higher drug levels (Evers et al. 2013). Treatment of the inflammatory disease may reduce cytokine-mediated influence on DMEs, leading to increased transcription, protein synthesis, and DME turnover, leading to lower drug levels. Due to the complexities in the downregulation process, neither in vitro nor preclinical models may serve as predictors of the clinical situation with the current state of knowledge (Dickmann et al. 2011).

Experimental Considerations In Vitro

Both the FDA and EMA DDI Guidance provide details on the design of in vitro experiments informing DDI studies. A major factor in assessing the fraction of metabolism (f_m) of an NCE (DDI potential as a victim) and in designing reversible in vitro inhibition experiments (DDI potential as a perpetrator) is selecting the appropriate enzyme source. Different sources of enzymes for in vitro experiments include subcellular fractions such as human liver microsomes, S9 or cytosolic liver fractions, recombinant enzymes, hepatocytes, or organ slices. Each of these in vitro systems contains a different compliment of DMEs and requires specific incubation conditions for optimal activity. Centrifugation of liver tissue homogenate produces three subcellular fractions: microsomes, which are isolated from the cell pellet after centrifugation; S9 fractions, which are the soluble enzyme constituents after centrifugation; and cytosolic fractions, which are obtained from the supernatant after resuspension of the pellet and a second centrifugation. Due to ease of use and availability of pooled lots across a large number of donors, human liver microsomes are often selected as the relevant in vitro source of enzyme to determine inhibition potency for membrane-bound DMEs that reside in the endoplasmic reticulum, such as the CYPs, FMOs, and UGTs. Relevant cofactors, such as the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH) or uridine 5′-diphosphoglucuronic acid (UDPGA) for CYPs/FMOs or UGTs, respectively, must be added to microsomal incubations to initiate and sustain enzyme activity. Additives, such as the pore-forming agent, alamethicin, are added to overcome latency of the UGTs to observe in vitro activity. Other additives, such as magnesium chloride, are often added to microsomal preparations to help stabilize CYP structure to support maximal activity. Liver S9 fractions contain soluble DME constituents, which may be appropriate

for examining the effect of inhibitors on soluble enzymes such as aldehyde oxidase (AO), epoxide hydrolase (EH), sulfotransferases (SULTs), or carboxylesterases (CES); S9 preparations may contain some residual CYP content as well. Cytosolic fractions contain soluble enzymes such as AO and the SULTs. Recombinant, expressed enzymes are often used to aid in determining which DMEs are involved in the metabolism of an NCE; they are less commonly used to determine inhibition potency, as the expression systems are artificial (typically expressed in a baculovirus vector), tend to overexpress enzyme relative to liver fractions, and may lack protein-protein interactions that exist in subcellular liver fractions. Hepatocytes contain the full complement of DMEs and the prerequisite cofactors for DME activity but are more difficult to prepare and use. The preparation of liver or other organ slices is technically challenging, which is why they are rarely used. Integration of the in vitro metabolism data in its entirety allows for initial assessment of f_m; regulatory guidance recommends follow-up for instances where greater than 25% of DME metabolism is believed to proceed by a single metabolism pathway (Fig. 7).

The Pharmaceutical Research and Manufacturers of America (PhRMA) have published an overview on the conduct of in vitro DDI studies (Bjornsson et al. 2003). Key factors identified in experimental design are reaction conditions including solvent effects, protein concentration, nonspecific binding to the incubation matrix, and probe substrate selection. The manuscript recommends screening initial reaction conditions so that the enzymatic reactions are linear with regard to both time and protein concentration. Low protein levels minimize inhibitor depletion in the incubations and reduce the impact of nonspecific binding to the incubation matrix. Measurement of the unbound drug fraction in the incubation matrix is often used to correct in vitro measurements for the prediction of clinical DDIs and therefore is a useful parameter to measure in vitro. Low organic solvent concentrations, particularly of DMSO, are used to minimize possible effects on DME activity. The PhRMA publication outlines the use of selective probe substrate reactions to assess inhibitory potential for individual DMEs. A list of commonly used selective probe substrates and inhibitors for CYP-mediated DDI are shown in Table 3. The FDA and EMA DDI Guidance are updated on a regular basis and provide another source for commonly used probe substrates. Due to its large active site and the potential for atypical kinetics, two probe substrates are recommended to assess inhibitory potency for CYP3A4 from both an FDA and EMA perspective (specifically midazolam and testosterone based on current EMA DDI Guidance). For situations where large differences in inhibition potency are observed

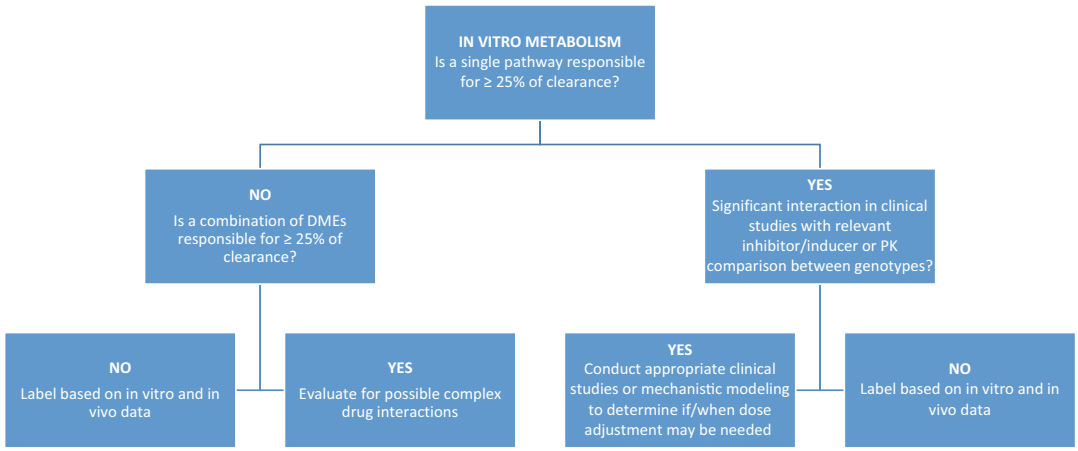


Fig. 7 Decision tree to determine victim DDI potential

Table 3 Typical CYP-selective probe substrates, inhibitors, and inactivators

CYP isoform	Substrate	K_m (μM)	Inhibitor	K_i (μM)	Inactivator	K_I (μM)	K_{inact} (min^{-1})
1A2	Phenacetin	20	α-Napthoflavone	0.016	Furafylline	4.4	0.14
2B6	Bupropion	50	Clotrimazole	0.022	Ticlopidine	0.32	0.43
2C8	Paclitaxel	5	Montelukast	0.004	Gemfibrozil glucuronide	35	0.022
2C9	Diclofenac	4	Sulfaphenazole	0.041	Tienilic acid	1.0	0.17
2C19	(S)-Mephenytoin	20	(+)-N-Benzylnirvanol	0.027	Ticlopidine	5.3	0.077
2D6	Dextromethorphan	4	Quinidine	0.012	Paroxetine	1.1	0.11
2E1	Chlorzoxazone	50	Diethyldithiocarbamate	16	Disulfiram	12	0.02
3A4/5	Midazolam	2	Ketoconazole	0.002	Mifepristone	3.6	0.079
3A4/5	Testosterone	20	Ketoconazole	0.005	Mifepristone	3.6	0.079

between the two probes, the most potent inhibition should be used for prediction of clinical DDIs (Foti et al. 2010).

PhRMA has also published an overview on the conduct and design of experiments to characterize TDI (Grimm et al. 2009). For TDI experiments, the additional experimental factors of concern are protein concentration in the preincubation, preincubation time, incubation time, dilution factor, and concentrations of the inactivator and probe substrate. Protein concentration in the preincubation is important because a balance of turnover versus nonspecific binding is desired; increasing protein will not necessarily translate to an increase in turnover or inactivation rate. Preincubation time is important, as it is generally desired to use initial rates for the inactivation calculations; too little or too much inactivation may increase the difficulty in calculating these kinetic parameters. For IC_{50} shift experiments, dilution from a preincubation is not done, so that the substrate concentration remains the same (at k_m) for the incubations with and without preincubation. Dilution is normally carried out for the K_i and k_{inact} experiment, as it is desired to minimize the potential for reversible inhibition; the dilution also reduces the potential for inactivation during the probe substrate reaction. High substrate concentrations are used (often five- to tenfold of k_m), to ensure that sufficient substrate is present to outcompete remaining inactivator from the enzyme active site.

For in vitro induction experiments, key features identified by PhRMA include the exposure time of the hepatocytes to the NME, the concentrations of the NME used, the potential of the NME to cause cellular toxicity, and the experimental output to measure (Chu et al. 2009). Regulatory guidance from EMA and FDA both recommend the use of three or more hepatocyte donors; the hepatocytes may be cryopreserved or fresh. Changes in mRNA levels versus positive controls are the assay readout for development studies; DME activity may be assessed in some cases as a secondary endpoint. Regulatory guidance from EMA guidance recommends exposure to test article over 72 h. Both FDA and EMA guidance recommend refreshing exposure to the NME daily; EMA recommends measuring the amount of parent drug from the incubation at several times on the final day of the experiment. Single concentrations of NME may be used to characterize induction potential in the discovery environment, while multiple NME concentrations are used to characterize EC_{50} and E_{max} for the development environment. EMA guidance recommends in vitro concentrations exceeding 50-fold the mean unbound C_{max} value for DMEs present in the liver, with in vitro concentrations of NME exceeding 0.1*dose/250 mL as an estimate for DMEs present in the gut (CYP3A). As the in vitro concentration of NME increases, the likelihood of cell toxicity may increase. Simultaneous loss of cell viability and reduction in mRNA levels are expected if cell viability alone is the cause; a reduction in mRNA levels without loss of cell viability may indicate downregulation of the DME by the NME.

Prediction of DDI for the Clinical Situation

PK Principles

Organ clearance models are used to estimate intrinsic clearance, as with the following equation:

$$CL = \frac{(Q_{\text{organ}} {}^{*} f_u {}^{*} CL_{\text{int}})}{(Q_{\text{organ}} + f_u {}^{*} CL_{\text{int}})} \tag{12}$$

where CL is organ clearance, Q is the blood flow to a particular organ, f_u is the fraction unbound in plasma, and CL_{int} is the intrinsic organ clearance of unbound drug (Iwatsubo et al. 1997). High extraction ratio drugs ($Q_{\text{organ}} < < f_u{}^{*}CL_{\text{int}}$) exhibit clearance that is independent of f_u.

$$CL \cong Q_{\text{organ}} \tag{13}$$

Low extraction ratio drugs ($Q_{\text{organ}} > > f_u{}^{*}CL_{\text{int}}$) are dependent upon both the f_u and the intrinsic clearance for activity.

$$CL \cong f_u {}^{*} CL_{\text{int}} \tag{14}$$

For orally dosed drugs, exposure is defined as the area under the plasma concentration-time curve (AUC).

$$AUC_{\text{oral}} = \frac{(F {}^{*} \text{Dose})}{Cl} \tag{15}$$

where bioavailability (F) is defined as

$$F = Fa {}^{*} Fg {}^{*} Fh \tag{16}$$

F_a is the fraction of drug absorbed through the gut wall, F_g is the fraction of drug that escapes the intestine unchanged, and F_h is the amount of drug that escapes from the liver unchanged. For oral drugs eliminated primarily by the liver, the Eqs. 15 and 16 can be combined such that

$$AUC_{\text{oral}} = \frac{(Fa {}^{*} Fg {}^{*} \text{Dose})}{fu {}^{*} CL_{\text{int}}} \tag{17}$$

This equation holds true for both high and low extraction ratio drugs (Benet and Hoener 2002).

Quantitative Prediction of Clinical DDIs

Equations for the reversible inhibition situation already been introduced. Under conditions observed in the clinic, often the concentration of substrate is much lower than the apparent K_m, such that

$$CL_{\text{int}} = \frac{\nu}{[S]} = \frac{V_{\max}}{K_m} \tag{18}$$

For situations with competitive inhibition, the following equation applies:

$$CL_{\text{int},I} = \frac{V_{\max}}{K_m \left(1 + \frac{[I]}{K_i}\right)} \tag{19}$$

Eqs. 18 and 19 can be combined to estimate a change in intrinsic clearance in the inhibited and uninhibited state:

$$\frac{CL_{\text{int}}}{CL_{\text{int},I}} = 1 + \frac{[I]}{K_i} \tag{20}$$

Converting this to a ratio of the area under the plasma-concentration time curves results in the equation

$$AUCR = \frac{AUC_i}{AUC} = 1 + \frac{[I]}{K_i} \tag{21}$$

For drugs with more than one pathway of clearance, Eq. 12 can be rewritten as

$$\begin{aligned} CL_{\text{int}} &= \frac{V_{\max,1}}{K_{m,1}} + \frac{V_{\max,2}}{K_{m,2}} \\ &= fm_1 {}^{*} CL\text{int} + (1 - fm_1) * CL_{\text{int}} \end{aligned} \tag{22}$$

where f_m is the fraction metabolized by a specific pathway (Ito 2005). If only one pathway is inhibited, then

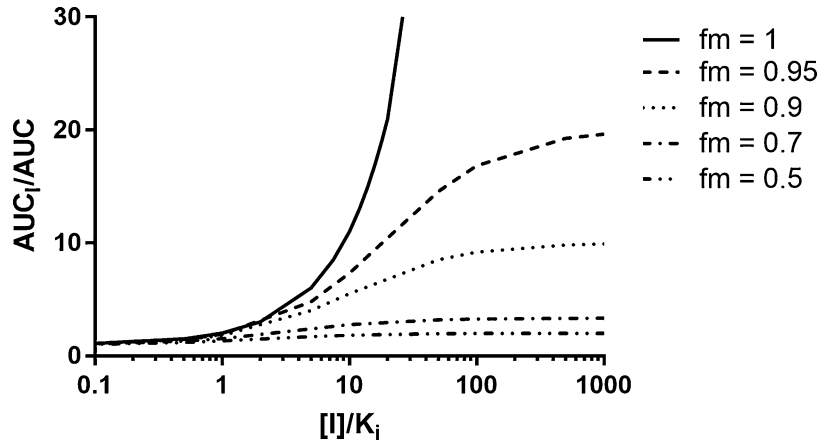


Fig. 8 Impact of f_m and inhibition potency on expected DDI magnitude

$$CL_{\text{int},I} = \frac{V_{\max,1}}{K_{m,1}\left(1+\frac{[I]}{K_i}\right)} + \frac{V_{\max,2}}{K_{m,2}} = \frac{fm_1{}^{*}CL_{\text{int}}}{1+\frac{[I]}{K_i}} + (1-fm_1)*CL_{\text{int}} \quad (23)$$

To obtain a ratio in the uninhibited and inhibited state, this becomes

$$\frac{CL_{\text{int},I}}{CL_{\text{int}}} = \frac{fm_1}{1+[I]/Ki} + (1-fm_1) \quad (24)$$

Equation 24 can be converted from intrinsic clearance to AUC ratio to give

$$AUCR = \frac{1}{\frac{fm_1}{1+[I]/K_i} + (1-fm_1)} \quad (25)$$

Equation 25 summarizes the key drivers to observed inhibition in the clinic. The magnitude of DDI effect is dependent upon characteristics of both the perpetrator and victim. For the victim drug, the fraction metabolized (f_m) by the inhibited pathway is a key driver of the magnitude of the DDI effect; the higher the f_m, the larger the potential magnitude of the DDI (Fig. 8). Key features of the perpetrator include the inhibition potency for the particular DME, the unbound fraction, and the free concentration of the inhibitor at the site of inhibition.

Regulatory Guidance and DDIs

Both the EMA and FDA Guidance propose a tiered approach for DDI prediction using basic models, mechanistic static models, and then complex models such as physiologically based pharmacokinetic modeling (PBPK). For basic models, a decision tree scheme for inhibition in line with FDA and EMA Regulatory Guidance is shown in Fig. 9. For reversible inhibition, the following equation holds:

$$AUCR = 1 + [I]/K_i \quad (26)$$

where $AUCR$ is the AUC ratio in the presence and absence of an inhibitor and the inhibitor concentration [I] is the unbound C_{max} value for liver-based DDIs or is calculated using the relationship of dose/250 mL for gut-based DDIs. Exceeding recommended cutoff values ($AUCR$ >1.02 for liver-based DDI and AUCR > 11 for gut-based DDIs) triggers a potential clinically relevant DDI result and a move down the decision tree to more complex mechanistic static models (MSM) or PBPK.

A similar decision tree for TDI is also shown in Fig. 9; the following equation is used to estimate DDI risk, where K_{deg} is the degradation rate of the

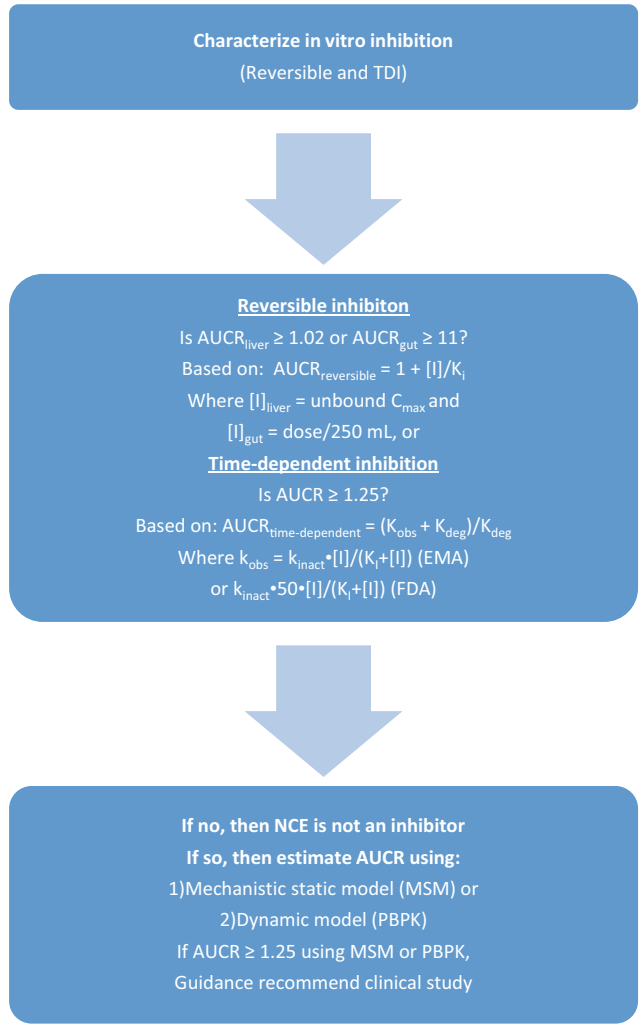


Fig. 9 Basic model decision tree for inhibitory DDIs

enzyme of interest. The same [I] definitions apply for this equation as for reversible inhibition, but the cutoff is $AUCR \geq 1.25$.

$$AUCR = \frac{\left(k_{obs} + K_{\mathrm{deg}}\right)}{K_{\mathrm{deg}}} \tag{27}$$

For induction, a decision tree in line with the FDA and EMA DDI Guidance in shown in Fig. 10. The following equation applies, where *AUCR* is the AUC ratio in the presence and absence of an inducer. In this instance, *AUCR* <0.8 triggers moving to the next step of the decision tree:

$$AUCR = \frac{1/(1 + dose * E_{\max}{}^{*}[I])}{(EC_{50} + [I])} \tag{28}$$

If a DDI is deemed possible by the basic models, the next step is to assess DDI risk through MSM or dynamic models, especially PBPK. While MSMs incorporate the unbound fraction in plasma to become closer to physiological relevance, the inhibitor concentration is assumed to

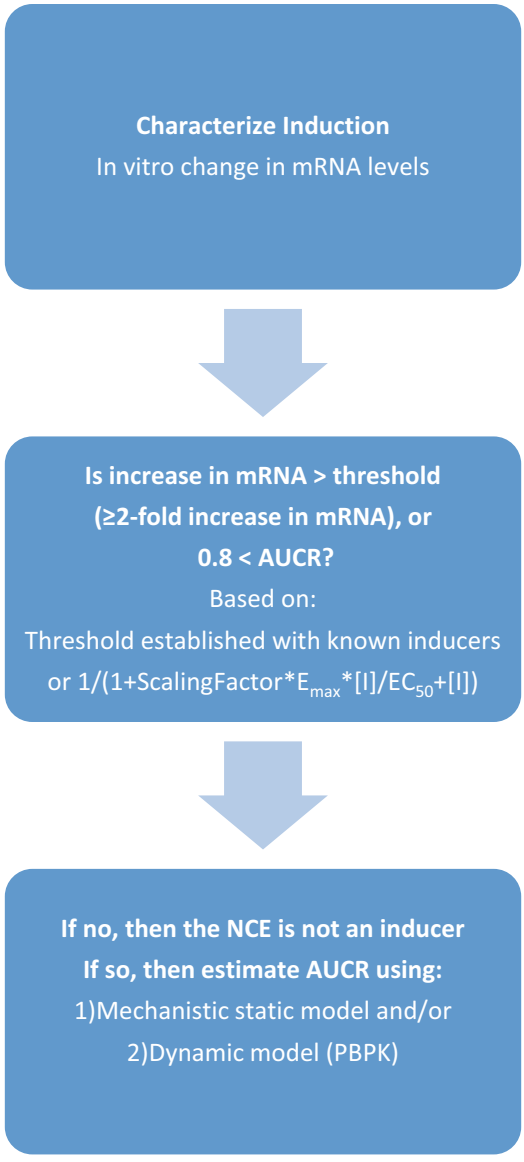


Fig. 10 Basic model decision tree for induction-based DDIs

be constant and maximal throughout the dose interval. This is a conservative assumption that may not accurately represent the clinical situation for many drugs. Estimation of the overall DDI magnitude is accomplished using Eq. 29 below, which integrates both the expected gut and liver contributions to DDI, as well as the potential contribution of inhibition, inactivation and induction.

$$AUCR = \left(\frac{1}{\text{Inhib}_g{}^{*}\text{Inact}_g{}^{*}\text{Induct}_g{}^{*}\left(1-F_g\right)+F_g}\right) \times * \left(\frac{1}{\text{Inhib}_h{}^{*}\text{Inact}_h{}^{*}\text{Induct}_h{}^{*}fm+\left(1-fm\right)}\right) \tag{29}$$

The equations for inhibition ($\text{Inhib}_{\text{organ}}$), inactivation ($\text{Inact}_{\text{organ}}$), and induction ($\text{Induct}_{\text{organ}}$) are shown below, respectively:

$$\text{Inhib}_{\text{organ}} = \frac{1}{1+\frac{[I]_{\text{organ}}}{K_i}} \tag{30}$$

$$\text{Inactiv}_{\text{organ}} \frac{K_{\text{deg, organ}}}{K_{\text{deg, organ}} + \frac{[I]_{\text{organ}} * k_{\text{inact}}}{[I]_{\text{organ}} + K_I}} \tag{31}$$

$$\text{Induct}_{\text{organ}} = 1 + \frac{\text{dose} * E_{\max}{}^{*}[I]_{\text{organ}}}{\left(EC_{50} + [I]_{\text{organ}}\right)} \tag{32}$$

The concentrations in the gut and liver are defined by the following equations, where $[\text{I}]_{\text{max, b}}$ is the maximum concentration of the inhibitor in blood.

$$[I]_{\text{gut}} = \frac{Fa^{*}ka^{*}\text{Dose}}{Q_{\text{enterocyte}}} \tag{33}$$

$$[I]_{\text{liver}} = fu, b^{*}\left([I]_{\max, b} + \frac{Fa^{*}ka^{*}\text{Dose}}{Q_{\text{liver}}}\right) \tag{34}$$

An additional option is PBPK modeling. PBPK is a modeling technology that has seen recent emergence both for internal decision-making for compound progression through drug discovery and development and for regulatory applications (Jamei et al. 2009). PBPK models have three main types of input parameters: demographic and genetic information (i.e., age or gender), physiological information (i.e., organ blood

flow and enzyme levels), and drug-specific parameters (i.e., pKa, logP, solubility). Differential equations and Monte Carlo-based simulations integrate the inputs together to simulate a variety of outcomes, including plasma concentration-time profiles, enzyme activity profiles, and drug-tissue concentrations. Because of the types of inputs and the modeling technique used, PBPK is well-suited for modeling where changes in physiology or populations may impact PK, changes in physicochemical properties or formulations may impact PK, or where dynamic simulations such as DDIs are desired. PBPK may be used to answer fundamental clinical pharmacology based questions such as (1) what are the intrinsic factors that may influence exposure, (2) what are the extrinsic factors that may influence exposure, and (3) what are situations in which dosing may need to be adjusted due to intrinsic and extrinsic factors. Because physiologically relevant parameters are included, PBPK may more closely represent the clinical situation than basic or MSM models. PBPK may also provide information on expected variability in studies based on demographic factors.

The potency of an inhibitor or inducer is determined based on the magnitude of its interaction with a sensitive probe substrate for a specific enzyme pathway. Strong inhibitors increase AUC $\geq$ fivefold, moderate inhibitors increase AUC two- $\geq$ to $<$ fivefold, and weak inhibitors increase AUC $\geq$ 1.25- to $<$ twofold. Strong inducers reduce AUC $\geq$ 80%, moderate inducers reduce AUC by $\geq$50 to $<$80%, and weak inducers reduce AUC by $\geq$20 to $<$50%.

Additional Considerations

Free Drug Hypothesis

The free drug hypothesis is a tenet of drug discovery and development, which has two fundamental premises (Smith et al. 2010) First, at steady state, drug concentration on either side of a biological membrane is the same. Second, it is free drug at the site of action that is pharmacologically active. While these two proposals may often be true, there are a numbers of situations where the free drug hypothesis may fail. The first premise may fail when drug permeability is low, when uptake transporters increase the concentration of a drug in a tissue, when efflux transporters decrease the concentration of drug in a tissue, or when low or disrupted blood flow may reduce the concentration of drug throughout a tissue. The second premise of the free drug hypothesis may fail when an irreversible inhibitor is involved, when a series of target-mediated events must occur in a series in order to elicit a pharmacological effect, or when in vitro conditions poorly represent the in vivo condition.

These concepts may be more specifically applied to DDIs. Low permeability or efflux of an NME out of the primary organ of clearance, such as the liver, may reduce the magnitude of a DDI. This is rarely observed, as leakiness often provides an alternate mechanism for tissue distribution of a drug. Conversely, uptake of an NME into a primary organ of clearance, such as the liver, may produce a DDI of unexpected magnitude. The concept of in vivo K_i, or $K_{i,iv}$, was developed to help substantiate these observations (Kunze and Trager 1996). $K_{i,iv}$ is a derived parameter shown in Eq. 35 below.

$$\frac{Cl_{\text{int}}}{Cl_{\text{int},I}} = 1 + \frac{[I]}{K_{i,iv}} \tag{35}$$

In an idealized situation, the ratio of in vitro K_i to $K_{i,iv}$ should equal one. Marked deviation from unity indicates the presence of a situation not accounted for in the in vitro experiment, such as mechanism(s) other than reversible inhibition, an active site environment in the in vivo situation that differs from the in vitro conditions, or active uptake or efflux that has altered the relative concentration of perpetrator drug at the enzyme active site. This type of phenomena has been observed clinically for drugs such as fluvoxamine, which has a high liver to plasma partition ratio. Using (S)-mephenytoin as a probe substrate for CYP2C19 activity, the unbound $K_{i,iv}$ was estimated to be 1.9 nM, compared to an unbound in vitro K_i of 76 nM, a nearly 40-fold increase in estimated inhibition potency in vivo (Yao et al. 2003). Similarly, the $K_{i,iv}$ for the fluvoxamine for its interaction with theophylline (CYP1A2) was estimated at 3.6 nM, while the in vitro K_i was

determined to be 36 nM. Conversely, the interaction of the reversible inhibitor fluconazole with the CYP2C9 substrate (S)-warfarin has an estimated $K_{i,iv}$ of 19.8 μM, similar to the measured in vitro K_i of 8 μM, indicating similarity between the in vitro and in vivo environments (Neal et al. 2003).

Is Concern over Plasma Protein Displacement DDIs Justified?

A common misconception is that plasma protein displacement is a major cause of drug interactions. Several early instances of DDI, initially believed to be caused by plasma protein displacement, were later proven to be based on metabolic DDIs. This is understandable based on theoretical considerations, as displacement of a drug from plasma proteins, if occurring, would be compensated for by an increase in intrinsic clearance as more free drug would be available for metabolism; the system would quickly revert back to equilibrium conditions. In a derivation by Benet and Hoener (2002), conversion of Eq. 17 into unbound drug concentration generates the following relationship:

$$AUC_{u,\,\mathrm{oral}} = {f_u}^{*} AUC_{\mathrm{oral}} = \frac{(Fa^{*}Fg^{*}\mathrm{Dose})}{CL_{\mathrm{int}}} \quad (36)$$

As can be seen, f_u is not involved in the final equation, and thus, changes in f_u are not expected to impact orally administered drugs cleared by the liver. Two unlikely instances where plasma protein displacement could conceivably play a role in DDIs are cases where a narrow therapeutic index drug is dosed intravenously and exhibits a high extraction ratio or where a narrow therapeutic index drug is dosed orally and exhibits a very fast PK/PD equilibrium (Benet and Hoener 2002).

Impact of Pharmacogenetics on DDIs

Pharmacogenetics is the study of how genetic factors impact drug response, where drug selection and dose may be altered depending upon the patient's genetic makeup. Genetic variation in genes expressing DMEs, called polymorphisms, may influence PK. Allelic variants of DMEs may be inactive or have gain or loss of function; the alteration in activity may also be substrate dependent. CYP2C9 is a well-known example, where CYP2C9*1 (wild type), CYP2C9*2, and CYP2C9*3 are the most common allelic variants. Expressed CYP2C9.1, CYP2C9.2, and CYP2C9.3 variants exhibit wild type, reduced, and markedly reduced activity in vitro, respectively (Rodrigues and Rushmore 2002). These CYP2C9 variants also exhibit differing inhibition potency when compared using the same inhibitors (Kumar et al. 2006b) The correlation of genotype to phenotype may be examined using a selective probe substrate of an enzyme. The four types of phenotypes commonly observed are extensive metabolizers (EMs) with typical activity, poor metabolizers (PMs) with no activity due to lack of enzyme expression or expression of inactive enzyme, intermediate metabolizers (IMs) with reduced activity due to one deficient allele, and ultra-rapid metabolizers (UMs) from gain of function variants, including copy number variants (Zanger and Schwab 2013). Examples of DMEs with clinically relevant genetic variation include CYP2C9 (phenytoin and S-warfarin), CYP2D6 (antidepressants and tamoxifen), CYP2C19 (proton pump inhibitors, S-mephenytoin, clopidogrel, and ticlopidine), UGT1A1 (bilirubin, SN-38), and NAT2 (isoniazid, procainamide, and hydralazine). As with DDI, both FDA (Guidance for Industry 2013) and EMA (guideline on key aspects for the use of pharmacogenomics in the pharmacovigilance of medicinal products) have regulatory guidance outlining the assessment of the impact of genotype and phenotype on PK and drug dosing. PK results with differing phenotypes can aid in establishing f_m for that drug using the following equation (Bohnert et al. 2016).

$$fm = 1 - \frac{AUC(EM)}{AUC(PM)} \quad (37)$$

Phenotype may have a marked role in DDIs in certain circumstances, particularly when the main route of clearance is due to a polymorphically expressed enzyme. If, for example, a NCE is primarily cleared by CYP2D6, secondary DMEs

may take on the primary role in clearance for CYP2D6 PMs. Co-administration of a drug that inhibits the secondary pathway would normally have minimal effect on the NCE clearance for a CYP2D6 EM but may exhibit a marked DDI effect for the CYP2D6 PM. In vitro characterization of routes of clearance and phenotype assessment of patient characteristics can be pivotal for informing and preventing these types of DDIs.

Conclusion

The theory and practice of understanding and providing quantitative estimates for metabolism-based DDIs has advanced dramatically. In vitro assay design and screening provide information on chemotypes that likely exhibit DDIs in the clinical situation and provide medicinal chemists with the basis to minimize DDI impact through the development of structure-activity relationships early in the discovery process. With the advent of technologies like PBPK, drug metabolism and clinical pharmacology scientists may quantitatively predict the clinical effects of DDIs based on high-content in vitro experiments, even for complex situations where inhibition, inactivation, and induction are all expected simultaneously or where multiple drug entities such as drug metabolites are present. Current frontiers in the field include the prediction of DDIs based on drug metabolism-transporter interplay and the prediction of DDIs for special populations including pediatrics and specific disease states. Understanding of the DDI potential for a clinical drug candidate is gained through the careful generation, examination, and integration of in vitro, preclinical, and clinical data to form a cohesive picture of the absorption, distribution, metabolism, and excretion (ADME) characteristics for that drug.

References and Further Reading

Benet LZ, Hoener BA (2002) Changes in plasma protein binding have little clinical relevance. Clin Pharmacol Ther 71:115–121

Bjornsson TD, Callaghan JT, Einolf HJ, Fisher V, Gan L, Grimm S, Kao J, King SP, Miwa G, Ni L, Kumar G, McLeod J, Obach SR, Roberts S, Roe A, Shah A, Snikeris F, Sullivan JT, Tweedie D, Vega JM, Walsh J, Wrighton SA (2003) The conduct of in vitro and in vivo drug-drug interaction studies: a PhRMA perspective. J Clin Pharmacol 43:443–469

Bohnert T, Patel A, Templeton I, Chen Y, Lu C, Lai G, Leung L, Tse S, Einolf H, Wang Y-H, Sinz M, Stearns R, Walsky R, Geng W, Sudsakorn S, Moore D, He L, Wahlstrom J, Keims J, Narayanan R, Lang D, Yang Q (2016) Evaluation of a new molecular entity as a victim of metabolic drug-drug interactions – an industry perspective. Drug Metab Dispos 44:1399–1423

Chu V, Einolf HJ, Evers R, Kumar G, Moore D, Ripp S, Silva J, Sinha V, Sinz M, Skerjanec A (2009) In vitro and in vivo induction of cytochrome P450: a survey of the current practices and recommendations: a pharmaceutical research and manufacturers of America perspective. Drug Metab Dispos 37:1339–1354

Chung J, Alvarez-Nunez F, Chow V, Daurio D, David J, Dodds M, Emery M, Litweiler K, Paccaly A, Peng J, Rock B, Wienkers L, Yang C, Yu Z, Wahlstrom J (2015) Utilizing physiologically based pharmacokinetic modeling to inform formulation and clinical development for a compound with pH-dependent solubility. J Pharm Sci 104:1522–1532

Day RO, Snowden L, McLachlan AJ (2017) Life-threatening drug interactions: what the physician needs to know. Intern Med J 47:501–512

Dickmann LJ, Patel SK, Rock DA, Wienkers LC, Slatter JG (2011) Effects of interleukin-6 (IL-6) and an anti-IL-6 monoclonal antibody on drug-metabolizing enzymes in human hepatocyte culture. Drug Metab Dispos 39:1415–1422

European Medicine Agency (EMA), Committee for Human Medical Products (CHMP) (2012) Guideline on the investigation of drug interactions. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf

European Medicine Agency (EMA), Committee for Human Medical Products (CHMP) (2015) Guideline on key aspects for the use of pharmacogenomics in the pharmacovigilance of medicinal products. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/11/WC500196800.pdf

Evers R, Dallas S, Dickmann LJ, Fahmi OA, Kenney JR, Kraynov E, Nguyen T, Patel AH, Slatter JG, Zhang L (2013) Critical review of preclinical approaches to investigate cytochrome P450-mediated therapeutic protein drug-drug interactions and recommendations for best practices: a white paper. Drug Metab Dispos 41:1593–1609

FDA Guidance for Industry (2013) Clinical pharmacogenomics: premarket evaluation in early-phase clinical studies and recommendations for labeling. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf

FDA Guidance for Industry (2017a) Clinical drug interaction studies – study design, data analysis, and clinical implications. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292362.pdf

FDA Guidance for Industry (2017b) In vitro metabolism and transporter-mediated drug-drug interaction studies. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm581965.pdf

Foti RS, Rock DA, Wienkers LC, Wahlstrom JL (2010) Selection of alternative CYP3A4 probe substrates for clinical drug interaction studies using in vitro data and in vivo simulation. Drug Metab Dispos 38:981–987

Foti RS, Rock DA, Pearson JT, Wahlstrom JL, Wienkers LC (2011) Mechanism-based inactivation of cytochrome P450 3A4 by mibefradil through heme destruction. Drug Metab Dispos 39:1188–1195

Frere J-M, Leyh B, Renard A (1983) Lineweaver-Burk, Hanes, Eadie-Hofstee and Dixon plots in non-steady-state situations. J Theor Biol 101:387–400

Grimm SW, Einolf HJ, Hall SD, He K, Lim HK, Ling KH, Lu C, Nomeir AA, Seibert E, Skordos KW, Tonn GR, VanHorn R, Wang RW, Wong YN, Yang TJ (2009) The conduct of in vitro studies to address time-dependent inhibition of drug-metabolizing enzymes: a perspective of the pharmaceutical research and manufacturers of America. Drug Metab Dispos 37:1355–1370

Hutzler JM, Tracy TS (2002) Atypical kinetic profiles in drug metabolism reactions. Drug Metab Dispos 30:355–362

Ito K, Hallifax D, Obach RS, Houston JB (2005) Impact of parallel pathways of drug elimination and multiple cytochrome P450 involvement on drug-drug interactions: CYP2D6 paradigm. Drug Metab Dispos 33:837–844

Iwatsubo T, Hirota N, Ooie T, Suzuki H, Shimada N, Chiba K, Ishizaki T, Green CE, Tyson CA, Sugiyama Y (1997) Prediction of in vivo drug metabolism in the human liver from in vitro metabolism data. Pharmacol Ther 73:147–171

Jacubeit T, Drisch D, Weber E (1990) Risk factors as reflected by an intensive drug monitoring system. Agents Actions 29:117–125

Jamei M, Marciniak S, Feng K, Barnett A, Tucker G, Rostami-Hodjegan A (2009) The Simcyp population-based ADME simulator. Expert Opin Drug Metab Toxicol 5:211–223

Korzekwa KR, Krishnamachary N, Shou M, Ogai A, Parise RA, Rettie AE, Gonzales FJ, Tracy TS (1998) Evaluation of atypical cytochrome P450 kinetics with two-substrate models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active site. Biochemistry 37:4137–4147

Kumar V, Locuson CW, Sham YY, Tracy TS (2006a) Amiodarone analog-dependent effects on CYP2C9-mediated metabolism and kinetic profiles. Drug Metab Dispos 34:1688–1696

Kumar V, Wahlstrom JL, Rock DA, Warren CJ, Gorman LA, Tracy TS (2006b) CYP2C9 inhibition: impact of probe substrate selection and pharmacogenetics on in vitro inhibition profiles. Drug Metab Dispos 34:1966–1975

Kunze KL, Trager WF (1993) Isoform-selective mechanism-based inhibition of cytochrome P450 1A2 by furafylline. Chem Res Toxicol 6:649–656

Kunze KL, Trager WF (1996) A rational approach to management of a metabolically based drug interaction. Drug Metab Dispos 24:429–435

Levy G (1998) What are narrow therapeutic index drugs? Clin Pharmacol Ther 63:501–505

Lin JH (2006) CYP induction-mediated drug interactions: in vitro assessment and clinical implications. Pharm Res 23:1089–1116

Lin JH, Lu AYH (1997) Role of pharmacokinetics and metabolism in drug discovery and development. Pharmacol Rev 49:403–449

Lin Y, Lu P, Tang C, Mei Q, Sandig G, Rodrigues AD, Rushmore TH, Shou M (2001) Substrate inhibition kinetics for cytochrome P450-catalyzed reactions. Drug Metab Dispos 29:368–374

Lineweaver H, Burk D (1934) The determination of enzyme dissociation constants. J Am Chem Soc 56:658–666

Michaelis L, Menten ML (1913) Die kinetic der invertinwirkung. Biochem Z 49:333–369

Neal JM, Kunze KL, Levy RH, O'Reilly RA, Trager WF (2003) Kiiv, an in vivo parameter for predicting the magnitude of a drug interaction arising from competitive enzyme inhibition. Drug Metab Dispos 31:1043–1048

Rodrigues AD, Rushmore TH (2002) Cytochrome P450 pharmacogenetics in drug development: in vitro studies and clinical consequences. Curr Drug Metab 3:289–309

Rowland M, Martin SB (1973) Kinetics of drug-drug interactions. J Pharmacokinet Biopharm 1:553–567

Silverman RB (1995) Mechanism-based inactivators. Methods Enzymol 249:240–283

Smith DA, Di L, Kers EH (2010) The effect of plasma protein binding on in vivo efficacy: misconceptions in drug discovery. Nat Rev Drug Discov 9:929–929

VandenBrink BM, Foti RS, Rock DA, Wienkers LC, Wahlstrom JL (2012) Prediction of CYP2D6 drug interactions from in vitro data: evidence for substrate-dependent inhibition. Drug Metab Dispos 40:47–53

Wahlstrom JL, Rock DA, Slatter JG, Wienkers LC (2006) Advances in predicting CYP-mediated drug interactions in the drug discovery setting. Expert Opin Drug Discov 1:677–691

Wienkers LC, Heath TG (2005) Predicting in vivo drug interactions from in vitro drug discovery data. Nat Rev Drug Discov 4:825–833

Yao C, Kunze KL, Kharasch ED, Wang Y, Trager WF, Ragueneau I, Levy RH (2001) Fluvoxamine-theophylline interaction: gap between in vitro and in vivo inhibition constants toward cytochrome P4501A2. Clin Pharmacol Ther 70:415–424

Yao C, Kunze KL, Trager WF, Kharasch ED, Levy RH (2003) Comparison of in vitro and in vivo inhibition potencies of fluvoxamine toward CYP2C19. Drug Metab Dispos 31:565–571

Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol Ther 138:103–141

Specific Studies for Formulation Development

47

Roland Wesch

Contents

Abstract

The bioavailability of a drug depends on the properties of the drug product, a combination of drug and formulation properties. The support for formulation development by means of clinical PK studies is multifaceted and, in fact, covers all routes of administration, intravascular routes as well as extravascular ones, like oral, intramuscular, or subcutaneous routes, and – in most instances – vaginal, dermal, ocular, topic, rectal, nasal, or pulmonary administration. The drugability (disease-tailored exposure profiles mediated by optimized delivery systems) of pharmacologically active substances will remain one of the major challenges in drug development, especially if poorly absorbable, poorly soluble compounds are considered. Rare exceptions include some drugs belonging to BCS class I (highly soluble, highly permeable) that rapidly dissolve from Immediate Release solid oral drug products.

R. Wesch (✉)
R&D Metabolism and PK Germany, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany
e-mail: roland.wesch@sanofi.com; Roland.Wesch@sanofi-aventis.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_15

Purpose and Rationale

The bioavailability of a drug depends on the properties of the drug product, a combination of drug and formulation properties. The support for formulation development by means of clinical PK studies is multifaceted and, in fact, covers all routes of administration, intravascular routes as well as extravascular ones, like oral, intramuscular, or subcutaneous routes, and – in most instances – vaginal, dermal, ocular, topic, rectal, nasal, or pulmonary administration. The drugability (disease-tailored exposure profiles mediated by optimized delivery systems) of pharmacologically active substances will remain one of the major challenges in drug development, especially if poorly absorbable, poorly soluble compounds are considered. Rare exceptions include some drugs belonging to BCS class I (highly soluble, highly permeable) that rapidly dissolve from Immediate Release solid oral drug products.

Formulation development requires close coordination of various functions, for example, galenics, analytics, process development, preclinical, and clinical pharmacology. Clinical PK is just one, however, an important component in this framework.

Details about how to deal with changes in components or composition of drug products are described in published regulatory guidances (Guidance for Industry Waiver of in vivo bioavailability and bioequivalence studies for immediate release solid oral dosage forms containing certain active moieties/active ingredients based on the biopharmaceutics classification system (BCS). U.S. Department of Health and Human Services et al. 1999; SUPAC-IR 1995; SUPAC-MR 1997), while no formal guidance exists that in particular covers all types of formulation interactions. Details about clinically relevant (drug–drug) interactions and assessment of equivalence of formulations can be found in Guidance for Industry Statistical Approaches to Establishing Bioequivalence (2001), Steinijans and Hauschke (1997), CPMP (2002), and CPMP (1998).

In order to optimize drug exposure, the development of modified release formulations is an option. In vitro tests (stability testing, dissolution) are routinely used as a first step in formulation development in order to build an absorption model for the prediction of in vivo exposure via the in vitro/in vivo correlation. The assessment of efficacy in disease models and exposure (PK) in animals will be applied to lead candidates only, before going to man.

Reasons for change of formulation include, but are not limited to, poor bioavailability of solid oral formulations, limitations in drug load for oral or parenteral formulations, profound food effect, too early/too late onset of action (absorption, distribution), too short/too long duration of action (metabolism, elimination), or high intra- and inter-individual variability.

Procedure

The design of an exploratory formulation development study with SAR001 is presented below in Part A.

In the project with SAR001, the formulation development study explored the relative bioavailability of three prototype nanocrystal (NC) formulations (tablet, granules, lyophilisate) versus a soft gelatine capsule (SGC) formulation that was used in early clinical phases. This study should help to determine selection and development of alternative formulations to be used in Phase III studies, as the current SGC had *limitations in the unit strength* possibly not suitable for long-term efficacy trials. The food effect was also investigated, as in an animal pilot study, the *magnitude of food effect* was more important for NC dispersion when compared to SGC.

In another project (HMR456) modified release (MR) formulations were developed in order to overcome the *short elimination half-life,* and thus, a *short duration of action* of the Immediate Release formulation. The study design and the main PK results are presented elsewhere in this textbook (▶ Chap. 38, "Effects of Food Intake"). In Part B, we will discuss the value of the deconvolution tool, which was applied to the PK results of the study.

Part A

Protocol Outline

Relative bioavailability of three prototype nanocrystal formulations of SAR001 in comparison to a soft gelatin capsule formulation of SAR001 under fasting and fed conditions to healthy subjects.

Primary Objective

To assess the relative bioavailability of three prototype nanocrystal formulations of SAR001, in comparison to the soft gelatin capsule formulation, by assessing plasma concentration of SAR001 under fasting and fed conditions.

Study Design

Open, randomized, two-group, four-treatment, four-period, four-sequence crossover study using two parallel groups under fasting or fed conditions (Groups I and II, respectively). All single oral drug administration periods within groups were separated by a washout of 7 days.

Inclusion Criteria

Healthy male subjects, aged between 18 and 45 years, with a Body Mass Index (BMI) between 18 and 28 kg/m^2 inclusive, and liver function tests and creatine kinase values within reference ranges.

Treatments

- Treatment A: Single dose of 80 mg SAR001 in soft gelatin capsule (SGC).
- Treatment B: Single dose of 80 mg SAR001 in uncoated tablet.
- Treatment C: Single dose of 80 mg SAR001 in granules for oral suspension.
- Treatment D: Single dose of 80 mg SAR001 as lyophilisate for oral suspension.

All treatments were administered under two food conditions: fasting (overnight fasting + 4 h postdose, Group I) and fed (high fat-high calorie breakfast starting 30 min and ending 5 min predose, Group II).

Pharmacokinetic Data

Concentration of SAR001 in plasma before and at predefined times after dosing.

Evaluation

Only part of the evaluation will be presented here.

Standard descriptive statistics were calculated for each parameter and each treatment.

To determine the relative bioavailability of any pair of formulations, 90% confidence intervals (CI) of formulations ratio for $AUC_{0–168}$ and C_{max} were displayed under each food regimen.

For C_{max}, $AUC_{0–72}$, and $AUC_{0–168}$, formulation effect was assessed using a linear mixed effects model separately for each food regimen on log-transformed parameters. Estimates with 90% confidence intervals of pairwise formulations ratio of geometric means were computed within the linear mixed effects model framework.

Similarly, food effect was assessed for each formulation using a linear mixed effects model on log-transformed parameters. Estimates and 90% confidence intervals of food regimens ratio of geometric means were computed within the linear mixed effects model framework.

To determine the food effect on each of the four formulations, 90% CI of fed/fasted ratio for $AUC_{0–168}$ and C_{max} were displayed for each formulation separately.

Critical Assessment of the Method

Due to the exploratory pilot character of such a formulation development study, a sample size calculation in its proper sense is not routinely performed. Often, a number of 12 study completers per cohort is considered sufficient for this purpose. This quite small cohort size "encourages" to implement high complexity with multiple objectives within a single trial. If solid formulations are to be compared, the same unit strength for each formulation is recommended in order to avoid an intrinsic source of variability, for example, if two tablets with 40 mg each are compared

to a single capsule with 80 mg drug load. The number of dose units should be kept low, again to avoid a source of variability. Standardization is questionable if up to ten units or even more have to be swallowed with a limited amount of non-carbonated water (e.g., 240 mL).

Preceding studies should help to define the necessary washout period in order to avoid carry-over effects. If the predose concentration is higher than 5% of C_{max} in a given study period, the predose value should be subtracted from all post-dose concentrations as corrective action.

Concerning food, not only composition (constituents, calories) but also start and end of food intake in relation to drug administration should be standardized.

In this example, the study subjects were stratified for food condition, which results in four high fat-high calorie meals for members of Group II, and in four 14 h lasting fasting periods for members of Group I. In order to avoid late dropouts in crossover studies that apply within comparisons of both food conditions, it may be advisable to define the compliance to high fat-high calorie breakfast as an inclusion criterion for the study.

In all studies with parallel groups, care should be taken that baseline demographics – for example, gender ratios, body weight, BMI, age – are similar between groups. A possibility to overcome this request is the implementation of the same reference treatment for each group. An example for this approach can be found in the food effect chapter.

Modifications of the Method

If in late clinical development the need for a formulation switch becomes evident, and a formulation comparison is necessary for bridging purposes, the exploratory character of a formulation development study gets lost and a bioequivalence study has to be conducted.

Part A

To illustrate the type of data that can be obtained using the discussed study, a high-level summary of the pharmacokinetic results obtained from the study described above under "Procedure" is presented below (Tables 1 and 2; Figs. 1 and 2).

Results: Pharmacokinetics

In fasting or in fed condition, whatever the formulation, t_{max} was about the same, with median values ranging from 2 to 4 h post administration.

In fasting condition, SAR001 exposure was lower with nanocrystal formulations compared to soft gelatin capsules (SGC). In comparison to SGC, variability of pharmacokinetic parameters was similar for the lyophilisate, slightly higher for the tablets and higher for the granules.

In fed condition, SAR001 exposure was close to that observed with SGC for granules and lower than that observed with SGC for tablets. Variability of pharmacokinetic parameters was about the same, whatever the formulation.

Food effect (relative bioavailability) based on $AUC_{0–72}$ was about twice as high with nanocrystal formulations compared to soft gelatin capsules.

Part B

Procedure

The immediate release PK properties of the active metabolite HMR123 of drug HMR456 were not sufficient to support a twice-a-day dosing (terminal elimination half-life too short, time above PD concentration threshold too short). Therefore

Table 1 Pharmacokinetic population: numbers by treatment/condition

Condition/ treatment	Soft gelatin capsules	Granules for oral suspension	Lyophilisate for oral suspension	Uncoated tablets
Fasting	13	13	13	13
Fed	13	14	13	12

Table 2 Relative bioavailability estimates and 90% confidence intervals between food conditions

Parameter	Food comparison	Ratio estimate and 90% CI
C_{max} (ng/mL)	Fed vs. fasted for capsule	2.03 (1.39, 2.96)
	Fed vs. fasted for granule	4.35 (2.99, 6.32)
	Fed vs. fasted for lyophilisate	5.04 (3.45, 7.35)
	Fed vs. fasted for tablet	5.77 (3.94, 8.47)
$AUC_{0–72}$ (ng·h/mL)	Fed vs. fasted for capsule	2.00 (1.44, 2.76)
	Fed vs. fasted for granule	3.93 (2.85, 5.43)
	Fed vs. fasted for lyophilisate	4.31 (3.11, 5.96)
	Fed vs. fasted for tablet	4.59 (3.31, 6.36)
$AUC_{0–168}$ (ng·h/mL)	Fed vs. fasted for capsule	2.07 (1.54, 2.77)
	Fed vs. fasted for granule	3.85 (2.84, 5.22)
	Fed vs. fasted for lyophilisate	3.69 (2.72, 4.99)
	Fed vs. fasted for tablet	3.86 (2.81, 5.31)

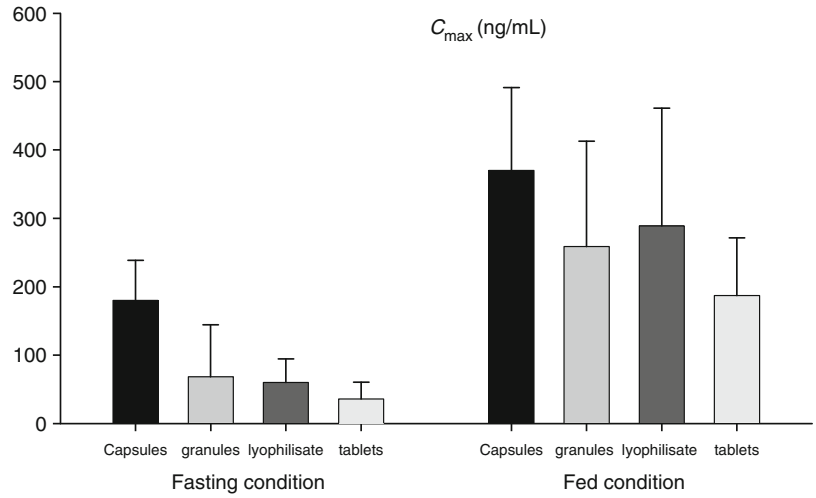


Fig. 1 Mean (SD) SAR001 C_{max} obtained after single oral administration to healthy young male subjects using four different formulations, in fasting or fed condition

modified release (MR) formulations were developed. The use of a deconvolution tool will be discussed in this section.

The design of the exploratory bioavailability study on modified release drug products is presented in brief below. Detailed information is given in the food effect chapter. For the assumptions of the hypothetical deconvolution tool, in vitro/in vivo dissolution data from a predecessor study with the same compound was used.

Protocol Outline

Comparison of pharmacokinetics and safety of Modified Release formulations of 600 mg HMR456 with that of an immediate release formulation.

Primary Objective

To compare the PK characteristics of modified release (MR) formulations of HMR456 with the PK of an immediate release (IR) formulation of HMR456.

Secondary Objective

To assess the influence of food on the PK of MR formulations of HMR456.

Hypothetical in vivo dissolution was performed in addition to the objectives mentioned in the study protocol. Application of the method, results, and interpretation will be discussed here.

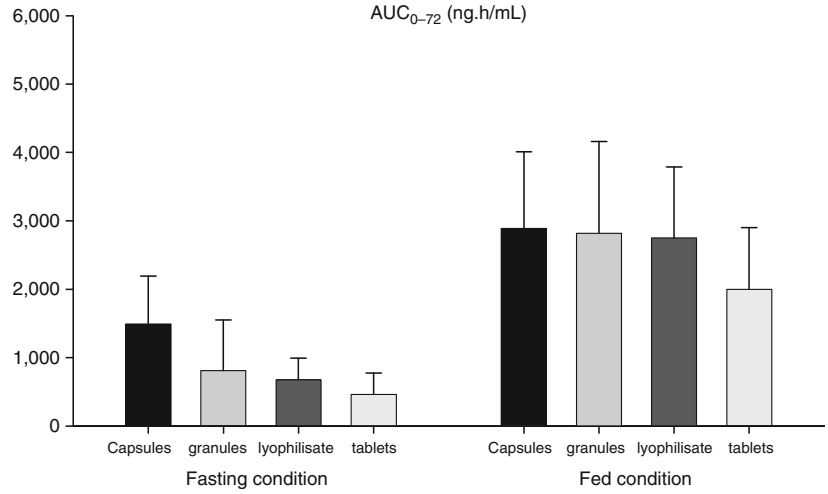


Fig. 2 Mean (SD) SAR001 AUC_{0-72} obtained after single oral administration to healthy young male subjects using four different formulations, in fasting or fed condition

Study Design

Single center, open-label, single-dose, four-period crossover study design with two parallel treatment groups.

Single oral doses of 600 mg HMR456 were given under fasting and under non-fasting conditions.

Inclusion Criteria

Healthy men aged 18–55 years.

Treatments

Treatment Group I

- Treatment A: 600 mg HMR456 (one film-coated tablet containing 200 mg + one film-coated tablet containing 400 mg given together) as IR formulation under non-fasting (NF) conditions (reference).
- Treatment B: 600 mg HMR456 in MR formulation (matrix tablet 1) under fasting (F) and NF conditions.
- Treatment C: 600 mg HMR456 in MR formulation (bilayer tablet 1) under F and NF conditions.

Treatment Group II

- Treatment A: 600 mg HMR456 (one film-coated tablet containing 200 mg + one film-coated tablet containing 400 mg given together) as IR formulation under non-fasting (NF) conditions (reference).
- Treatment D: 600 mg HMR456 in MR formulation (matrix tablet 2) under F and NF conditions.
- Treatment E: 600 mg HMR456 in MR formulation (bilayer tablet 2) under F and NF conditions.

MR tablet formulation 1 contains hydroxypropyl methyl cellulose, MR tablet formulation 2 contains carrageenan.

Pharmacokinetic Data

Concentration of HMR123 in plasma before and at predefined times after dosing.

Evaluation

Bioanalytical data: Individual plasma concentrations of HMR123 were tabulated together with standard descriptive statistics for each treatment. Individual and median profiles were presented graphically.

In vivo *dissolution data*: The individual hypothetical in vivo dissolutions for the four MR formulations administered under fasting and non-fasting conditions were estimated by

Table 3 Hypothetical dissolution data for HMR123 obtained by deconvolution using Treatment A(NF) as impulse function. Median, range

Measures	B(NF)	B(F)	C(NF)	C(F)	D(NF)	D(F)	E(NF)	E(F)
Maximum amount absorbed (mg)	358.97	290.50	485.97	561.49	582.56	460.12	629.25	652.95
	239.61–727.06	173.93–484.88	287.67–994.55	300.14–1,817.29	465.29–1,076.11	339.60–795.42	430.72–5,674.72	452.19–3,183.05
Maximum amount absorbed[a] (% of dose)	63.22	51.16	85.58	98.88	102.59	81.03	110.82	114.99
	42.20–128.04	30.63–85.39	50.66–175.15	52.86–320.04	81.94–189.51	59.81–140.08	75.85–999.35	79.63–560.55
Time to reach maximum amount (h)	15.00	24.00	15.00	15.00	9.00	15.00	2.50	15.00
	4.00–24.00	15.00–24.00	6.00–15.03	0.50–36.00	2.00–24.00	10.00–24.00	0.50–24.00	0.50–15.00
Time to reach 20% of maximum amount (h)	1.81	0.55	0.57	0.27	0.76	0.34	0.51	0.26
	0.43–2.64	0.33–1.58	0.26–1.25	0.05–1.64	0.28–1.92	0.16–0.81	0.10–1.54	0.12–0.31
Time to reach 40% of maximum amount (h)	2.57	2.14	0.83	0.35	0.91	0.62	0.69	0.33
	0.97–3.38	0.54–5.47	0.33–1.73	0.11–1.97	0.36–2.62	0.31–1.16	0.20–1.83	0.24–0.81
Time to reach 50% of maximum amount (h)	2.80	3.34	0.91	0.40	1.04	0.69	0.81	0.37
	1.35–4.10	0.58–7.45	0.36–2.21	0.14–2.07	0.39–2.68	0.34–1.59	0.26–1.98	0.29–1.02
Time to reach 60% of maximum amount (h)	3.57	4.54	1.06	0.44	1.54	0.94	1.04	0.41
	1.51–6.00	0.62–9.33	0.40–3.19	0.16–2.25	0.43–2.75	0.37–2.22	0.30–2.13	0.33–1.22
Time to reach 80% of maximum amount (h)	5.27	6.95	2.95	1.64	1.79	2.54	1.44	0.48
	1.80–8.61	0.71–15.28	0.46–6.17	0.22–6.18	0.53–2.87	0.44–7.96	0.40–2.64	0.42–4.49

B(F/NF): 600 mg HMR456 in ER formulation (HPMC matrix tablet) under fasting (F) and non-fasting (NF) conditions, respectively
C(F/NF): 600 mg HMR456 in ER formulation (HPMC, bilayer tablet) under fasting (F) and non-fasting (NF) conditions, respectively
D(F/NF): 600 mg HMR456 in ER formulation (carrageenan matrix tablet) under fasting (F) and non-fasting (NF) conditions, respectively
E(F/NF): 600 mg HMR456 in ER formulation (carrageenan, bilayer tablet) under fasting (F) and non-fasting (NF) conditions, respectively
[a]Dose calculated for 567.84 mg HMR123 (corresponding to 600 mg HMR456)

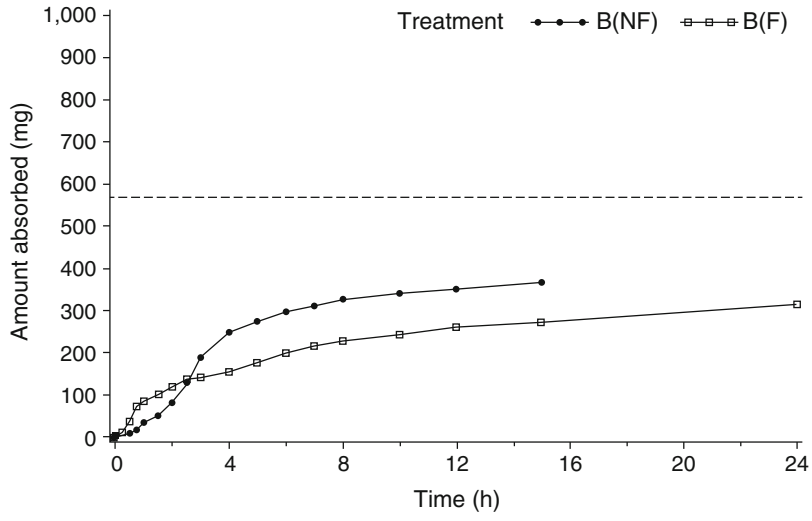


Fig. 3 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments B (F) and B(NF) (mg)

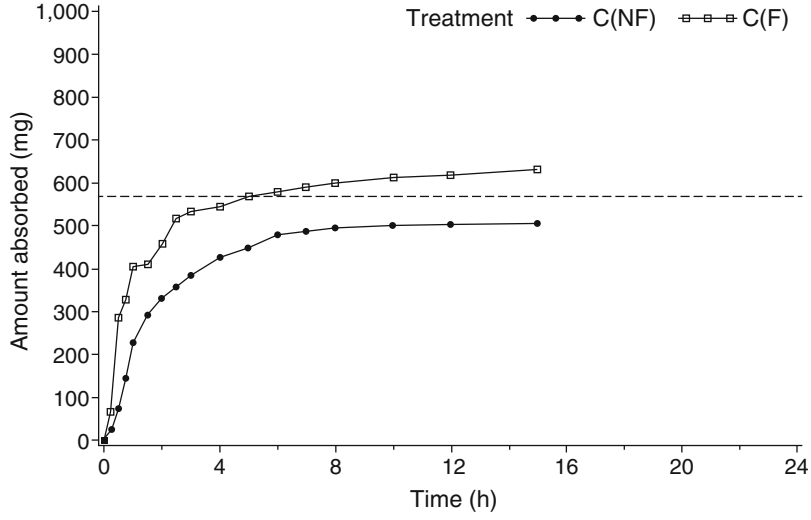


Fig. 4 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments C (F) and C(NF) (mg)

numerical deconvolution using the individual response to the IR formulation given under non-fasting conditions as the weighting (impulse) function using a hidden function of the validated HOEREP-PC software.

Plateau time data: The additional pharmacokinetic characteristics, i.e., plateau times (h) of HMR123 (time above 200, 500, 800, and 1,000 ng/mL) were calculated in the interval from administration ($t = 0$) to exactly 12 h thereafter from the plasma concentration–time data pairs and subjected to ANOVA. Points of intersection with a specific plateau concentration were obtained by linear interpolation.

PK data are presented elsewhere.

Critical Assessment of the Method

The study described here has a very complex design for its exploratory approach. It combines four different MR formulations, each tested under fasting and non-fasting conditions, and compares the results to the IR drug product as the reference formulation in two separate study groups. The

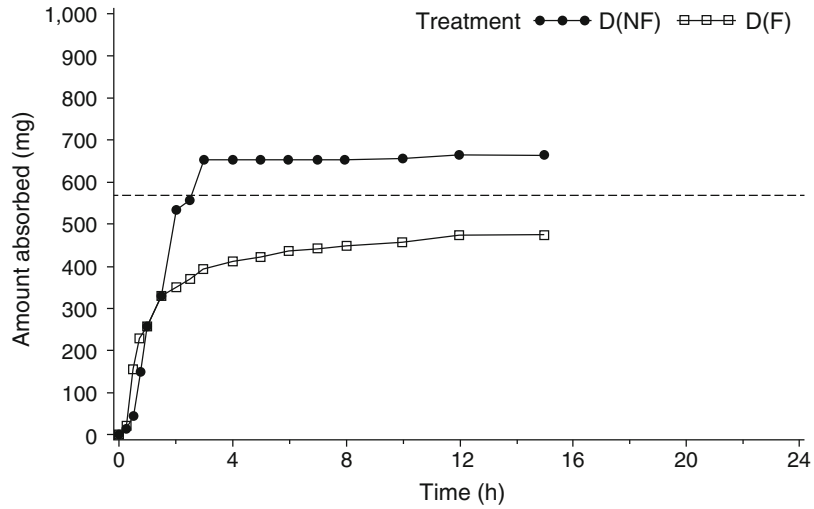


Fig. 5 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments D (F) and D(NF) (mg)

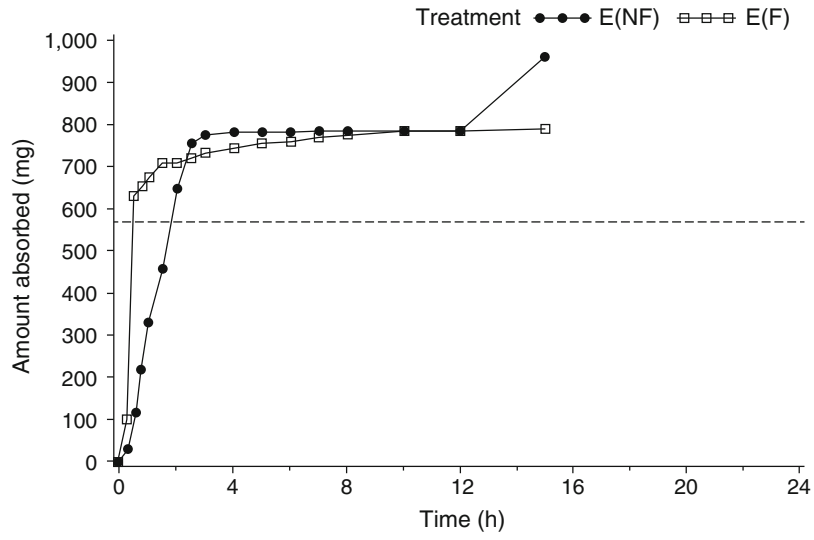


Fig. 6 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments E (F) and E(NF) (mg)

bilayer tablets combine an IR component and an MR component in one vehicle. In this project, a close cooperation between the galenics department, analytical science department, and the clinical pharmacokinetic function (including study management, bioanalysis and PK evaluation) was mandatory. The in vitro/in vivo correlation was done by means of the deconvolution which is an appropriate surrogate to describe the in vivo dissolution.

The mismatch of surpassing 100% absorption of the active metabolite, that we observed in our study, is probably due to method constraints in combination with the immediate release data, as the deconvolution method requires data from a formulation with zero-order absorption for the impulse function, for example, an oral solution (oral bolus input); the immediate release formulation only provides an approximation to the required properties.

Modifications of the Method

The application of in vitro/in vivo correlation (IVIVC) and the tools for obtaining IVIVC

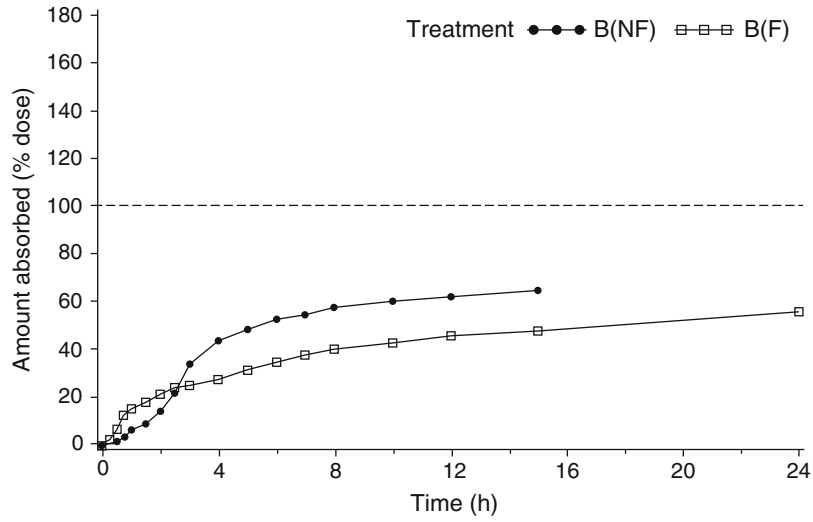


Fig. 7 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments B (F) and B(NF) (%dose)

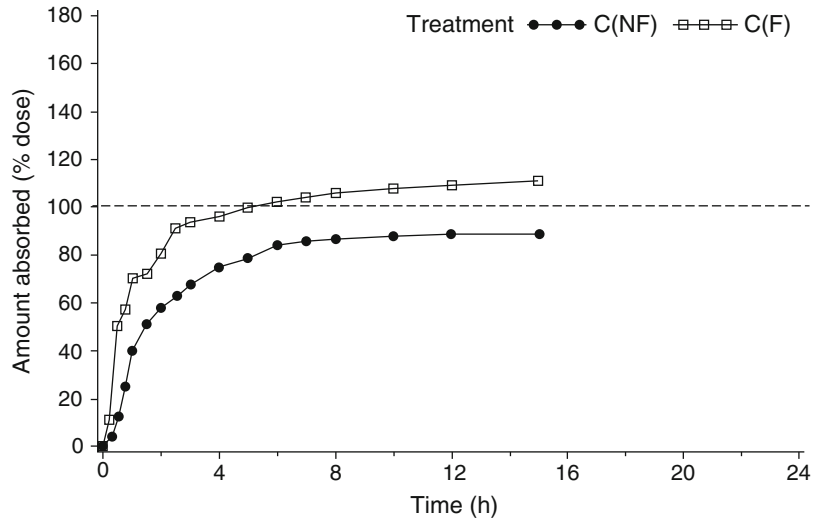


Fig. 8 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments C (F) and C(NF) (%dose)

including deconvolution are reviewed in FDA Guidance for Industry (1997).

References and Further Reading

FDA Guidance for Industry: Extended release oral dosage forms: development, evaluation, and application of in vitro/in vivo correlations. September 1997

Example

To illustrate the amount of data that can be obtained using the deconvolution tool obtained from the study described above under "Procedure" is presented below.

Results – Hypothetical In Vivo Dissolution

Deconvolution is used to evaluate in vivo drug release and drug absorption from orally administered drug formulations (i.e., extended release) when data from a known drug input are available. The applied deconvolution method requires data from a formulation with zero-order absorption as known input, for example, an oral solution (oral bolus input); the immediate release formulation used as

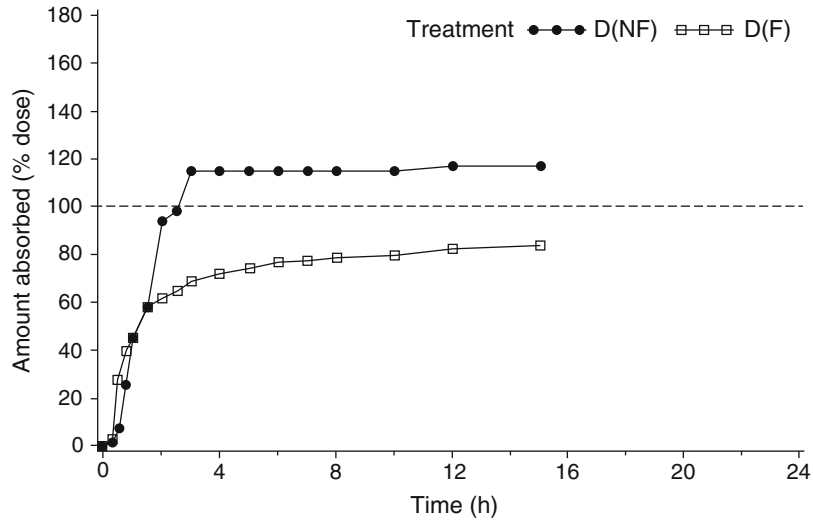


Fig. 9 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments D (F) and D(NF) (%dose)

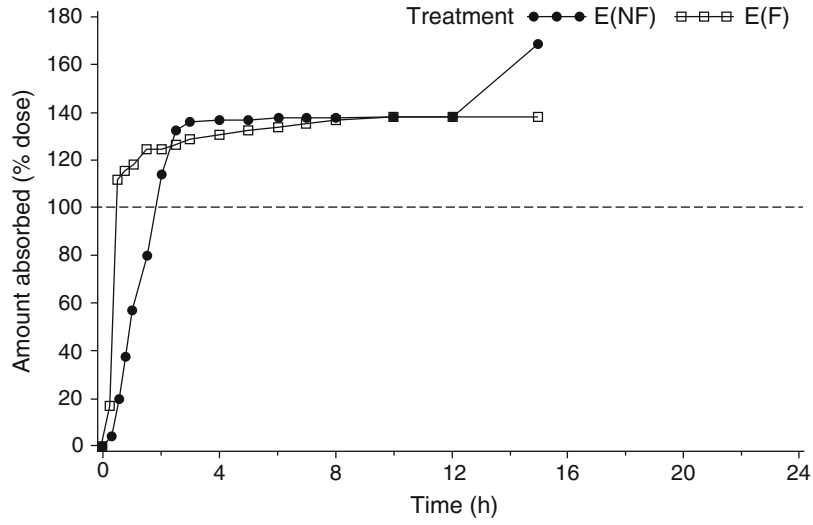


Fig. 10 Geometric means of the hypothetical in vivo dissolution profiles of HMR123. Treatments E (F) and E(NF) (%dose)

known input only provides an approximation to the required properties.

The medians and ranges of the hypothetical dissolution data for the active metabolite HMR123 obtained by deconvolution are listed in the following Table 3.

The following Figs. 3, 4, 5, 6, 7, 8, 9, and 10 show the hypothetical geometric mean in vivo dissolution profiles for the metabolite HMR123 (absolute amount absorbed vs. time as well as percentage of theoretical dose of the metabolite vs. time) (Table 4).

As can be seen in the above figures, as well as in Table 3, Treatments C and E (the bilayer tablets that contain the IR component) had a steeper amount absorbed profile as compared to the parallel matrix tablets (Treatments B and D). For example, the time for 50% of the maximal absorption looks much shorter (especially when Treatment C is compared to B). This effect was more pronounced under fasting conditions. Only with Treatment E (carrageenan bilayer tablets), the hypothetical in vivo dissolution profiles surpassed the 100% absorption, both under fasting and non-fasting conditions. For Treatment C, this occurred only under fasting conditions and for Treatment D only under non-fasting conditions.

Table 4 Relative bioavailability estimates and 90% confidence intervals between formulations

Food condition	Parameter	Formulation comparison	Ratio estimate and 90% CI
Fasted	C_{max} (ng/mL)	Granule vs. capsule	0.30 (0.20, 0.45)
		Lyophilisate vs. capsule	0.29 (0.19, 0.43)
		Tablet vs. capsule	0.17 (0.11, 0.25)
	$AUC_{0–72}$ (ng·h/mL)	Granule vs. capsule	0.47 (0.34, 0.64)
		Lyophilisate vs. capsule	0.44 (0.32, 0.60)
		Tablet vs. capsule	0.29 (0.21, 0.39)
	$AUC_{0–168}$ (ng·h/mL)	Granule vs. capsule	0.50 (0.39, 0.63)
		Lyophilisate vs. capsule	0.52 (0.41, 0.67)
		Tablet vs. capsule	0.37 (0.29, 0.49)
Fed	C_{max} (ng/mL)	Granule vs. capsule	0.67 (0.51, 0.88)
		Lyophilisate vs. capsule	0.72 (0.54, 0.95)
		Tablet vs. capsule	0.47 (0.35, 0.62)
	$AUC_{0–72}$ (ng·h/mL)	Granule vs. capsule	0.95 (0.80, 1.12)
		Lyophilisate vs. capsule	0.94 (0.79, 1.11)
		Tablet vs. capsule	0.65 (0.54, 0.77)
	$AUC_{0–168}$ (ng·h/mL)	Granule vs. capsule	0.92 (0.79, 1.07)
		Lyophilisate vs. capsule	0.94 (0.81, 1.08)
		Tablet vs. capsule	0.70 (0.60, 0.81)

References and Further Reading

CPMP/EWP/560/95 Note for Guidance on the Investigation of Drug Interactions (CPMP, June 1998)

CPMP/EWP/QWP/1401/98 Note for Guidance on the Investigation of Bioavailability and Bioequivalence (CPMP, January 2002)

FDA Guidance for Industry: Extended release oral dosage forms: development, evaluation, and application of in vitro/in vivo correlations. September 1997

Guidance for Industry Scale-up and post-approval changes (SUPAC-IR): Chemistry, manufacturing, and control; in vitro dissolution testing and in vivo bioequivalence documentation. U.S. Department of Health and Human Services, Food and Drug Administration; Center for Drug Evaluation and Research (CDER). September 1995

Guidance for Industry Scale-up and post-approval changes (SUPAC-MR): Chemistry, manufacturing, and control; in vitro dissolution testing and in vivo bioequivalence documentation. U.S. Department of Health and Human Services, Food and Drug Administration; Center for Drug Evaluation and Research (CDER). September 1997

Guidance for Industry Statistical Approaches to Establishing Bioequivalence. U.S. Department of Health and Human Services, Food and Drug Administration; Center for Drug Evaluation and Research (CDER). January 2001

Guidance for Industry Waiver of in vivo bioavailability and bioequivalence studies for immediate release solid oral dosage forms containing certain active moieties/active ingredients based on the biopharmaceutics classification system (BCS). U.S. Department of Health and Human Services, Food and Drug Administration; Center for Drug Evaluation and Research (CDER). January 1999

Steinijans V-W, Hauschke D (1997) Individual bioequivalence. A European perspective. J Biopharm Stat 7 (1):31–34

Absolute and Relative Bioavailability

48

Khaled Abo-EL-Sooud

Contents

Abstract

Attaining desired blood drug concentration and, consequently, augmenting the bioavailability of poorly absorbed drugs have always been an essential aspect for the pharmaceutical agency. The achievement of this target gives positive economic benefits as reducing drug dosage and medical impacts in decreasing toxicity and bacterial resistance in case of antimicrobials. Various factors may reduce the availability of drugs. There are numerous ways to estimate bioavailability. Various software's models have been developed to simplify such analyses. The newly developed programs should provide a range of modules for pharmacokinetic and pharmacodynamic analysis with a more user-friendly interface.

Attaining desired blood drug concentration and, consequently, augmenting the bioavailability of poorly absorbed drug molecules have always been an essential aspect of development plans for the pharmaceutical agency. The achievement of this target gives positive economic benefits as reducing drug dosage and frequency and medical impacts in decreasing toxicity and bacterial resistance in case of antimicrobials. A drug may be well absorbed orally because of good lipid solubility and yet not has a good oral bioavailability because of extensive presystemic loss. Various physiological factors reduce the availability of drugs prior to their entry into the systemic circulation. There are numerous ways that are

K. Abo-EL-Sooud (✉)
Pharmacology Department, Faculty of Veterinary Medicine, Cairo University, Giza, Egypt
e-mail: kasooud@cu.edu.eg

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_16

followed to estimate bioavailability. To simplify such analyses, various software's models have been developed. The newly developed programs should provide a range of modules for pharmacokinetic and pharmacodynamic analysis with a more user-friendly interface. Factors that reduce the availability of drugs prior to their entry into the systemic circulation should be considered during prescription. Understanding the difference between absolute and relative bioavailability and be able to convert between these values is an essential issue. Relative bioavailability is one of the significant measures used to assess bioequivalence between several drug products.

Purpose and Rationale

The part of parent drug that absorbed, i.e., reached, the systemic circulation (blood). It is symbolized by the letter **F** (expressed in percent). The dosage form and molecular size of the drug determine to a great extent how much of a drug reaches the systemic circulation (Musther et al. 2014).

The rate and extent to which a drug is absorbed and available at the site of action are of the concerns of the bioavailability (Zweig et al. 2007). Rate means "how fast" the drug is absorbed per time unit. The extent means the amount of the dose enters systemic circulation from the site of administration.

Regulatory decisions concerning marketed drugs depend upon their ability to achieve the C_{max} of the generic product and T_{max} (Endrenyi and Yan 1993). Therefore, it is recommended that the 90% confidence limits for the percentage ratio of the C_{max}/AUC (area under the curve) values of two drug products should be (based on their logarithmic averages or medians) between 75% and 133%.

The overall drug exposures are markedly higher after administration of solution as compared to capsule (De Beule and Van Gestel 2001). Consequently, FDA guidelines clearly state that the two formulations should not be used correspondently. The mean relative bioavailability of itraconazole capsule was 85% that of the solution, but drug absorption was variable, and overall drug concentrations were similar between formulations. Mean elimination half-lives of both formulations were nearly identical at approximately 33 h (Hasbach et al. 2017). The consequences of failing to recognize a potential difference could be substantial and include avoidable cost, increased risk of toxicity, inadequate tissue drug concentrations, and treatment failures.

Relative bioavailability is one of the measures used to assess bioequivalence between two drug products. For FDA approval, a generic manufacturer must demonstrate that the 90% confidence interval for the ratio of the mean responses (usually of AUC and, C_{max}) of its product is within the limits of 80–125% (Chow and Liu 2009).

Some drugs administered orally are poorly bioavailable as they readily undergo first-pass metabolism and incomplete absorption. Drug efficacy can be severely limited by poor aqueous solubility, and some drugs also show side effects due to their poor solubility (Chaudhary et al. 2012). Thus, there is need of molecules which themselves have no same therapeutic activity but when combined with other drugs/molecules enhance their bioavailability. Many natural compounds from medicinal plants have the capacity to augment the bioavailability when co-administered with another drug (Tatiraju et al. 2013).

Procedure

The absolute bioavailability is the dose-corrected AUC of extravascular route divided by AUC intravenous. For example, the formula for calculating **F** for a drug administered depends upon the calculated $AUC_{0-\infty}$ using the linear trapezoidal rule according to the following equations (Abo-EL-Sooud et al. 2017):

$$\mathbf{F} = [\text{mean } AUC_{\text{NON-IV}} / \text{mean } AUC_{\text{IV}}] \times 100.$$

Therefore, a drug given by the intravenous route will have an absolute bioavailability of 100% ($\mathbf{F} = 1$), whereas drugs given by other

routes usually have an absolute bioavailability of less than one. When the bioavailability of extravascular route is very close to unity, this indicates that the absorption of the tested drug from this site is nearly completed. If we compare the two different dosage forms having same active ingredients and compare the two-drug bioavailability, it is called relative bioavailability (Gray and Jones 2004).

Relative Bioavailability and Bioequivalence

In pharmacology, relative bioavailability measures the bioavailability (estimated as the AUC) of a formulation (A) of a certain drug when compared with another formulation (B) of the same drug, usually an established standard or through administration via a different route. When the standard consists of an intravenously administered drug, this is known as absolute bioavailability (Shargel et al. 2004).

While the mechanisms by which a formulation affects bioavailability and bioequivalence have been extensively studied in drugs, formulation factors that influence bioavailability and bioequivalence in nutritional supplements are largely unknown (Cuiné et al. 2008). As a result, in nutritional sciences, relative bioavailability or bioequivalence is the most common measure of bioavailability, comparing the bioavailability of one formulation of the same dietary ingredient to another.

Evaluation

The pharmacokinetic analysis is performed using compartmental and noncompartmental models. From both models, the obtained AUC is the unique parameter for bioavailability estimation. Noncompartmental analysis (NCA) used to calculate pharmacokinetic parameters does not assume the number of compartments (Foster 2007). The compartmental analysis (CA) describes the decline in blood drug concentration as a function of time and predicts the drug concentrations at any time. This type classify the body to central (blood) and peripheral compartments (tissues). Drug that is in the peripheral compartment can return to the central compartment or other different tissues. The model may be best described by the compartmental open model which may be mono-exponential or bi-exponential or tri-exponential (Okusanya et al. 2007).

Biexponential expression

$$\mathrm{Cp} = \mathrm{A}e^{-\alpha t} + \mathrm{B}e^{-\beta t}$$

where Cp is the drug concentration in serum at time T; A and B are the intercepts of the distribution and elimination phases, respectively, with the concentration axis expressed as μg/ml; α and β are the distribution and elimination rate constants, respectively, expressed in units of reciprocal time (h^{-1}); and e is the natural logarithm base. The pharmacokinetic parameters and AUC are calculated according to Baggot and McKellar (1994).

The noncompartmental parameters are calculated by using the statistical moment theory (Wolfsegger and Jaki 2009). The elimination half-life ($T_{½el}$) is calculated as $\ln 2/\beta$. The AUC was calculated according to the trapezoidal rule. The mean residence time (MRT) is calculated as AUMC/AUC, where AUMC is the area under the first moment curve and AUC is the area under the curve.

There are numerous ways that are followed to estimate bioavailability. Site-specific analysis can offer essential information that can influence a risk assessment. Bioavailability can likewise be resolved for other extravascular courses of the organization, for example, intramuscular, subcutaneous, rectal, and sublingual (Morris et al. 2011). Sublingual and rectal routes are commonly used to avoid hepatic first-pass impact (Narang and Sharma 2011). Bioavailability of most small molecular weight drugs administered intramuscularly alternately subcutaneous injections depends upon the mechanism and the rate of transport to general circulation. High molecular weight formulations administered enter the blood to a limited extent through the lymphatic pathway (Khan et al. 2013).

The AUC is one of the essential pharmacokinetic parameters used to calculate others such as clearance or bioavailability. The AUC tells us how much drug is in the body and has units of concentration*time (e.g., mg*h/L). There are several methods to calculate AUC, ranging from more complex but more accurate methods like the integration of the equation that describes the pharmacokinetic profile to an easier, but less accurate method is called the trapezoidal method. This latter method is commonly used because of its ease and is based on the idea of calculating the area of a trapezoid (Persky 2012).

When a bioavailability study is carried out, at least two dosage forms are administered to each subject. One dosage form is the product to be tested, while the other dosage form is a standard or reference dosage form. This may be an IV dose, oral solution, or most commonly the original manufacturer's product. The doses are given with sufficient time between administrations for the drug to "wash out" or are completely eliminated. We usually assume that each subject eliminates each dosage form at the same rate (Toutain and Bousquet-Mélou 2004).

Relative Bioavailability

The relative bioavailability is concerned with the extent to which an extravascular drug product (e.g., a generic drug product) is absorbed in comparison with the trade name, or currently marketed drug product. This is usually determined by comparing the AUC of the plot of plasma drug concentration vs. time of the new product to that of the trade name product, i.e.,

$$\text{relative F} = \text{AUC}_{\text{generic}}/\text{AUC}_{\text{trade name}}$$

In case of evaluating pharmacokinetic interactions, the AUC $_{0-\infty}$ is calculated using the linear trapezoidal rule alone or with co-administered drug, and relative bioavailability is calculated according to the following equations (Abo-EL-Sooud et al. 2017):

$$F_{\text{rel}} = \left[\frac{\text{mean AUC}_{\text{with co-administered}}}{\text{mean AUC}_{\text{without co-administered}}}\right] \times 100.$$

Critical Assessment of the Method

A drug may be well absorbed orally because of good lipid solubility and yet not has a good oral bioavailability because of extensive presystemic loss. While the intravenous bioavailability of drugs is always 100%, the oral bioavailability is usually less than 100% because of incomplete absorption and/or first-pass elimination (Sim 2015).

The absolute bioavailability of a drug, when administered by an extravascular route, is usually less than one (i.e., $\mathbf{F} < 100\%$). Various physiological factors reduce the availability of drugs prior to their entry into the systemic circulation. Whether a drug taken with or without food will also affect absorption, other drugs taken concurrently may alter absorption and first-pass metabolism, intestinal motility alters the dissolution of the drug, and may affect the degree of chemical degradation of the drug by intestinal microflora (Guerville and Boudry 2016).

Factors affecting bioavailability:

Physical properties of the drug molecules
The form pharmaceutical formulation (short- or long-acting)
The presence of food in the stomach
Gastric emptying rate
Drug-drug and drug-food interactions
Gastrointestinal tract integrity
Microsomal enzymes inducers or inhibitors
First pass effect enterohepatic circulation, diet, gender
Disease condition especially hepatorenal functions

In clinical trials, interindividual variation is a critical measurement used to assess the bioavailability differences from patient to patient in order to ensure predictable dosing (Howgate et al. 2006).

Table 1 Residual analysis for drug concentrations after intravenous input

Time	Conc	Ln(Conc)	Conc_pre	Residual	Weight
0.25	1.64	0.49469624	1.63994069	5.931E-05	1
0.5	1.139	0.13015068	1.13988013	−0.0008801	1
1	1.06	0.05826891	1.03101809	0.02898191	1
2	0.81	−0.210721	0.92523711	−0.1152371	1
4	0.78	−0.2484614	0.74587998	0.03412002	1
6	0.77	−0.2613648	0.60129125	0.16870875	1
8	0.52	−0.6539265	0.48473103	0.03526897	1
10	0.221	−1.5095926	0.39076598	−0.169766	1

Table 2 Compartmental analysis of plasma data after intravenous input

Parameter	Unit	Value
A	μg/ml	5.219348692
Alpha	1/h	9.204329903
B	μg/ml	1.147722998
Beta	1/h	0.107742638
Parameter	Unit	Value
k10	1/h	0.567500333
k12	1/h	6.997086385
k21	1/h	1.747485823
t1/2Alpha	h	0.075306642
t1/2Beta	h	6.433360013
C0	μg/ml	6.36707169
V	(mg/kg)/(μg/ml)	0.785290357
CL	(mg/kg)/(μg/ml)/h	0.445652539
V2	(mg/kg)/(μg/ml)	3.144371412
CL2	(mg/kg)/(μg/ml)/h	5.494744465
AUC 0-t	μg/ml*h	7.59265674
AUC 0-inf	μg/ml*h	11.219503
AUMC	μg/ml*h^2	98.9310014
MRT	h	8.817770398
Vss	Mg/kg/(μg/ml)	3.929661769

Table 3 Noncompartmental analysis of plasma data after extravascular input

Parameter	Unit	Value
Lambda_z	1/h	0.118444235
t1/2	h	5.85209726
Tmax	h	1
Cmax	μg/ml	29
Tlag	h	0
Clast_obs/Cmax	μg/ml	0.155172414
AUC 0-t	μg/ml*h	88
AUC 0-inf_obs	μg/ml*h	125.9925626
AUC 0-t/0-inf_obs	μg/ml*h	0.698453926
AUMC 0-inf_obs	μg/ml*h^2	1160.299044
MRT 0-inf_obs	h	9.209266163
Vz/F_obs	(mg)/(μg/ml)	55.28344692
Cl/F_obs	(mg)/(μg/ml)/h	6.548005555

Modifications of the Method

Tiresome mathematical calculations, optimization algorithms, and graph plotting are essential for pharmacokinetic data analysis. To simplify such analyses, different software's models have been developed. Many of these commercially available packages are expensive or have a steep learning curve.

Several models and add-in programs have previously been constructed for diverse applications (Meineke 2000; Brown 2006**)**. The pharmacokinetic analysis has been adopted for such NCA computation (Dansirikul et al. 2005; Jaki and Wolfsegger 2012) and bioavailability/bioequivalence trials (Abdallah and Ludden 1995; Chow et al. 2011**).** However, several programs have only one type of analysis mostly NCA calculation functions or need a specific spreadsheet templates with limiting input data.

Therefore, it is valuable exploring the possibility of cost-effective and easy-to-use alternatives for pharmacokinetic and pharmacodynamic analysis. PKSolver is an available menu-driven add-in program used Microsoft Excel in Visual Basic for Applications (VBA) (Zhang et al. 2010). The program provides a range of modules for pharmacokinetic and pharmacodynamic analysis other than NCA, including CA (Watabe et al. 2006),

Table 4 Compartmental analysis of enterohepatic circulation model after oral dose input

Parameter	Unit	Value
Ka	1/h	1.1651223
k10	1/h	0.698827735
k1 g	1/h	0.466307328
Ttom	h	5.557105687
V/F	Mg/(μg/ml)	0.321034708
Parameter	Unit	Value
t1/2 ka	1/h	0.594913668
t1/2 k1 g	1/h	0.991871309
t1/2 k10	1/h	1.486459979
D_rec	Mg	7.912095782
D_rec/dose	Mg	0.395604789
Tmax1	h	0.858274297
Cmax1	μg/ml	22.9182339
Tmax2	h	6.415379984
Cmax2	μg/ml	9.330701874
AUC 0-t	μg/ml*h	74.52201017
AUC 0-inf	μg/ml*h	74.62153654
AUMC	μg/ml*h^2	128.0914937
MRT	h	1.716548594
CL/F	(mg)/(μg/ml)/h	0.268019139

pharmacodynamic modeling (Felmlee et al. 2012), multiple absorption sites (Plusquellec et al. 1999; Kota et al. 2007; Abo-EL-Sooud et al. 2017), and enterohepatic circulation (Gabrielsson and Weiner 1999; Roberts et al. 2002; Abo-El-Sooud et al. 2016), which were developed for fitting the double-peak concentration-time profile based on the classical one-compartment model (Tables 1, 2, 3, and 4). The parameters estimated with PKSolver are satisfactory. In conclusion, the PKSolver simplified the pharmacokinetic and pharmacodynamic data analysis process and its output could be generated in Microsoft Word in the form of an integrated report. The program provides pharmacokinetic researchers with a fast and easy-to-use tool for routine and basic pharmacokinetic and pharmacodynamic data analysis with a more user-friendly interface.

References and Further Reading

Abdallah HY, Ludden TM (1995) A spreadsheet program for simulation of bioequivalence and bioavailability studies. Comput Biol Med 25:349–354. https://doi.org/10.1016/0010-4825(95)00015-V

Abo-El-Sooud K, Swielim GA, Wally YR, El-Gammal SM (2016) Comparative biliary and serum kinetics of doxycycline after oral and intramuscular routes with special reference to its unique entero-hepatic circulation in turkeys. Wulfenia 23:206–219

Abo-EL-Sooud K, Mouneir SM, Fahmy MAF (2017) Curcumin ameliorates the absolute and relative bioavailabilities of marbofloxacin after oral administrations in broiler chickens. Wulfenia 24:284–297

Baggot JD, McKellar QA (1994) The absorption, distribution and elimination of anthelmintic drugs: the role of pharmacokinetics. J Vet Pharmacol Ther 17:409–419

Brown AM (2006) A non-linear regression analysis program for describing electrophysiological data with multiple functions using Microsoft Excel. Comput Methods Prog Biomed 82:51–57. https://doi.org/10.1016/j.cmpb.2006.01.007

Chaudhary A, Nagaich U, Gulati N et al (2012) Enhancement of solubilization and bioavailability of poorly soluble drugs by physical and chemical modifications: a recent review. J Adv Pharm Educ Res 2:32–67. https://doi.org/10.1016/S0040-6031(02)00451-3

Chow S-C, Liu J-P (2009) Design and analysis of bioavailability and bioequivalence studies. J Chem Inf Model 53:758. https://doi.org/10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C

Chow S-C, Endrenyi L, Chi E et al (2011) Statistical issues in bioavailability/bioequivalence studies. J Bioequiv Availab S1:007. https://doi.org/10.4172/jbb.S1-007

Cuiné JF, Mcevoy CL, Charman WN et al (2008) Evaluation of the impact of surfactant digestion on the bioavailability of danazol after oral administration of lipidic self-emulsifying formulations to dogs. J Pharm Sci 97:995–1012. https://doi.org/10.1002/jps.21246

Dansirikul C, Choi M, Duffull SB (2005) Estimation of pharmacokinetic parameters from non-compartmental variables using Microsoft Excel? Comput Biol Med 35:389–403. https://doi.org/10.1016/j.compbiomed.2004.02.008

De Beule K, Van Gestel J (2001) Pharmacology of itraconazole. Drugs 61:27–37. https://doi.org/10.1016/S1043-6618(02)00088-9

Endrenyi L, Yan W (1993) Variation of C_{max} and C_{max}/AUC in investigations of bioequivalence. Int J Clin Pharmacol Ther Toxicol 31:184–189

Felmlee MA, Morris ME, Mager DE (2012) Mechanism-based pharmacodynamic modeling. Methods Mol Biol 929:583–600. https://doi.org/10.1007/978-1-62703-050-2-21

Foster DM (2007) Noncompartmental versus compartmental approaches to pharmacokinetic analysis. Princ Clin Pharmacol 8:102–130. https://doi.org/10.1016/B978-0-12-369417-1.50048-1

Gabrielsson JL, Weiner DL (1999) Methodology for pharmacokinetic/pharmacodynamic data analysis. Pharm Sci Technol Today 2:244–252

Gray M, Jones DP (2004) The effect of different formulations of equivalent active ingredients on the performance of two topical wound treatment products. Ostomy Wound Manag 50:34–40

Guerville M, Boudry G (2016) Gastrointestinal and hepatic mechanisms limiting entry and dissemination of lipopolysaccharide into the systemic circulation. Am J Physiol Gastrointest Liver Physiol 311: G1–G15. https://doi.org/10.1152/ajpgi.00098.2016

Hasbach AE, Langlois DK, Rosser EJ, Papich MG (2017) Pharmacokinetics and relative bioavailability of orally administered innovator-formulated itraconazole capsules and solution in healthy dogs. J Vet Intern Med 31:1163–1169. https://doi.org/10.1111/jvim.14779

Howgate EM, Rowland Yeo K, Proctor NJ et al (2006) Prediction of in vivo drug clearance from in vitro data I: impact of inter-individual variability. Xenobiotica 36:473–497. https://doi.org/10.1080/00498250600683197

Jaki T, Wolfsegger MJ (2012) Non-compartmental estimation of pharmacokinetic parameters for flexible sampling designs. Stat Med 31:1059–1073. https://doi.org/10.1002/sim.4386

Khan AA, Mudassir J, Mohtar N, Darwis Y (2013) Advanced drug delivery to the lymphatic system: lipid-based nanoformulations. Int J Nanomedicine 8:2733–2744

Kota J, Machavaram KK, McLennan DN et al (2007) Lymphatic absorption of subcutaneously administered proteins: influence of different injection sites on the absorption of darbepoetin alfa using a sheep model. Drug Metab Dispos 35:2211–2217. https://doi.org/10.1124/dmd.107.015669

Meineke I (2000) An add-in implementation of the RESAMPLING syntax under Microsoft EXCEL. Comput Methods Prog Biomed 63:99–104. https://doi.org/10.1016/S0169-2607(00)00077-8

Morris CA, Duparc S, Borghini-Fuhrer I et al (2011) Review of the clinical pharmacokinetics of artesunate and its active metabolite dihydroartemisinin following intravenous, intramuscular, oral or rectal administration. Malar J 10:263

Musther H, Olivares-Morales A, Hatley OJD et al (2014) Animal versus human oral drug bioavailability: do they correlate? Eur J Pharm Sci 57:280–291. https://doi.org/10.1016/j.ejps.2013.08.018

Narang N, Sharma J (2011) Sublingual mucosa as a route for systemic drug delivery. Int J Pharm Pharm Sci 3:18–22

Okusanya O, Forrest A, DiFrancesco R et al (2007) Compartmental pharmacokinetic analysis of oral amprenavir with secondary peaks. Antimicrob Agents Chemother 51:1822–1826. https://doi.org/10.1128/AAC.00570-06

Persky AM (2012) The impact of team-based learning on a foundational pharmacokinetics course. Am J Pharm Educ 76:1–10. https://doi.org/10.5688/ajpe76231

Plusquellec Y, Efthymiopoulos C, Duthil P, Houin G (1999) A pharmacokinetic model for multiple sites discontinuous gastrointestinal absorption. Med Eng Phys 21:525–532

Roberts MS, Magnusson BM, Burczynski FJ, Weiss M (2002) Enterohepatic circulation: physiological, pharmacokinetic and clinical implications. Clin Pharmacokinet 41:751–790. https://doi.org/10.2165/00003088-200241100-00005

Shargel L, Wu-Pong S, Yu A (2004) Applied biopharmaceutics and pharmacokinetics, 7th edn. McGraw-Hill education, New York, pp 471–475. https://doi.org/10.7326/0003-4819-94-6-826_2

Sim DSM (2015) Drug absorption and bioavailability. In: Chan YK, Ng KP, Sim DSM (eds) Pharmacological basis of acute care. Springer International Publishing, Switzerland, pp 17–26

Tatiraju DV, Bagade VB, Karambelkar PJ et al (2013) Natural bioenhancers: an overview. J Pharmacogn Phytochem 2:55–60

Toutain PL, Bousquet-Mélou A (2004) Bioavailability and its assessment. J Vet Pharmacol Ther 27 (6):455–466

Watabe H, Ikoma Y, Kimura Y et al (2006) PET kinetic analysis – compartmental model. Ann Nucl Med 20:583–588. https://doi.org/10.1007/BF02984655

Wolfsegger MJ, Jaki T (2009) Non-compartmental estimation of pharmacokinetic parameters in serial sampling designs. J Pharmacokinet Pharmacodyn 36:479–494. https://doi.org/10.1007/s10928-009-9133-9

Zhang Y, Huo M, Zhou J, Xie S (2010) PKSolver: an add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel. Comput Methods Prog Biomed 99:306–314. https://doi.org/10.1016/j.cmpb.2010.01.007

Zweig AS, Tress ML, Vanden Hoek TL (2007) Absorption of drugs transport of a drug from the GI tract factors affecting drug absorption related to patient. J Pharm Pharmacol 7:23–25

Bioequivalence

49

Henning H. Blume

Contents

Abstract

Exposure in the central circulation is an essential prerequisite for efficacy and safety of systemic drugs. After oral administration the rate and extent of absorption from the gastrointestinal tract may be significantly determined by the biopharmaceutical properties of both, the drug substance as well as the drug formulation. This is the reason why rate and extent of bioavailability of oral dosage forms needs to be characterized thoroughly during drug product development for regulatory submission. In case of generic medicinal products assessment of bioequivalence in comparison to an appropriate reference product is the basis for marketing authorization applications and the approval process. The

H. H. Blume (✉)
Frankfurt am Main, Germany

SocraTec C&S GmbH, Oberursel, Germany
e-mail: henning.blume@socratec-pharma.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_17

requirements for this procedure are therefore clearly defined in the bioequivalence guidelines published by the competent regulatory authorities, e.g., EMA in Europe or FDA in the USA.

Most relevant parameters to be considered in this context are discussed in this chapter. Since there are still certain differences between the main guidelines, activities have been started to discuss the basis for science-driven regulations and to harmonize the existing requirements. In particular the Global Bioequivalence Harmonization Initiative (GBHI) of the European Federation for Pharmaceutical Sciences (EUFEPS) has achieved significant contributions to this process.

Introduction and Definitions

Bioavailability (BA) and bioequivalence (BE) are essential elements in clinical research and drug product development. The characterization of systemic exposure is equally relevant for innovative as well as generic preparations considering that systemic drugs develop their efficacy only if the active ingredient is absorbed into the bloodstream and becomes available at its site of action. It is, therefore, mandatory for all newly developed medicinal products to demonstrate their appropriate bioavailability in a dossier submitted along with the Marketing Authorisation Application. In case of generic products, the assessment of bioequivalence is the most essential prerequisite in order to confirm their therapeutic equivalence with the innovator's reference product in terms of efficacy as well as safety.

The definitions of bioavailability and bioequivalence are very similar in the global regulatory landscape. As an example the following definitions were taken from the CHMP Note for Guidance 2001, which was the last European guideline describing standards for BA and BE:

- **BA**: "Bioavailability means the rate and extent to which an active substance or active moiety is absorbed from a pharmaceutical form and becomes available at the site of action." However, considering that in the majority of cases measurement at the site of action is not possible, a more practical definition was suggested taking into account that the substance in the general circulation must be in exchange with its presence at the site of action: "Bioavailability is understood to be the extent and the rate at which a substance or its active moiety is delivered from a pharmaceutical form and becomes available in the general circulation."
- **BE**: "Two medicinal products are bioequivalent if they are pharmaceutically equivalent or pharmaceutical alternatives and if their bioavailabilities after administration in the same molar dose are similar to such degree that their effects, with respect to both efficacy and safety, will be essentially the same."

It is important to consider that the assessment of bioavailability should reflect the in vivo performance of a drug product after its administration in line with the recommendations given by the labeling. In case of oral products, this includes several important biopharmaceutical processes, starting with disintegration of a solid dosage form in the stomach or proximal intestine, followed by drug release from the preparation and dissolution of the compound. All these processes may have an essential impact on drug absorption into the body. Moreover, also a pre-systemic "first-pass" effect is of importance, which may occur during absorption in the intestinal enterocytes (gut wall first-pass metabolism) or after absorption in the liver (hepatic first-pass metabolism). Major steps in this context are elucidated in Fig. 1:

But there are also other aspects which can significantly affect the bioavailability of oral drug products. In this context the administration conditions – fasted or fed – play a relevant role as they may impact the transit of non-disintegrating solid oral dosage forms through the gastrointestinal (GI) tract. In particular gastric emptying will be significantly delayed in fed state. However, GI transit can also be modified by certain excipients, e.g., accelerated by sugar alcohols such as mannitol, or vary depending on tablet size and shape.

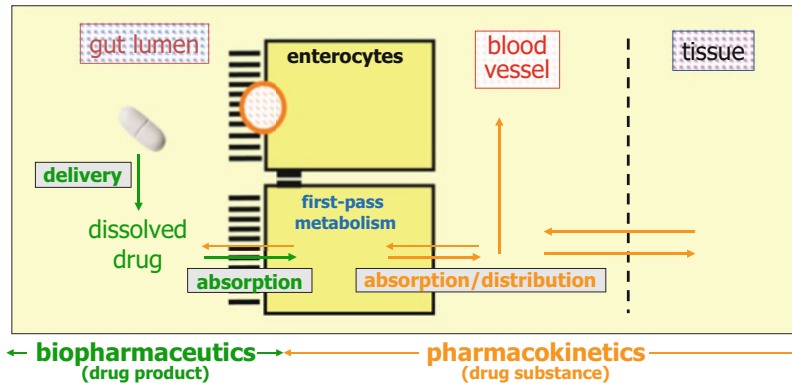


Fig. 1 Disposition of systemic drugs: pharmacokinetic processes

As a consequence, for product comparisons – e.g., with the intention of BE assessment in case of generic drug development – constant administration conditions are essential, while various study settings, fasted and fed state, should be investigated in order to properly characterize and understand product performance in changing situations of treatment.

Concept of BA/BE Assessment

Even though the general study design for BA and BE assessments is very similar (except the statistical sample size estimation and confirmative evaluation of study objectives in case of BE), all regulatory guidelines clearly differentiate between both scenarios. This seems useful considering that only more general suggestions are given for BA studies, while very restrictive requirements have been defined for BE projects. The latter includes also procedures how the pharmacokinetic parameters should be calculated and which to be used for evaluation.

Characterization of Bioavailability

As bioavailability is an essential prerequisite for the efficacy and safety of systemic products, proper characterization of their systemic total and peak exposure as well as pharmacokinetic properties is requested for regulatory purposes of all medicinal products – with only very few exceptions of preparations with self-explanatory exposure, e.g., gases for inhalation.

In this context it should be carefully differentiated between properties of the drug substance and characteristics of the drug product. Absorption rate constant, volume of distribution, drug biotransformation, and elimination rate constant as well as total, hepatic, and renal clearances are primarily drug substance-related properties. They should be properly characterized for all active ingredients. However, some of those parameters can also considerably be affected by the properties of the pharmaceutical form, in particular drug absorption. As consequence, systemic total and peak exposure need to be determined as well.

Pharmacokinetic Properties

In order to determine the pharmacokinetic characteristics of an active compound, the drug substance should be administered in a way which allows investigating its in vivo properties (absorption, distribution, metabolism, and elimination/excretion) widely unmodified by the biopharmaceutical properties of a pharmaceutical formulation, which can in particular affect drug absorption but also first-past metabolism to certain extent.

This goal is, however, not easily achievable as any administration of a drug substance needs certain application form which, at the same time, can modify GI transit and drug absorption. The best option in this context is to investigate an aqueous solution either for intravenous (i.v.) or oral

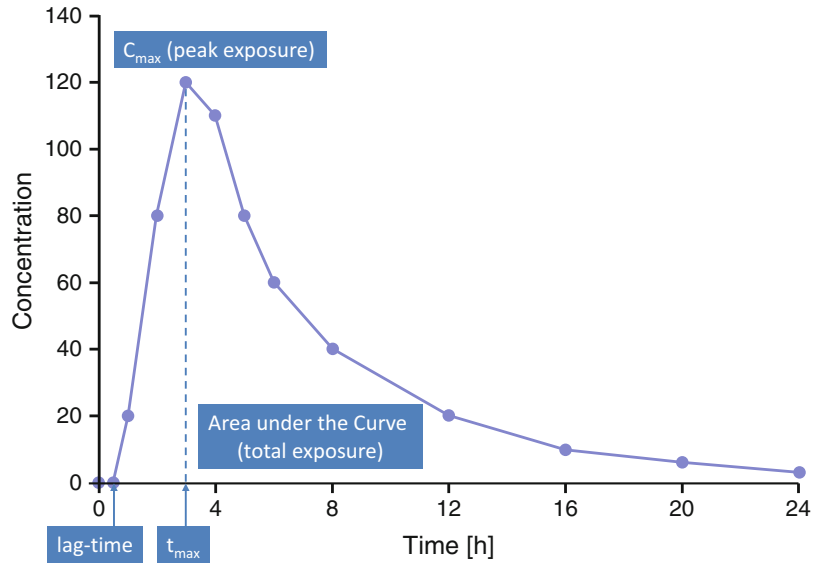


Fig. 2 Typical plasma profile after fasted oral administration of a conventional IR tablet

administration. Under such conditions uptake of the active ingredient should occur widely unaffected by excipients or pharmaceutical technology measures, e.g., tabletting.

- **i.v. application**: Pharmacokinetic parameters determined after i.v. application of an aqueous solution are the best measures to characterize drug substance properties as they will not be affected by any modification of drug absorption.
- **Oral versus i.v. administration**: Comparison of orally and intravenously administered aqueous solutions opens the opportunity to determine the rate of drug absorption widely unaffected by a pharmaceutical formulation and its potential impact on other pharmacokinetic parameters like first-pass metabolism which may be affected to certain extent also by the rate of absorption.

Bioavailability

Bioavailability studies generally need to be designed as comparative investigations. Reliable conclusions will not be possible from clinical trials which include only one study arm as the pharmacokinetic (PK) parameters determined in such a case may be modified significantly by the properties of the study population as well as other study conditions. Only comparisons with known (e.g., approved reference product) or well-established (e.g., oral solution) in vivo performance will allow proper conclusion on the bioavailability of a newly developed product, which should reflect the rate and extent of drug absorption.

The main BA characteristics derived from plasma profiles (Fig. 2) are:

- **Total exposure**: The degree of intestinal drug absorption (extent of bioavailability) can be determined as the area under the drug concentration versus time curve (AUC) which has been shown to be proportional to the amount of the active ingredient reaching the systemic circulation.
- **Peak exposure**: This parameter represents the highest value (C_{max}) in a plasma concentration-time profile. It is dependent on both, the rate of absorption (the more rapid absorption occurs, the higher concentration peaks will be) and the extent of absorption (C_{max} normally increases with rising AUC). As consequence, this is a "hybrid" PK parameter and, thus, not a perfect characteristic for the rate of drug absorption. On the other hand, C_{max} is clinically important as it will impact both, efficacy and safety of a drug.

- **Time-point of peak exposure (t_{max})**: This parameter represents the time-point when the maximum concentration (C_{max}) of a plasma concentration-time profile is achieved. More rapid absorption will result in shorter t_{max} values.
- **Lag time**: In cases of a delayed onset of drug absorption, the period between the time-point of drug administration and the occurrence of the first measurable plasma concentration is defined as lag time (t_{lag}). This parameter is primarily affected by drug formulation properties, e.g., in case of gastro-resistant products.

Determination of the parameters C_{max}, t_{max}, and t_{lag} deviates between BA and BE studies. While certain degrees of freedom are accepted in BA studies with respect to the calculation of these parameters (including an assessment by use of pharmacokinetic models), restrictive requirements have been defined for BE assessment as clearly stated in the guidelines. C_{max} and t_{max} should be directly derived from the measured plasma concentration-time profile as the highest observed concentration and the time-point of its occurrence.

With regard to the study concepts and the preparations administered, the following scenarios should be differentiated:

- **Assessment of absolute bioavailability**: In this case preparations administered on a non-parenteral route will be compared with an intravenous application form. Values determined as absolute bioavailability allow conclusions on certain loss in exposure due to limited absorption or first-pass metabolism. Such conclusions will be best if the compound administered via the non-parenteral route is given as (aqueous) solution.
- **Assessment of relative bioavailability**: In this case preparations are administered via the same route in order to allow comparison of their biopharmaceutical properties. However, also systemic exposure of products applied on different ways – e.g., orally versus topically – may be compared. In the latter case an appropriate interpretation of the results needs to consider also information on divergent conditions for drug absorption (and first-pass metabolism) at the different sites of application.

While the absolute bioavailability is more a drug substance-related characteristics, the relative bioavailability should allow relevant conclusions on the biopharmaceutical properties of drug products and, thus, the quality of their formulation.

Assessment of Bioequivalence

In contrast to more general suggestions for BA studies, very strict and binding requirements have been defined in the guidelines for the assessment of bioequivalence as basis for an "abridged application" in Europe resp. an "Abbreviated New Drug Application" in the USA. This includes, among others, the following aspects in particular:

- **Prospective study planning**: All conditions for study conduct and evaluation of the results should be pre-specified in the protocol. Later changes or post hoc extensions of the project will not be acceptable in the majority of cases.
- **Statistical sample size estimation**: This essential element of a prospective study planning should assure that the number of subjects included into the clinical part will be sufficient for confirmation of BE (in case the products are comparable enough for such conclusion).
- **Administration conditions**: In general fasted state is considered most discriminative and, thus, most sensitive to detect differences between products. However, BE may deviate between fasted and fed administration. The US-FDA therefore often requests BE assessments under both conditions. In the EU those conditions should be applied for BE assessments which are suggested for the reference product in its labeling.
- **Reference product**: The reference product should be a medicinal product which has been approved based on full clinical documentation. In the USA these products are defined by the

FDA in the *Orange Book* as "Reference Listed Drugs." In Europe the identification of the appropriate reference product is often not as easy, especially in cases of older products ("grandfather's drugs"), and in such cases it may be advisable to contact the agencies for confirmation of the right selection.

- **Characterization of systemic exposure**: As mentioned above clear rules have been defined for the determination of plasma profiles (e.g., use of a validated bioanalytical procedure, sufficient number and proper distribution of sampling time-points over the entire curve, etc.). Compliance with these requirements is normally strictly enforced by the reviewers in BE studies conducted for regulatory purposes, e.g., for generic drug applications.
- **Statistical evaluation and bioequivalence assessment**: The statistical methods to be used for BE assessment and the general acceptance criteria for a conclusion on bioequivalence are also clearly defined in the guidelines and need to be considered in studies for regulatory submission.

Even though guidelines only present suggestions concerning the appropriate study design and clinical conduct of BE investigations as well as the statistical evaluation of study results, applicants are well advised to principally follow those suggestions. However, one should be aware that BE requirements can never apply to a "one-size-fits-all" principle, and thus, it is essential to adequately "adjust" those suggestions given by the guidelines to the individual question(s) to be investigated.

A proper clinical rationale should be the guiding principle to define all details in study design, clinical conduct, and statistical evaluation. All these aspects should be laid down a priori in the study protocol, and all measures need to follow these definitions. This includes also all relevant aspects of the "Good Clinical Practice" requirements described in the international guidelines as well as the ethical principles laid down in the Declaration of Helsinki.

Design and Conduct of In Vivo Bioequivalence Studies

The general concept and basic principles of bioequivalence studies will be described in the following. More practical details may be found in regulatory guidelines and product-specific guidances relevant for each jurisdiction.

Concept and Essential Elements of BE Studies

Assessment of bioequivalence as basis for generic drug approval should be established on a comparison of medicinal products containing same active ingredient(s) in same molar dose in "comparable" (US-FDA: "same") dosage forms, e.g., tablets with tablets or also tablets with capsules.

Normally these studies are conducted in healthy volunteers following a crossover design as this is considered most appropriate in order to reduce the within-subject variability and, thus, should allow a more discriminative comparison between the products. Only in exceptional cases BE studies will be conducted in a patient populations, in particular if investigations in healthy subjects do not seem adequate and acceptable due to safety or tolerability concerns, e.g., in oncology. Thus, medical and ethical aspects need to be taken into account as well.

Single-Dose Studies

Assessment of systemic peak and total exposure after single-dose administration is the most conventional approach in the clinical development of generic drug products. Such design is generally considered most sensitive in detecting potential differences between formulations.

Administration occurs in fasted or fed state together with 240 mL of non-carbonated water. In this context different requirements are relevant in Europe (similar in Canada and Australia) or in the USA. While the EU guidelines generally suggest an administration according to the labeling of the reference product, the US-FDA more often requires BE assessments after both fasted and fed administration, also in cases of immediate-

release dosage forms. Detailed recommendations are given by product-specific guidances.

Plasma concentration versus time profiles will be determined in order to properly characterize the absorption, distribution, and elimination phases of the active ingredient. All relevant parts of a profile should be adequately described by measured values. In order to determine the bioavailability of the medicinal products sufficiently, minimum of 80% of the area under the curve (representing the extent of absorption) should be covered by measuring points.

For BE assessment total and peak exposure need to be considered and compared between the products on an individual basis (test and reference comparison for each subject). Total exposure – representing the extent of absorption – should be appropriately described by the AUC (area under the plasma concentration-time curve) which is calculated from all measured values. Peak exposure (C_{max}) will be taken directly from the plasma profiles as the highest determined concentration.

Other pharmacokinetic parameters of interest are the time of peak concentration (t_{max}), the lag time (t_{lag}), or the elimination half-life ($t_{1/2}$). They are, however, only considered descriptively (and not for confirmative analysis) as they are normally highly variable and, moreover, are predefined by the time-points for plasma sampling (t_{max}, t_{lag}). Thus, they are not considered suitable – and not mandatory – for assessment of bioequivalence.

Multiple-Dose Studies

Even though multiple dosing can be considered most relevant for the treatment of chronic diseases and, thus, should reflect therapeutic practice more realistically, such setting is not recommended for BE assessment. Rationale for this regulatory decision is that the within-subject variability will be dampened by multiple dosing, and consequently, studies at steady state may be less discriminative and, thus, also less sensitive to detect differences between products.

Exceptions from this general rule are only (1) studies in patient populations who need chronic treatment of their diseases and (2) investigations of modified-release products which release their active ingredient(s) so slowly that significant accumulation may occur after multiple dosing. The latter may be expected in cases of preparations developing plateau-like plasma profiles with less than 90% of total AUC covered within the intended dosing interval ($AUC_{0-\tau}$). In such cases the European authorities request for additional multiple-dose studies. This is a Europe-specific requirement which is not supported by the US-FDA which does generally not recommend multiple-dose studies for BE assessment except in patient populations.

Administration conditions (fed or fasted) for multiple-dose studies should be defined in line with the recommendations given by the Summary of Product Characteristics of the reference product. This includes normo-caloric instead of high-fat, high-calorie meals and also fasting periods adapted to realistic treatment situations (e.g., if fasted state is suggested by the SmPC, the post-dose fasting period must not be maintained for 4 h).

It is essential to confirm that steady state is achieved at the profiling day. Plasma profiles need to be characterized for at least one dosing interval; however, blood sampling over 24 h may be considered in order to take also diurnal variations in pharmacokinetics into account.

Similar to the assessment of BE in single-dose settings, total and peak exposure should be considered for comparison between the products also at steady state. Total exposure will be appropriately described by the AUC calculated from the measured values during the dosing interval (it has been demonstrated that this area [$AUC_{0-\tau}$] equals the $AUC_{0-\infty}$ in case of single-dose studies). Peak exposure (C_{max}) will be taken directly from the plasma profiles as the highest determined concentration. An additional important parameter for BE assessment at steady state is the trough value determined at the time-point of next dosing (end of dosing interval). This concentration is an important efficacy-related characteristic.

Another pharmacokinetic parameter of interest is PTF, the peak-trough fluctuation, which describes the range of fluctuation between maximum and minimum concentrations of a profile. PTF is, however, only considered descriptively as its variability is normally high and, thus, is

normally not suitable for confirmative BE analysis. Relevant steady-state parameters are shown in Fig. 3.

Drug-Food Interactions (Food Effect Studies)

The intake of food (or caloric beverages, including alcoholic) will immediately initiate a significant change in gastrointestinal physiology. Relevant alterations include an increase in gastric pH, bile secretion, and blood supply to the intestine. These modifications may impact the solubility of compounds and, consequently, their dissolution from medicinal products and often also the rate and extent of drug absorption. First-pass effect may be affected as well – especially in the gut wall – and decrease in certain cases with the consequence of rising systemic exposure of the drug substance. On the other hand, drug absorption might also be reduced due to binding or complexation of compounds by food ingredients.

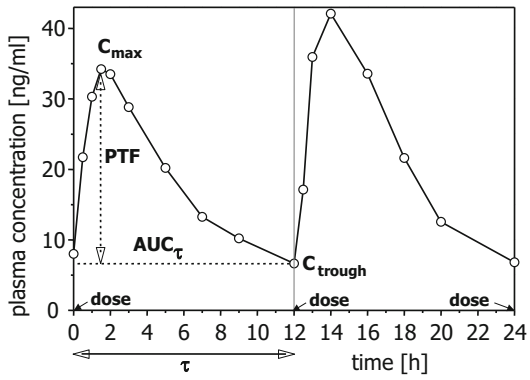


Fig. 3 Typical plasma profile and PK parameters after multiple dosing/at steady state

These are reasons why the US-FDA and some other authorities request BE studies in both, fed and fasted state, even in case of immediate-release drug products. So far such investigations are less frequently required in Europe.

There is, however, another essential change in gastrointestinal physiology induced by food intake: the GI transit of chyme – and non-disintegrating solid oral dosage forms as well – will occur retardedly after administration in the fed state. In this context a delay in gastric emptying is of particular relevance. Prolongation of gastric residence time can significantly impact the systemic drug exposure, especially in case of modified-release dosage forms.

Such meal-induced delay in gastric emptying is elucidated in Fig. 4: This example shows a generic alternative of the innovator's nifedipine OROS (*or*al *os*motic pump) preparation for once daily administration (Schug et al. 2000). The generic matrix tablet had been developed as enteric-coated form. The rationale for this surprising measure (there is no obvious reason for such coating in case of nifedipine tablets) was the attempt to mimic the delayed-release character of the innovator product (typical for an OROS). The coating avoids drug release during gastric

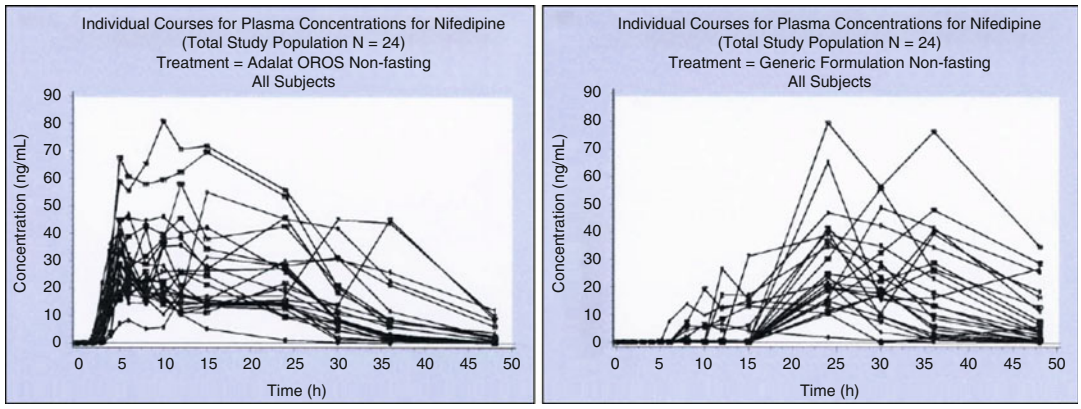


Fig. 4 Nifedipine plasma profiles after fed administration of enteric-coated tablets (left) or the OROS preparation (right) (Schug et al. 2000)

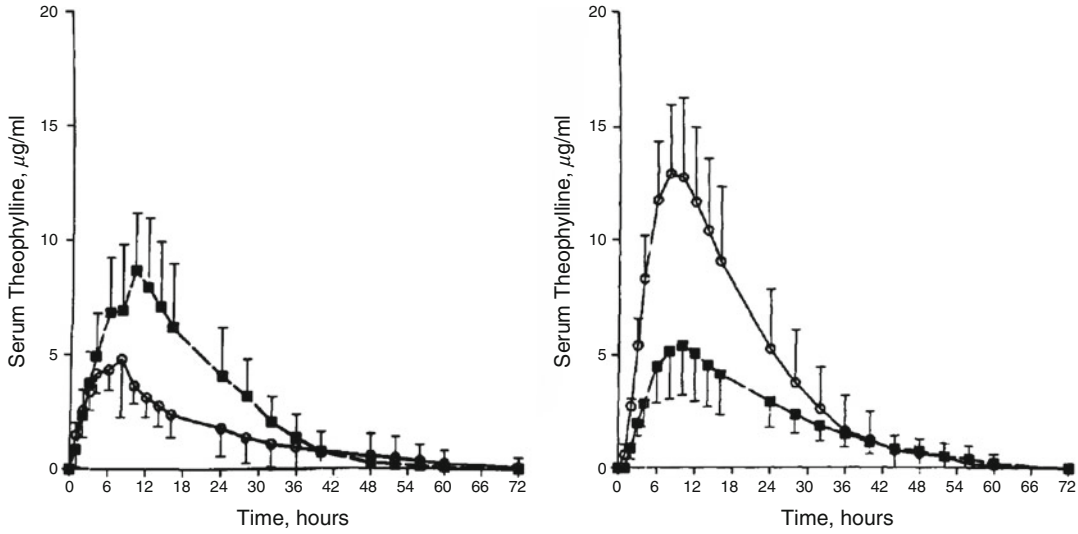


Fig. 5 Food-induced changes in exposure of nifedipine prolonged-release tablets: increase in AUC in one case (left graph) and reduction in another (right graph) (Karim et al. 1986)

residence. The consequences are, however, drastic – and certainly unintended: in fed state onset of drug absorption was delayed by 16 h or even longer in the majority of subjects while only by 2–4 h after intake of the osmotic system.

Numerous examples can be found in the literature demonstrating food-induced changes in exposure which can, moreover, vary between modified-release products containing the same active ingredient. Systematic investigations in the 1980s with several theophylline prolonged-release preparations marketed in the USA discovered considerably different and sometimes even opposite food effects between the products. Two very distinctive examples are shown in Fig. 5 (Karim et al. 1986).

These examples indicate that the changes initiated by food ingestion are not drug substance related but caused by the biopharmaceutical properties of the formulation. Based on these findings – which elucidate the general importance of food-induced changes in the exposure of modified-release preparations – such food interaction studies are meanwhile requested for all MR products by the international guidelines.

Food effect studies should be designed in order to investigate "worst-case" scenarios. Thus, the products should be administered immediately after a high-fat, high-calorie meal (50% fat, 800–1000 kilocalories) in comparison with intake after an overnight fast. All other conditions for BE assessment are identical as in studies after fasted administration.

Thus, in case of modified-release generic products, BE needs to be demonstrated after fasted as well as fed administration in comparison to the approved reference preparation.

Bioanalysis

Optimum conditions for the analytical determination are necessary in order to obtain reliable information from bioequivalence studies. Moreover, results reported from the studies need to be transparent and traceable from raw data to the finally listed values. All measures should assure reproducibility of the data. Parameters like selectivity, accuracy, and precision of the method as well as stability of the samples in all phases of the determination are essential elements in this context.

Bioanalytical method validation has, therefore, become one of the major issues in developing science-based regulations. Detailed requirements for pre-study and within-study validation procedures have been defined in CHMP and US-FDA guidelines, and compliance with these is critically reviewed during the regulatory assessment of

reports submitted along with the Marketing Authorisation Application. It is, thus, recommendable in BE studies for regulatory submission to carefully consider the suggestions given by these guidelines.

Analytes to Be Measured

Generally BE assessments should be established on the plasma concentration-time profiles determined for the parent compound. This requirement is supported by the assumption that these data reflect best the in vivo performance of the drug product after administration.

In certain cases measurement of metabolites may be recommended, especially if the parent drug concentrations are very low – e.g., in case of compounds undergoing pronounced first-pass metabolism – and, thus, their reliable analytical determination is difficult even with high-sophisticated analytical procedures. However, also in such situations, other attempts should be taken into consideration with the intention to achieve measurable plasma concentrations of the parent drug, e.g., administration of higher (event supratherapeutic) single doses.

These rules should also be applied to inactive prodrugs as even in these cases the parent compound should reflect most appropriately the performance of the products.

Statistical Analysis and Conclusion on Therapeutic Equivalence

The statistical procedures to be used for BE assessment are also clearly defined in the major international guidelines. After intensive discussions in the 1980s and 1990s, broad consensus has been achieved, and, accordingly, 90% confidence intervals for the ratio of the population geometric means should be calculated for the parameters used for confirmative BE analysis (e.g., AUC and C_{max}). This approach is equivalent to a two one-sided test procedure. Basis for such statistical calculations is an analysis of variance (ANOVA) applied to the parameters under consideration. The use of nonparametric analyses is generally not recommended.

Details of the procedure used for statistical analysis should be pre-specified in the protocol.

Results obtained from parametric statistical analysis need to be evaluated and interpreted for BE assessment by use of the preset bioequivalence acceptance criteria. There is meanwhile international consensus to apply an 80.00–125.00% acceptance range in this context. As long as the 90% confidence interval lies completely inside this range, conclusion on BE is accepted. Otherwise bioequivalence cannot be confirmed by the study results. In case of narrow therapeutic index drugs, even more restrictive requirements may need to be applied, e.g., 90.00–111.00% acceptance limits. In specific cases this will be indicated in the product-specific guidances published by EMA or US-FDA.

It is common understanding to conclude from bioequivalence on therapeutic equivalence of generic medicinal products. This is in line with the concept of an abridged application in the EU (resp. the Abbreviated New Drug Application procedure in the USA) considering that all information of the clinical documentation submitted for the reference product may also apply for a generic alternative with confirmed BE.

Immediate-Release Versus Modified-Release Oral Products

Conditions and requirements for BE assessment deviate considerably between immediate-release and modified-release medicinal products. Consequently, separate guidelines have been edited for both types of products in the majority of jurisdictions.

Immediate-Release Products

This group of products includes solid oral dosage forms like tablets or capsules but also some other preparations, e.g., effervescent tablets, granules, as well as oral dispersible forms.

In these cases, BE assessment based on the PK characteristics AUC (for total exposure) and C_{max} (for peak exposure) is generally considered sufficient to describe the products' BA and to conclude on BE. In most cases studies under one condition – in the more discriminative fasted state (FDA) or considering the suggested conditions for the

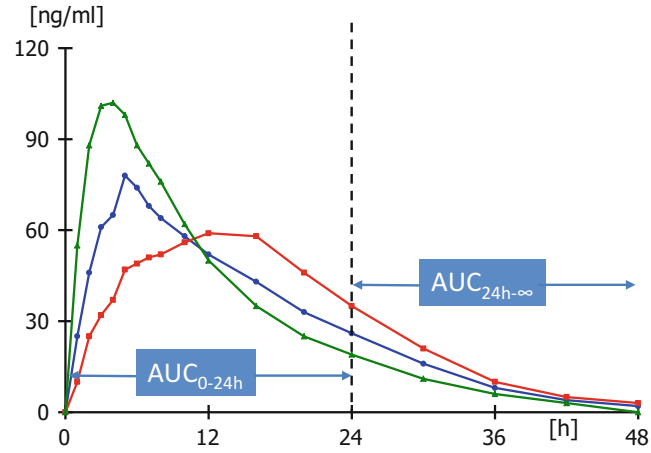


Fig. 6 Calculation of partial AUCs for better description of product's profiles according to the CHMP guideline for MR products

reference product (EU) – may suffice; however, fasted and fed administration are requested with increasing frequency.

In certain cases, in vivo studies may be replaced by in vitro bioequivalence assessments as described more in detail in section "BCS-Based Biowaiver Concept."

Modified-Release Products

This group of medicinal products is more complex, and thus, the biopharmaceutical properties of the products require a more comprehensive characterization. Thus, consideration of AUC and C_{max} as exposure parameters is not sufficient – despite the regulatory practice for decades.

Prolonged- or sustained-release preparations, delayed-release (e.g., enteric-coated) forms, and – more recently – also products with "pulsatile" or "multiphasic" release are allocated to this group, and often the terminology used in this context is not precise, and thus, deviating names may be used for the same type of products.

Considering the more complex nature of these preparations and their biopharmaceutical properties, additional pharmacokinetic parameters need to be included in order to describe, e.g., a "plateau-like" or "biphasic" profile. The most appropriate procedure should be defined case by case. A way suggested by the revised CHMP guideline (2010) is the calculation of partial AUCs in order to describe the course of the profile more properly and allow adequate BE comparison. Figure 6 describes such an example:

Moreover, European authorities also request for conduct of multiple-dose studies in case of such MR products where an accumulation is "likely" during maintenance treatment (thus after multiple dosing). This may be assumed if less than 90% of the total AUC after single-dose administration is achieved during the intended dosing interval (Fig. 7).

The BE requirements of MR products deviate, thus, considerably between EMA and FDA. While FDA principally applies same criteria as for IR preparations, EMA suggests additional parameters for single-dose studies (i.e., partial AUC in order to compare the shape of profiles more appropriately) and requests for investigations at steady state in case of products with certain likelihood for accumulation. The requirement of conducting fed state studies in case of all MR forms is identical in both regions.

Special Application Forms

The principles developed for BE assessment of orally administered systemic products can generally also be applied to non-oral formulations and under certain conditions even to non-systemic products.

Non-oral Systemic Drug Products

Considering that the active ingredient needs to be absorbed into the central circulation in case of all systemic drugs, conclusions on bioequivalence

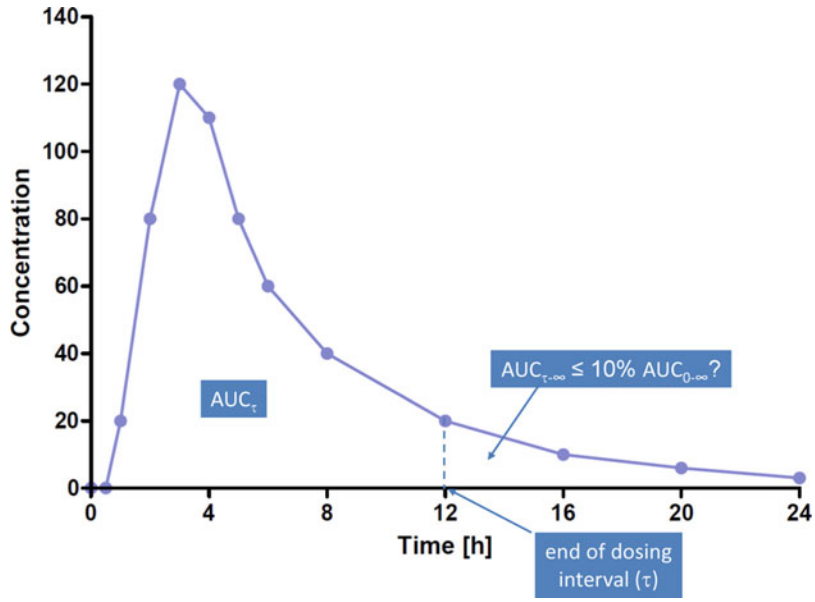


Fig. 7 MR products: evaluation of likelihood of accumulation ($AUC_{0-\tau}$ covers less than 90% of $AUC_{0-\infty}$)

should also be established in case of non-oral forms based on PK parameters derived from plasma concentration-time profiles. Thus, in this context there is no relevant difference between oral and non-oral preparations. However, the specificities of site and way of administration via the non-oral route need to be carefully taken into account.

Topical (Non-systemic) Drug Products

The situation is more complicated in case of non-systemic ("topical") drugs. In certain cases concentrations in plasma remain on such low level that reliable measurement and profiling is not possible, even with advanced analytical techniques. On the other hand, conclusions from the exposure in plasma to the concentration-time courses at the site of action should also be possible in certain situations. These, however, need to be defined and justified case by case.

Generally, the following paradigms have been developed for non-systemic drugs:

- **Plasma concentrations as safety measure**: Comparison of plasma concentrations measurable after administration of the generic alternative and the innovator product is considered an appropriate option to conclude on potential safety differences between both forms.
- **Assessment of therapeutic equivalence**: As long as the bioequivalence approach is not applicable (with convincing justification), clinical studies in patient populations will be needed in order to confirm therapeutic equivalence of both products.

Finally it will be a case-by-case decision how therapeutic equivalence can be appropriately confirmed in case of non-systemic drugs.

Special Challenges and Specificities in Bioequivalence Assessment

In some cases BE assessments require special approaches, e.g., if certain particularities of a reference product hamper the confirmation of bioequivalence of generic medicinal products.

Highly Variable Drugs

High variability of pharmacokinetic parameters makes BE assessment difficult and requires special measures in most cases. Sources of variability can be drug substance related but may also be caused by unfavorable properties of the drug product. Often the application of carefully controlled and consistent conditions in the clinical conduct and the bioanalytical determinations is not sufficient, and inclusion of a considerably higher number of subjects can be unavoidable.

Such increase in the number of volunteers can, however, approach or even exceed limits where reliable clinical conduct cannot be guaranteed any more. Special features have been developed for such exceptional situations, in particular scaling procedures considering the high within-subject variability determined for the reference product in the same study. This approach is only applicable to those drugs with a within-subject variability above 30%. Moreover, scaling is limited to AUC in Europe, while it is also applicable to C_{max} in the USA.

Details for study design and evaluation are described in the relevant guidelines.

Narrow Therapeutic Index (NTI) Drugs

There is a long and controversial international discussion on the question whether more restrictive BE acceptance criteria (e.g., 90.00–111.00%) need to be applied for NTI drugs.

On the other hand, there are strong arguments suggesting that especially the NTI drugs will tend to exhibit relatively low variabilities as otherwise safe treatment would not be possible with such drugs. Improvement of constant dosing from tablet to tablet (content uniformity) and stability of the given dose confirmed in smaller acceptance ranges until end of shelf life is often considered a more relevant criterion than tightening the BE acceptance range. This was, for example, the outcome of the discussion on bioequivalence in case of levothyroxine.

Hierarchy in BE Assessment Procedures

In general there are various alternatives for the investigation of bioequivalence or therapeutic equivalence, i.e., in vitro testing, PK comparison, and pharmacodynamic or clinical studies. There is, however, a clear hierarchy of these options with absolute priority for BE assessment by means of PK measurements or the BCS-based biowaiver approach, if applicable.

On the other hand, pharmacodynamic or clinical investigations are only acceptable in such rare cases where PK measurements are not reliably possible, and thus, generally preferred approaches cannot be realized.

BCS-Based Biowaiver Concept

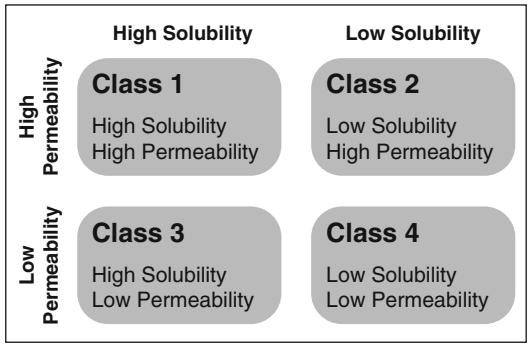
In certain cases the conclusion on bioequivalence can be established on comparative in vitro dissolution testing, and thus, in vivo clinical studies may be waived.

History and General Concept

For this purpose a Biopharmaceutics Classification System (BCS) has been developed in the 1990s (Amidon et al. 1995) and subsequently introduced into regulatory guidelines, first in the USA and later also in Europe.

The development of the BCS concept was initiated by a general discussion on the dissolution properties of solid oral dosage forms and their predictive value for the in vivo performance at the BIO-international conference 1994 in Munich, Germany. It was hypothesized that the systemic exposure should approach that of an oral solution and thus be self-evident in cases of immediate-release forms which dissolve their active ingredient(s) rapidly and completely during gastric residence. In such cases the active ingredient is emptied in solution from the stomach to the intestine where absorption will occur. Consequently, in vivo BE assessments might be waived for such products.

This concept was considered applicable for highly soluble and highly permeable compounds in rapidly dissolving drug products. Consequently drug substances were classified according to their biopharmaceutical properties as follows:



The conditions for allocation of compounds to these BCS classes as well as the dissolution requirements were defined considering the physiological environment in the GI tract:

- **Solubility**: "High solubility" is concluded if the highest dose strength (EU: highest dose) is completely soluble in 250 mL (considering that tablets should be taken with a full glass of water) of aqueous media of pH 1.2 (representing the fasted stomach), 4.5 (representing the fed stomach), and 6.8 (representing the small intestine).
- **Permeability**: Considering that highly permeable drugs should be completely absorbed, compounds with measured extent of absorption exceeding 85% are classified as "high permeability" drugs.
- **Dissolution**: Dissolution is considered as "complete" if minimum 85% of labeled content are in solution and as "rapid" if this limit is achieved within 30 min, which is assumed as mean residence time of tablets in the stomach. Experiments should be performed in buffered solutions of pH 1.2, 4.5, and 6.8 (see above).

For a BCS-based biowaiver application, the dissolution properties of the test formulation (e.g., a generic development product) and the reference product (e.g., the innovator product) need to be compared and their "similarity" confirmed. This is the case if dissolution is complete (>85%) within 15 min, while in cases of slower drug release, dissolution profiles should be compared statistically, e.g., by use of the f2 equation (details defined in the guidelines).

Extension of BCS-Based Biowaiver Applications

Already during the development of the BCS concept, it was argued that class III drugs (high solubility, low permeability) should be even better candidates (than class I) for a BCS-based biowaiver (Blume and Schug 1999). Consequently, an extension to class III compounds was suggested. However, it took more than 10 years to implement this proposal into the current guidelines, first in the EU (2010) and later in the USA (2015).

Nowadays BCS-based biowaivers are applicable to BCS class I and class III compounds in rapidly dissolving products. In this context the guidelines define more restrictive dissolution requirements for class III (>85%/15 min) than for class I drugs (>85%/30 min) in order to be on the safe side with this extension. There are, however, findings from simulation experiments indicating that the risk of a false-positive conclusion on bioequivalence based on the BCS concept is considerably lower in case of class III than in class I compounds.

On the other hand, compounds with incomplete – and often site-dependent – absorption may be more sensitive to interactions with excipients which might impact GI transit, intestinal absorption, and metabolism in the gut wall. Thus, the potential influence of differences in excipients used in both investigational products (test and reference) needs to be carefully considered and justified.

Regulatory Requirements and Need for Harmonization

The scientific basis for regulatory requirements in bioequivalence has continuously been developed and further optimized during the last four decades. On the other hand, deviations between the regions – e.g., Europe, North America, and Japan – still exist in several details, and this makes global development of medicinal products, especially generics, difficult. As consequence more

than one BE study may be necessary to fulfill the regional requirements.

With the intention to fill this gap and to harmonize the divergent regulations, several initiatives have been started, in particular the series of BIO-international conferences in the 1990s and currently the EUFEPS Global Bioequivalence Harmonization Initiative (GBHI).

The BIO-International Conferences: Toward Science-Driven Regulations

Primary intention of the BIO-international conferences was originally to discuss open issues in bioequivalence and to support the development of science-driven regulations. Scientist from academia and industry contributed to this process and important progress was achieved in exchange with experts from regulatory authorities. Examples for essential advancements were, among many other detailed achievements, the development of the Biopharmaceutics Classification System or the scaling procedure for the investigation of highly variable drugs.

Results and conclusions of the discussions have been summarized in conference reports (McGilveray et al. 1990; Blume and Midha 1993; Blume et al. 1995; Midha et al. 1996, 2005) and, even more important, were incorporated in new or revised guidelines. Considering that this implementation was not identical in all cases, discussions on harmonization were started more recently in smaller BIO-international conferences held in the 2000s in London.

The EUFEPS Global Bioequivalence Harmonization Initiative

A more systematic approach in harmonization of BE requirements was started by EUFEPS, the European Federation for Pharmaceutical Sciences, with its network on biopharmaceutics and bioavailability. Essential for the success of this Global Bioequivalence Harmonization Initiative is that key regulatory scientists from EMA as well as the US-FDA were prepared to support this process significantly from the very beginning.

This harmonization initiative is structured by means of specific conferences which take place every 18 months alternatively in Amsterdam, the Netherlands, or Rockville, USA. Meanwhile three conferences have been held in March 2015, September 2016, and April 2018. Results and conclusions from the discussions are summarized in conference reports (Chen et al. 2018, 2019; The global bioequivalence harmonization initiative: report of EUFEPS/AAPS third conference Manuscript in preparation) in order to give the scientific community the chance to further contribute to those issues still open even after the intensive debate during the meetings.

It is desirable that these activities will continue further on and will achieve essential success in harmonization. For this purpose significant contribution by all relevant jurisdictions globally is aspired. This initiative might also be supportive for the ICH process which obviously plans taking up BE issues as well, e.g., as first topic the BCS-based biowaiver concept.

Conclusions and Future Perspective

Concepts and requirements for BE assessment have been developed and further optimized during the previous decades. Relevant open issues have been identified and systematically resolved. Meanwhile, most of the international guidelines include appropriate requirements and suggestions for the majority of relevant BE issues. Nonetheless, the scientific community should continue their effort in contributing to further improvement of the current regulations.

Therapeutic Equivalence and Interchangeability

For the clinical use of medicinal products approved based on BE assessments, in particular generic alternatives, the question of their proven

therapeutic equivalence in comparison to the innovator product is of essential relevance.

Considering that assessment of BE is generally understood as confirmation of therapeutic equivalence, all generic products may be declared as therapeutic alternatives of the approved reference products. On the other hand, comparisons between the generic products are normally missing, and thus, therapeutic equivalence between all of them may be questionable.

In this context a discussion on "individual bioequivalence" in the 1990s in the USA may be reflected which concluded that BE assessment confirmed the "prescribability" of the generic products, while their "interchangeability" has not been reliably demonstrated. The therapeutic experience with generic substitution in all major jurisdictions did, however, not discover major clinical problems. Nonetheless, every exchange of products during maintenance treatment should be handled with care.

Approval Policy and Reimbursement Decision

In most jurisdictions different committees are responsible for the decisions on marketing authorization and reimbursement. Therefore, marketing authorization for medicinal products does not automatically include reimbursement. In increasing number of countries, the latter depends on the outcome of price negotiations between pharmaceutical industry and health insurance companies. In other countries, e.g., Germany, specific fixed-price limits have been defined for the reimbursement of generic medicinal products. These fixed-price limits are set case by case on a compound- and dose-strength-specific basis by a national reimbursement committee. This system, at the same time, supports the general understanding in the public that all medicinal products listed in such a fixed-price group are interchangeable and, thus, may be used for generic substitution.

References and Further Reading

Amidon GL, Lennernäs H, Shah VP, Crison JR (1995) A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm Res 12:413–420

Blume H, Midha KK (1993) Bio-international '92, conference on bioavailability/bioequivalence and pharmacokinetic studies. J Pharm Sci 82:1186–1189

Blume HH, Schug BS (1999) Class III drugs – better candidates for BA/BE waiver. Eur J Pharm Sci 9:117–121

Blume HH, McGilveray IJ, Midha KK (1995) BIO-international '94, conference report on bioavailability, bioequivalence and pharmacokinetic studies. Eur J Drug Metab Pharmacokinet 20:3–13

Chen M-L, Blume H, Beuerle G, Davit B, Mehta M, Potthast H, Schug B, Tsang YC, Wedemeyer R-S, Weitschies W, Welink J (2018) The global bioequivalence harmonization initiative: summary report for EUFEPS international conference. Eur J Pharm Sci 111:153–157

Chen M-L, Blume H, Beuerle G, Mehta M, Potthast H, Brandt A, Schug B, Ducharme M, Endrenyi L, Gallicano K, Schuirmann D, Welink J (2019) The global bioequivalence harmonization initiative: report of EUFEPS/AAPS second conference. Eur J Pharm Sci 127: 24–28

Karim A (1986) Effects of food on the bioavailability from controlled-release products in adults. J Allergy Clin Immunol 78:695–70

McGilveray IJ, Midha KK, Skelly JP, Dighe S, Doluisio JT, French IW, Karim A, Burford R (1990) Consensus report from bio international '89: issues in the evaluation of bioavailability data. J Pharm Sci 79:945–946

Midha KK, Nagai T, Blume HH, Hubbard JW, McGilveray IJ, Williams R (1996) Conference report of bio-international '96, conference on bioavailability, bioequivalence and pharmacokinetic studies

Midha KK, Shah VP, Singh GJP, Patnaik R (2005) Conference report: bio-international 2005. J Pharm Sci 96:747–754

Schug BS, Wolf D, Nilius R, Martin W, Schall R, Blume HH (2000) Product related food effects for nifedipine once daily solid oral dosage forms. Poster at the APVWorld Conference, Berlin

The global bioequivalence harmonization initiative: report of EUFEPS/AAPS third conference. Manuscript in preparation

Population Pharmacokinetics and Pharmacokinetic-Pharmacodynamics in Clinical Pharmacology

50

Daniel F. B. Wright, Chihiro Hasegawa, and Hesham S. Al-Sallami

Contents

D. F. B. Wright (✉) · H. S. Al-Sallami
School of Pharmacy, University of Otago, Dunedin, New Zealand
e-mail: dan.wright@otago.ac.nz; hesham.al-sallami@otago.ac.nz

C. Hasegawa
School of Pharmacy, University of Otago, Dunedin, New Zealand

Translational Medicine Center, Ono Pharmaceutical Co., Ltd., Osaka, Japan
e-mail: chihiro.hasegawa@otago.ac.nz

Abstract

Clinical pharmacology is a broad professional and scientific discipline concerned with all aspects of drug use in humans. One of the primary goals of this field is to improve health outcomes by supporting the development, rational use, and safety of medicines. Clinical pharmacology and pharmacometrics are closely related and share common goals and research

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_18

themes. Notable amongst these are pharmacokinetics and pharmacodynamics. In the pharmaceutical industry, population pharmacokinetic and pharmacokinetic-pharmacodynamic studies aid dose selection, assess links between drug exposure and efficacy and safety metrics, and inform the dosing information that will be presented on the drug label. In the clinical environment, population pharmacokinetic and pharmacokinetic-pharmacodynamic studies are conducted to aid dose optimization for an individual patient. The aim of this chapter is to present an overview of population pharmacokinetic and pharmacokinetic-pharmacodynamic concepts and methodology as they apply in the industrial and clinical setting. The chapter is divided into four parts: Part 1 will provide a board overview of general concepts and definitions related to population pharmacokinetic and pharmacokinetic-pharmacodynamic analyses; Part 2 will look at commonly used models; Part 3 will explore methodology; primarily nonlinear mixed effects modeling, and Part 4 will present examples of pharmacokinetic and pharmackinetic-pharmacodynamic analyses, presented in the style that is typical for a regulatory submission involving phase I data.

In the pharmaceutical industry, population pharmacokinetic (PK) and pharmacokinetic-pharmacodynamic (PKPD) studies play an important role in the analysis of Phase 1, 2, and 3 data, and to a lesser extent, data that arises from preclinical and post-marketing trials. The goals of PK and PKPD analyses align with those of the broader clinical pharmacology team, i.e., to aid decisions about dose selection for the prospective drug, to assess efficacy and safety data, and ultimately, to inform the dosing information that will be presented on the drug label. Post-marketing, in the clinical setting, PK and PKPD are often viewed as an applied field within the medical discipline of clinical pharmacology, focused primarily on optimizing drug therapy for individual patients. In both cases, the types of analyses conducted may extend beyond population PK and PKPD to include exposure-response (ER) analyses, quantitative systems pharmacology modeling (QSP), physiological-pharmacokinetic (PBPK) modeling, and optimal design work, amongst others. These approaches are bound together by the application of mathematical and statistical methodologies, referred to collectively as "pharmacometrics."

In this chapter, we assume that pharmacometric analyses performed in the industrial or clinical setting will share broadly similar goals, particularly with regards to optimizing drug doses. The only assumed difference is that the research conducted in the drug industry will tend to focus on the optimization of dose to the level of regulatory requirement (e.g., dosing guidance for the drug label) while in the clinical environment, the focus is usually dose optimization for an individual patient. The aim of this chapter is to present an overview of population pharmacokinetic and pharmacokinetic-pharmacodynamic concepts and methodology as they apply in the industrial or clinical setting. The chapter is divided into four parts: Part 1 will provide a board overview of general concepts and definitions related to population PK and PKPD analyses, Part 2 will look at commonly used PK and PKPD models, Part 3 will explore methodology, primarily nonlinear mixed effects modeling, and Part 4 will present an example of PK and PKPD analyses, presented in the style that might be expected for a regulatory submission involving phase I data.

Part 1: General Concepts and Definitions

Clinical pharmacology is a broad professional and scientific discipline concerned with all aspects of drug use in humans. One of the primary goals of this field is to improve health outcomes by supporting the development, efficacious use, and safety of medicines. While a diverse discipline, this chapter will focus primarily on two central themes in clinical pharmacology research, pharmacokinetics and pharmacodynamics. Here pharmacokinetics is defined as the time course of drug concentrations in the body, a science focused on the relationship between the drug dose and exposure (Fig. 1a). The generic term "drug exposure" will be used to refer to drug concentration at any time point, i.e., $C(t)$. Note that it is convenient and

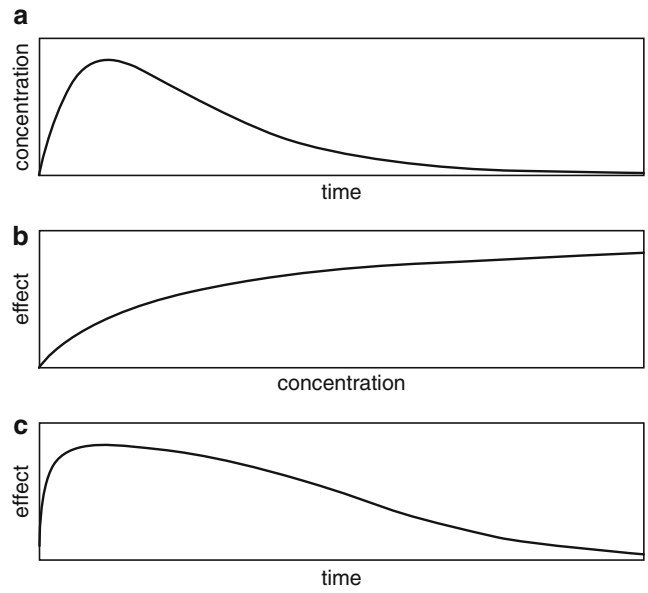


Fig. 1 Conceptual framework for PK, PD, and PKPD data (Adapted from Wright et al. 2011, used with permission from John Wiley and Sons licence number 4277280083455)

common to summarize $C(t)$ as a time-invariant measure such as the maximum, minimum, or steady-state average plasma concentration post dose ($C_{p_{max}}$, $C_{p_{min}}$, and $C_{p_{ss,ave}}$, respectively), as well as the area under the plasma concentration time curve (AUC). Pharmacodynamics is concerned with the relationship between drug exposure and response (Fig. 1b) and is independent of time. Pharmacokinetics-pharmacodynamics (PKPD) is a combination of PK and PD that allows us to explore the time course and magnitude of drug response (Fig. 1c).

In this chapter, we will consider pharmacodynamic responses that are quantified by measuring a biomarker or surrogate (e.g., blood pressure, prothrombin time to measure anticoagulation), rather than dichotomous measures of drug effect (e.g., seizure or no seizure) or other clinical endpoints (e.g., death, hospitalization, etc.). A definition of a pharmacodynamic biomarker will be presented below.

Pharmacometrics

Pharmacometrics is concerned with the analysis and interpretation of data that arises from drug studies. The discipline can be traced to the seminal work of Louis Sheiner (Sheiner 1977) and, in particular, his collaboration with Stuart Beal to develop nonlinear mixed effects modeling methodology for drug studies (Sheiner and Beal 1980). Pharmacometrics involves the use of mathematical and statistical models to predict drug exposure, physiological response, and clinical outcomes, and to describe the variability in these measures between (and within) individuals. As such, pharmacometric analyses play an important role in drug development as well as clinical practice. The terms "population analysis," "population approach," and "population modeling" are also commonly used to refer to pharmacometric analyses.

Optimizing the Dose

> All things are poison and not without poison; only the dose makes a thing not a poison
> Paracelsus (physician and botanist, 1493–1541)

The notion that a relationship exists between the amount of drug given to an individual and the intensity of the resulting response (desired or adverse) is intuitive and has likely been understood since antiquity. It was certainly recognized by Paracelsus in the sixteenth century who noted, in the above quote, that a poison and a therapeutic agent differ only in the dose that is administered.

However, it was not until the last century that it was possible to quantify the relationship between drug dose and physiological response.

Research intended to inform dose selection or optimization, whether to support drug labeling decisions for regulatory submission or to individualize therapy in the clinic, requires a quantitative understanding of drug exposure and physiological response, and how these vary between and within individuals. The optimal dose is expected to have the highest probability of achieving a desired physiological response in an individual patient or a population of patients while carrying a minimal risk of adverse effects. The general concept is that the drug will achieve the desired physiological response in most people once a threshold exposure (or steady-state plasma concentration) has been reached. Increasing the exposure may increase the magnitude of the response in some cases, depending on the shape of the exposure-response curve (see Fig. 3), but may also increase the risk of adverse effects and toxicity. This concept underpins the clinical practice of therapeutic drug monitoring, where plasma concentrations are measured and doses adjusted to achieve a specified target range. TDM is particularly useful for drugs with a narrow therapeutic range, i.e., where the desired physiological response and toxicity can occur within a narrow exposure range, and where the physiological response is difficult to measure, i.e., cases where the clinical endpoint is to prevent an adverse event (e.g., seizure).

Biomarkers for Physiological Response

The ideal measure of drug effectiveness is the consistent achievement of the clinical outcome of interest in the intended patient population. However, when the outcome is the long-term prevention of an adverse clinical event (e.g., stroke with anticoagulant therapy or cardiovascular disease with lipid-lowering drugs) or requires long-term observation (e.g., the cessation of gouty attacks with urate-lowering therapy), a biomarker for the physiological response is required.

A biomarker is an indicator of the biological, pathological, or pharmacological response to drug therapy (Biomarkers Definitions Working Group 2001). The biomarker may be on the causal pathway between the disease and the clinical outcome or may be correlated in some way with the outcome. The drug may alter the underlying disease which in turn will alter the biomarker or the drug may act directly on the biomarker itself which, in turn, will be correlated with the clinical outcome (e.g., symptomatic treatment).

The term biomarker can be distinguished from "biomeasure" and "surrogate." A biomeasure is any physiological measurement that can be obtained clinically (e.g., blood pressure). A biomarker is therefore a special case of a biomeasure. A surrogate is a biomarker that is intended to substitute for a clinical outcome. A surrogate is usually approved by a regulatory agency and can be used in the drug approval process.

Part 2: Pharmacokinetic and Pharmacokinetic-Pharmacodynamic Models

What Is a Model?

> Essentially, all models are wrong, but some are useful
>
> George E. P. Box (statistician, 1918–2013)

A model is a construct that allow us to simplify reality. A toy car is a trite example. It is small enough to fit into a child's hand yet, to a child, it retains all of the essential features of the real thing such as the right body shape, wheels, and perhaps a flashy paint job. Similarly, a pharmacometric model is a mathematical representation of the complex interaction between a drug and a biological system. It is used to describe the relationship between input variables, such as drug dose and time, and output variables including plasma drug concentration and drug response.

For a pharmacometric model to be useful, it must be simple enough for practical use yet retain the essential mathematical features that allow us to understand the relationship between input (e.g.,

drug dose, time) and output (e.g., plasma drug concentrations, drug response). This concept is embodied in the often-used quote above from George Box. All models are *wrong* because a model cannot recreate reality; however, a model can serve a *useful* purpose if it retains a close relationship with important aspects of reality.

The term "population model" is commonly used in pharmacometrics and may seem misleading. In the context of a pharmacometric analysis, "population" refers to the group (population) of individuals who are the intended recipients of the medicine. This may also include healthy volunteers who may be recruited, for example, in a Phase I pharmacokinetic study. A population model will, therefore, provide typical drug exposure and response information, estimates of random between subject variance, predictable variance in drug exposure or response associated with measurable patient factors (e.g., weight), and random within subject variance. Hence, the information contributed by an individual patient is retained and contributes to the overall understanding of drug behavior.

Pharmacokinetic Models

Pharmacokinetics is the science that relates the dose administered to the time course of measured drug concentrations in the body (usually the plasma) and therefore provides information about drug exposure. A PK model can predict the typical time course of drug concentration $C(t)$ as a function of the administered dose (D) over time (t) and is dependent on unknown pharmacokinetic parameters (θ_{pk}). The primary PK parameters of interest are clearance (CL), the apparent volume of distribution (V), the absorption rate constant (k_a) (for extravascular administration), systemic availability (F) (for extravascular administration), and the secondary parameter elimination rate constant (k_e), which is given by CL/V. CL is a constant that relates the rate of elimination to the measured drug concentration and is related to the functional capacity of the body. V is the apparent volume into which the drug distributes and is related to body composition. F is the fraction of drug that reaches the systemic circulation.

A PK model can be constructed using a compartmental structure with an input model and a disposition model (Fig. 2). The input model describes the time course of drug movement from the site of administration (e.g., the gut) to the site of drug measurement (e.g., the plasma). The disposition model describes the time course of drug distribution, metabolism, and elimination from the body and can be depicted as single or multiple compartments. Input, distribution, and elimination occur simultaneously and the relative time course of each will determine the pharmacokinetic behavior of the drug.

Figure 2 depicts one and two compartment PK models for an orally administered drug. The body is represented as a series of discrete units into which the drug distributes and from which drug elimination occurs. The compartments do not represent true physiological spaces, and it is assumed that drug behavior is similar within each compartment. In a one compartment model (Fig. 2a), the drug is assumed to distribute evenly throughout a single compartment. In a two compartment model (Fig. 2b), the drug distributes into an additional peripheral compartment.

For a one compartment extravascular administration model, the concentration of drug in the plasma at any time (after a single dose) is given by

Equation 1 *The equation for a one compartment PK model with first-order absorption and elimination*

$$C(t) = \frac{D \times F \times k_a}{V \times (k_a - k_e)} \times \left(exp^{(-k_e \times t)} - exp^{(-k_a \times t)} \right)$$

By convention the number of compartments in a PK model is defined by the number of exponential terms needed to describe the disposition of the data. Therefore, the model above, while having two exponential terms, has only one term related to disposition and hence is considered a one compartment model.

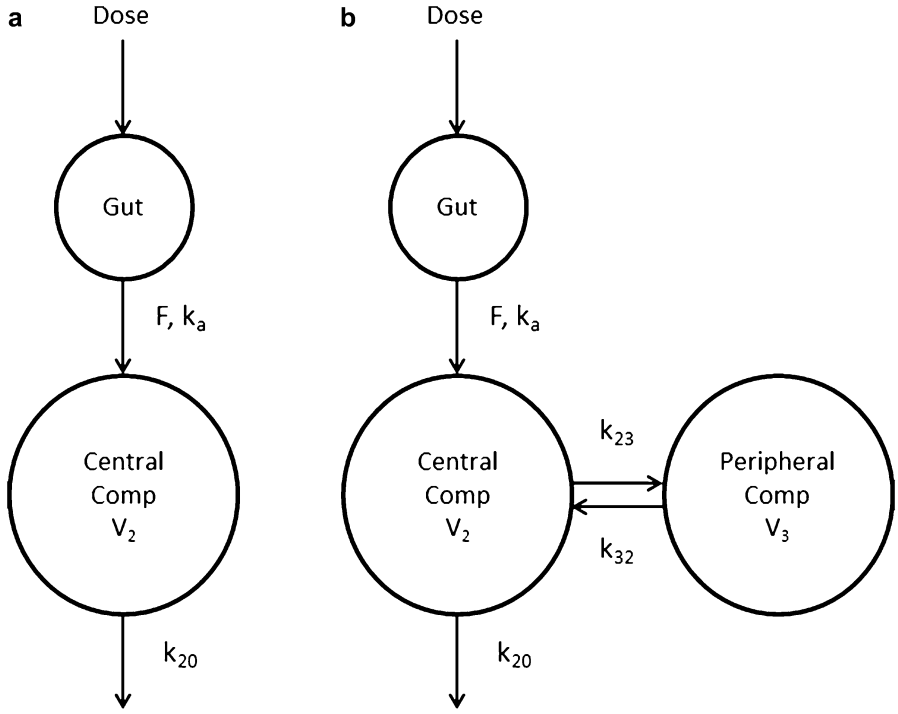


Fig. 2 Schematic of a one compartment PK model (**a**) and a two compartment PK model (**b**) for an orally administered drug with first-order absorption and elimination

Pharmacodynamic Models

Pharmacodynamics relates drug concentrations to the observed pharmacological response. Since *in vitro* pharmacodynamic experiments are often conducted in equilibrium conditions, the relationship between concentration and response is typically independent of time. A PD model can predict the typical response (denoted E for effect) of a drug as a function of the drug concentration (C) and unknown pharmacodynamic parameters (θ_{pd}). Note that we use the terms 'drug effect' and 'drug response' interchangeably in this chapter.

The relationship between drug concentration and response is usually characterized by a hyperbolic function, which yields a nonlinear relationship between effect and concentration. Hence, doubling the concentration (or dose) will not necessarily result in a doubling of drug response and the drug response will asymptote to a maximum despite increased concentrations.

Much of the underpinning theory, and the models used to describe the exposure-response relationship, is based on receptor-binding theory. However, in pharmacometrics, we are primarily concerned with the magnitude of drug response that will result from a given dose or exposure. . Therefore, a model that will predict drug response, not drug binding is required. The relationship between receptor occupancy and drug effect can be described using a proportionality constant known as 'intrinsic activity', which can range from 1 for a full agonist to 0 for an antagonist. This enables the relationship between concentration and effect to be expressed using an empirical version of the E_{max} model:

Equation 2 *The E_{max} model*

$$E = E_{max} \times \frac{C}{C_{50} + C}$$

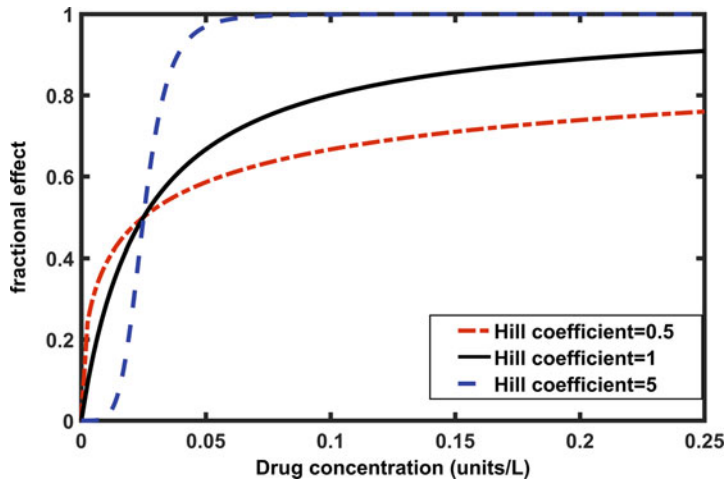


Fig. 3 The E_{max} model showing the change in shape with different Hill coefficients, which will dictate the shape of the curve. In this model $E_{max} = 1$, $C_{50} =$ 0.025 units/L

where E_{max} is the maximum effect of the drug, C is the drug concentration, and C_{50} is the drug concentration which results in 50% of maximum response.

The E_{max} model has been widely applied in pharmacodynamics and is the basis for many models in pharmacometrics. There are two parameters of interest: E_{max} and C_{50}. A generalization of the E_{max} model, called the sigmoidal E_{max}, includes an empirical exponent termed the Hill coefficient (λ) which changes the shape of the exposure-response curve (Fig. 3):

Equation 3 *The sigmoidal E_{max}model*

$$E = E_{max} \times \frac{C^{\lambda}}{C_{50}{}^{\lambda} + C^{\lambda}}$$

Values of λ greater than 1 produce a steep exposure-response curve and predict that relatively small changes in drug concentration will produce a rapid change in effect which, at the extreme, can be observed as an "on-off" phenomenon (e.g., anti-arrhythmic agents). Values of λ less than 1 produce a shallow concentration-response curve where response increases rapidly at low concentrations but approaches the E_{max} asymptote slowly at high concentrations (Fig. 3).

The E_{max} model can be further modified to include baseline physiological status (S_0) in the absence of drug. This model can be used to predict the change in this status (e.g., blood pressure) after drug administration.

Equation 4 *The E_{max}model with a constant baseline status* (S_0)

$$E = S_0 + E_{max} \times \frac{C}{C_{50} + C}$$

Pharmacokinetic-Pharmacodynamic Models

A PKPD model can predict the typical response (E) of a drug as a function of the dose (D) and time (t) dependent on pharmacokinetic (θ_{pk}) and pharmacodynamic parameters (θ_{pd}).

PKPD models conventionally include two basic subtypes:

- Immediate effects model
- Delayed effects model

Immediate Effects PKPD Model

Under nonequilibrium conditions, an immediate effects PKPD model predicts that drug effects will reach a maximum at approximately the same time as the maximum plasma concentration of the drug

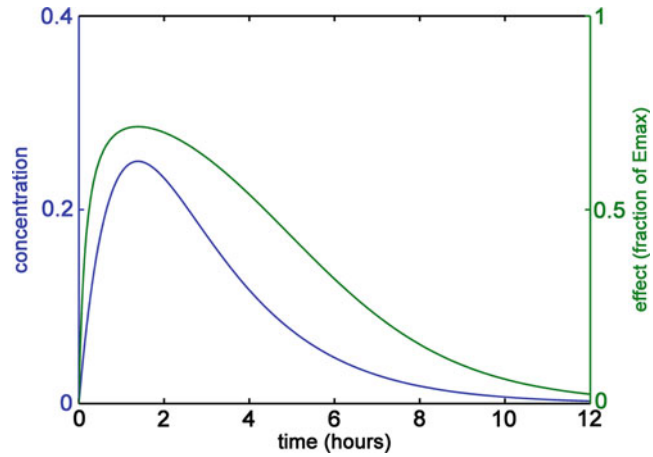


Fig. 4 The time course of drug concentrations (*blue line*) and drug effects (*green line*) for an immediate effects model. In this model $E_{max} = 1$, $C_{50} = 0.1$ units/L. The concentrations were generated using a one compartment model with first-order input. Dose = 1 unit, $CL = 1$ L/hour, $V = 1$ L, and $k_a = 0.5$ h

(see Fig. 4). The PK model provides the time-dependent plasma concentration ($C(t)$) so that the immediate drugs effects model can be described by the Emax model (equation 5).

Equation 5 *Immediate effects PKPD model*

$$E(t) = E_{max} \times \frac{C(t)}{C_{50} + C(t)}$$

By contrast, under equilibrium conditions, it is assumed that the steady state average concentration ($C_{ss,\,ave}$) is sufficient to describe the important pharmacokinetic characteristics of the drug. This is given by

Equation 6 *A model for $C_{ss,ave}$ under equilibrium conditions*

$$C_{ss,\,ave} = \frac{dose\ rate}{CL}$$

where *dose rate* is the maintenance dose (e.g., mg/day). This greatly reduces the complexity of the PK model, but it will only predict the magnitude of drug effects not the onset or duration. The PK model can be substituted into a PD model to provide a steady-state PKPD model:

Equation 7 *Immediate effects PKPD model under equilibrium conditions*

$$E = E_{max} \times \frac{C_{ss,\,ave}}{C_{50} + C_{ss,\,ave}}$$

Delayed Effect PKPD Models

An important limitation of many PKPD models is that the data used to develop the model will usually lack detail concerning underlying PK and PD mechanisms, i.e., they are empirical in nature. In the absence of this mechanistic information it will only be possible to model the rate-limiting step in the time course of drug response. Three mechanisms for delayed effects will be considered as possible rate-limiting processes in the time course of drug effects: drug-receptor binding, distribution to the biophase, and delay related to biological systems (e.g., secondary messengers).

Delay in drug-receptor binding. If the rate-limiting step in the time course of drug effects is driven by the drug-receptor interaction, the onset of drug effects will be related to the dissociation constant, k_{off} (Wright et al. 2011). If the k_{off} value is large, providing a short equilibration half-life, then the drug will be seen to behave as if it has an immediate effect (also termed an "immediate effect" or "direct action").

Delay in distribution to the biophase. For most drugs the site of action is distal to the venous compartment. A model can therefore be constructed to account for the delay in drug distribution to the so-called "effect compartment" or

"biophase." The distinction between the terms "biophase" and "effect compartment" is one of semantics. The biophase is the supposed true site of action. The effect compartment is a theoretical space that appears to have the same distributional properties as the biophase. The effect compartment is therefore not a true model for the biophase but rather a model for the delay in effect due to drug distribution to the biophase. This distribution is essentially a pharmacokinetic phenomenon, but it is described entirely and empirically by the observed delay in pharmacodynamic behavior.

An effect compartment model includes a link function between the PK model and PD model (Holford and Sheiner 1982). If the distribution to the biophase is a first-order process, the delay between peak plasma concentration and peak drug effect will be independent of dose and clearance but dependent only on the rate constant of elimination from the effect compartment. The model assumes that only a small amount of drug distributes into the biophase so that the overall impact on mass-balance is negligible. While the volume of the effects compartment cannot be determined, the rate constant for the loss of drug from the effect compartment can be estimated during the modeling analysis. The full effect compartment model therefore includes a PK model, a link model for the effect compartment, and a simple E_{max} model driven by the effect compartment concentration (C_e):

Equation 8 *Effect compartment PKPD model. $k_{e0}$is the equilibrium rate constant for the effect compartment*

$$\frac{dC_e}{dt} = k_{e0} \times (C(t) - C_e); C_e(t=0) = 0$$

$$E(t) = E_{max} \times \frac{C_e(t)}{C_{50} + C_e(t)}$$

Delay in system response. The mechanism of action for many drugs involves the activation or blockade of a receptor which, in turn, initiates a physiological response mediated by a series of biological processes (e.g., second messengers) (Fig. 5) (Dayneka et al. 1993; Jusko and Ho 1994; Sharma and Jusko 1998). These processes have a time course of their own and often constitute the rate-limiting step in the time course of drug effects. Like effect compartment models, these models are characterized by a delay in the observed effect with respect to the measured plasma

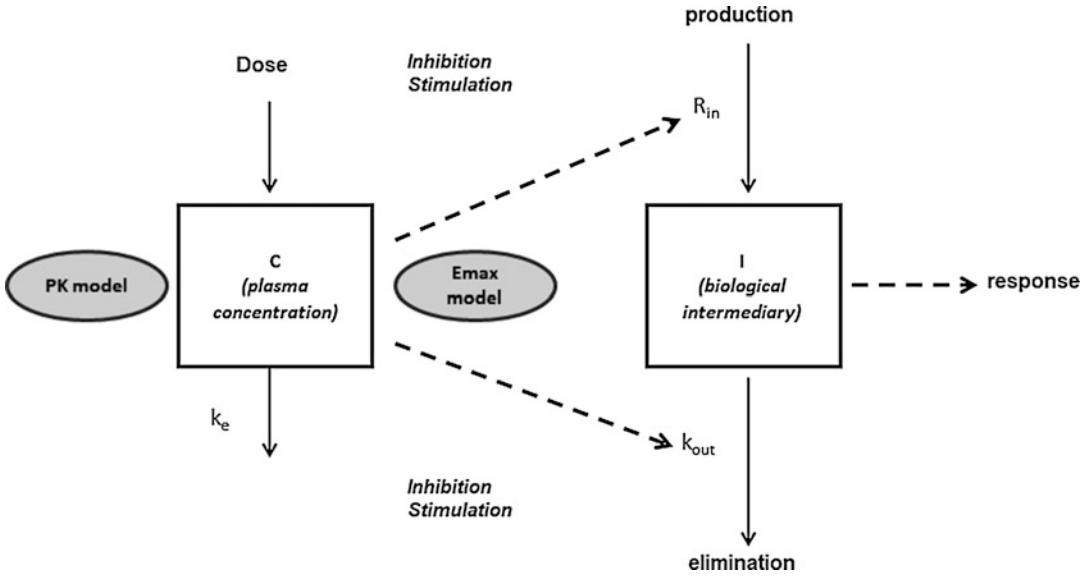


Fig. 5 Schematic of four turnover models including inhibition or stimulation of the intermediary production or inhibition or stimulation of the intermediary elimination (Adapted from Wright et al. 2011, used with permission from John Wiley and Sons licence number 4277311151819)

Table 1 Models for the turnover of a single biological intermediate (I). R_{in} is the zero-order production rate of I, k_{out} is the elimination rate constant for I

	Mechanism	Model
A	Inhibition of input	$\frac{dI}{dt} = R_{in} * \left[1 - E_{max} * \frac{C(t)}{C_{50}+C(t)}\right] - k_{out} * I$
B	Stimulation of input	$\frac{dI}{dt} = R_{in} * \left[1 + E_{max} * \frac{C(t)}{C_{50}+C(t)}\right] - k_{out} * I$
C	Inhibition of loss	$\frac{dI}{dt} = R_{in} - k_{out} * \left[1 - E_{max} * \frac{C(t)}{C_{50}+C(t)}\right] * I$
D	Stimulation of loss	$\frac{dI}{dt} = R_{in} - k_{out} * \left[1 + E_{max} * \frac{C(t)}{C_{50}+C(t)}\right] * I$

concentration and are commonly referred to as "turnover models" (and sometimes "indirect-response models") (Jusko and Ho 1994). An important difference between models based on physiological intermediaries and those based on an effect compartment is that the time of maximum effect for the turnover model will be dose dependent (Fig. 5).

There are four proposed mechanisms for drugs acting on a single biological intermediary (I) resulting in four turnover PKPD models (Dayneka et al. 1993; Jusko and Ho 1994; Sharma and Jusko 1998). These are presented in Table 1, while the typical concentration-effect profiles for each model are shown in Fig. 6.

Selection of the most appropriate PKPD model should be guided by prior knowledge of the drug pharmacology. Data from more than one dose level greatly helps to distinguish between models (immediate or delayed action and, if delayed, an effect compartment or a turnover model, see Fig. 6). For a direct effect model, the time of the peak concentration is also the time of peak effect for all dose levels. The effect compartment model introduces a delay between the peak concentration and peak effect; this could be due to drug transport to the site of action. If this transport is a first-order process (e.g., diffusion), the time delay is independent of dose. All four turnover models exhibit a dose-dependent delay between peak concentration and peak effects.

Statistical Models for Uncertainty in Drug Response Measurement

A population model includes a statistical model to describe the variability between the model predictions and the observations. This is termed residual unexplained variability (RUV) or "uncertainty." It is assumed that uncertainty arises typically from four sources:

1. Process error – where the dose or timing of dose or timing of blood samples are not conducted at the times that they are recorded
2. Measurement error – where the response is not measured exactly due to assay error
3. Model misspecification
4. Moment to moment variability within a patient

In mathematical terms, a model for one individual can be described as:

Equation 9 *The mathematical form of a model with residual error*

$$y_j = f\left(D, x_j, \theta\right) + \varepsilon_j$$

where the j^{th} observation (e.g., drug concentration) for the individual y_j is a function of the administered dose (D) and time, and θ is an n_p-by-1 vector of unknown mean parameters for the individual. In this model, the j^{th} observation deviates from the model prediction by an error, ε_j, which is assumed to be normally distributed with a mean of zero and a variance of σ^2:

Equation 10 *The distribution of* ε_j

$$\varepsilon_j \sim N\left(0, \sigma^2\right)$$

Statistical Models for Variability Between Individuals (Heterogeneity)

An expansion of a model for a single subject to a population of subjects includes an additional

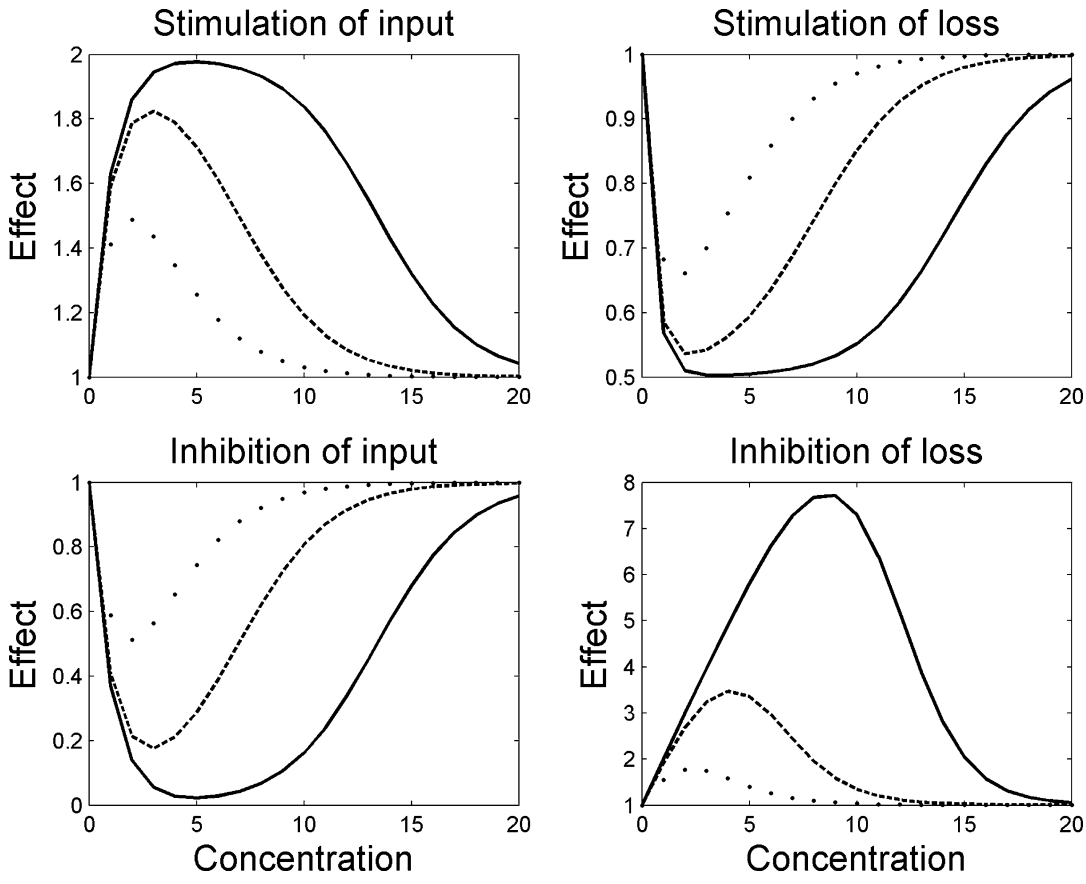


Fig. 6 Concentration-effect profiles using the four turnover PKPD models and three dose levels: 50 mg (*dotted curve*), 500 mg (*dashed curve*), and 2000 mg (*solid curve*). $CL = 10$, $V = 20$, $E_{max} = 1, C_{50} = 1$, $E_0 = 1$

consideration of the variability between people. This yields the well-known population approach model. An important goal of population modeling is to establish the relationship between pharmacokinetic and pharmacodynamics parameters of interest and covariates (i.e., observable patient characteristics such as sex, age, weight, height, organ function indices, and concomitant drugs). Genetic covariates may also predict differences in PKPD parameters for some drugs. Covariates explain some of the variability between individuals and therefore provide the basis for decisions about individualized dosing in clinical practice. In population analyses, variability between individuals is also called heterogeneity, between subject variability (BSV), intra-individual variability (IIV), and population parameter variability (PPV).

In mathematical terms, a population model for repeated measures in a series of individuals can be generalized as:

Equation 11 *The mathematical form of a population model with between-subject variability and residual error*

$$y_{ij} = f\left(D_i, x_{ij}, \theta; \eta_i\right) + \varepsilon_{ij}$$

where the j^{th} observed concentration for the i^{th} individual (y_{ij}) is a function of the administered dose (D_i) and time(x_{ij}) and η_i is an n_p-by-1 vector of the difference between the parameter estimates for the i^{th} individual and the typical values (geometric mean values) for the population. The distribution of η_i for all subjects in the study population is often assumed to be normally

distributed (although transformations can be applied to account for other distributions) with a mean of zero and a variance-covariance given by Ω:

Equation 12 *The distribution of* η_i

$$\eta_i \sim N(0, \Omega)$$

BSV can be described by two distinct models. Firstly, a model is developed to describe predictable reasons why individuals are different (BSVP, where "P" means predictable) and then a model is developed to quantify random variability (BSVR, where "R" means random) (Holford and Buclin 2012). Simply put, BSVP is the variability that can be explained by covariates while BSVR is the remaining random component. By quantifying both BSVP and BSVR in a population it is possible to predict the likely range of PKPD responses that may occur. BSVP can be reduced by accounting for influential covariates (e.g., size, organ function, disease state, or genetics) on parameter estimates. However, random variability in PK parameters across a patient population still remains even after accounting for patient covariates. A recent review showed that the average CV% for clearance (based on 181 population PK studies) was about 40% (IQR 26–48) (Al-Sallami et al. 2014). This corresponds to a fivefold variability in steady state average concentration ($C_{ss,\ ave}$) which clinically necessities a fivefold difference in dose-requirements to achieve a target C_{ss} across the population (Fig. 7). The authors suggest a recalibration of current perception of what constitutes normal PK variability. Traditionally, between-subject variability in PK parameters was considered "low" for CV% $\leq$ 10%, "medium" for CV% of around 25%, and "high" for CV% $>$ 40% (Rowland and Tozer 2011). We propose that a CV of 25–50% for clearance should be considered normal variability for most drugs.

Part 3: Methodology (Nonlinear Mixed Effects Modeling)

The following is a brief overview of nonlinear mixed effects modeling. It is not intended to be an exhaustive review of technical details or of the statistical methodology. For further details, please refer to Bonate 2011, Davidian and Giltinan 1995, and Vonesh and ChinChilli 1997.

Population analyses have traditionally employed three approaches: the naïve pooled approach; the two stage analysis; and the population nonlinear mixed effects modeling. The naïve pooled method, in essence, assumes that all observations arise from a single individual and hence differences between individuals cannot be quantitated (Sheiner and Beal 1980). For the two stage approach, the parameters of interest are estimated for each individual in the data set separately using ordinary least squares or a similar estimation method. The population parameters are then determined by calculating the arithmetic or geometric mean of the parameter values across all of the subjects. While a simple method for estimating population parameter values and BSV, the "two-stage" method requires rich data from each individual and may result in inflated and/or biased estimates of BSV.

In nonlinear mixed effects modeling, the population parameters and variance terms are estimated simultaneously for all individuals. Sparse data and unbalanced sampling designs can potentially be accommodated (although see Siripuram et al. 2017 for limitations in this regard), and both population and individual parameter values can be estimated. The term "mixed effects" refers to the combination of fixed effect parameters (e.g., mean drug CL in a patient population) and random effect parameters (e.g., between-subject variance for CL in a patient population).

Nonlinear mixed effects modeling provides a means of assessing the probability of the data arising from particular structural and variance models, given the parameters θ, Ω, and σ^2. In other words, the parameter estimation will often involve the computation of the likelihood of the observed data arising from the given model. It is usual to search for the best set of parameter values (often maximum likelihood estimates) via iterative algorithms.

There are several nonlinear mixed effects modeling software packages, but NONMEM® is commonly used in pharmacometrics. It was originally developed by Lewis Sheiner and Stuart

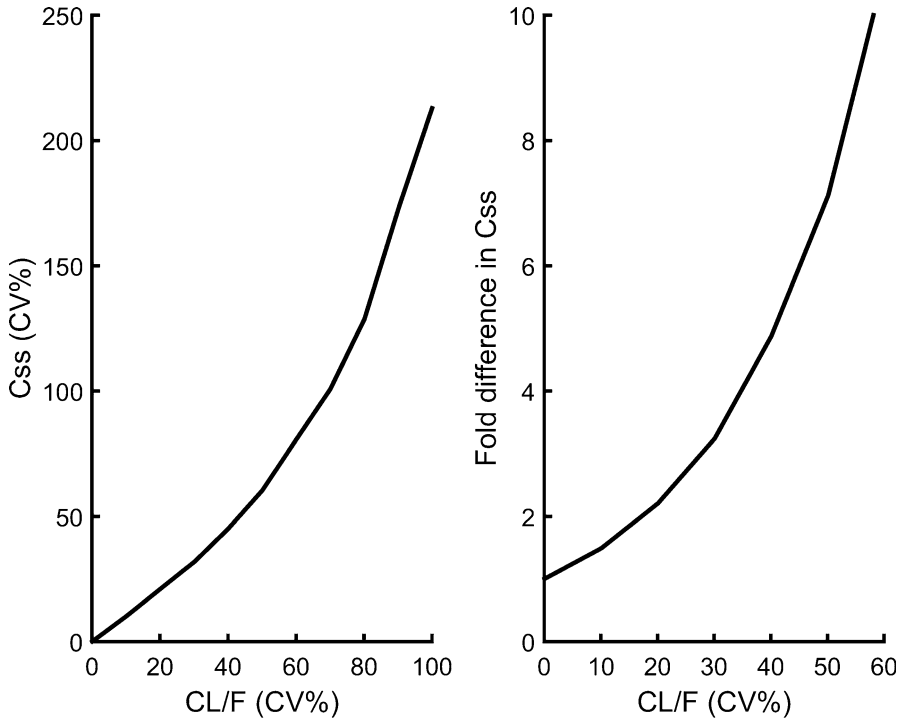


Fig. 7 Illustration of variability and fold-difference in C_{ss} as a result of BSV in CL of a hypothetical drug with PK described by a one compartment model with a CL of 1% and 100% bioavailability. C_{ss} calculated for a dose of 1 unit

Beal (Beal et al. 2011) and was the first software developed specifically for analyzing pharmacokinetic and pharmacodynamic data. Because nonlinear mixed effects models are nonlinear with respect to the random effects parameters, no closed-form solutions are available to solve the integrals for the expectation and variance of the likelihood. This problem has been addressed by using a linearization process during parameter estimation. The first-order (FO) method for approximating the likelihood in NONMEM® was first proposed by Sheiner and Beal and uses a first-order Taylor series approximation around the random effects (evaluated at $\eta_i = 0$). The first-order conditional estimation (FOCE) method uses a similar principle, but the expansion is evaluated at each iteration based on a conditional estimate of η_i, i.e., the empirical Bayes estimates (EBEs) of the BSV. An interaction term can be added (FOCE-I) for heteroscedastic error models to account for interaction between η and ε. The Laplacian approximation method uses a second-order Taylor series around the conditional estimates of η_i (Davidian and Giltinan 1995). For NONMEM the objective function value (OFV) is proportional to minus twice the log likelihood (−2LL).

Once a population model is fitted to the data, the model performance is evaluated based on statistical significance, predictive performance, and biological plausibility. To show statistically whether one model performs better than another, the likelihood ratio test is usually used. As the OFV is proportional to –2LL, and the likelihood ratio is asymptotically and approximately chi-squared distributed, a decrease in OFV between two nested models of $\geq$3.84 points denotes a p-value $<$0.05 with one degree of freedom. The precision of parameter estimates is computed using either the asymptotic standard errors obtained via maximum likelihood estimation or through nonparametric bootstraps. The predictive performance of the model can be assessed through the use of diagnostic plots and visual predictive

checks which can show whether model predictions are biased. The predictive performance can also be assessed by using the model to predict into a new data set. Additionally, parameter values and covariate effects in the final model are checked for biological and/or mechanistic plausibility.

Part 4: Example of Population PK and PKPD Analyses

The following example presents a population PKPD analysis that is similar to the type of work conducted as part of a drug development process in the pharmaceutical industry. The drug in this case is a fictitious example of a lipid-lowering drug simply termed "Drug P." No similarity to any marketed drug is intended. The structure of Part 4 is based on the style that is often presented in population PKPD analysis reports which may be included in regulatory submission documents.

Introduction

Drug P is a lipid-lowering drug that inhibits 3-hydroxy-3-methyglutaryl coenzyme A reductase (HMG-coA reductase) in the liver. HMG reductase is the rate-limiting enzyme in the biosynthesis of low-density lipoprotein (LDL)-cholesterol. Drug-P is inactive but is rapidly converted to an active metabolite, Metabolite-M. The conversion of Drug-P to Metabolite-M is approximately 30%.

This document provides a description of the population pharmacokinetics and pharmacokinetic-pharmacodynamic analyses of Drug-P in healthy volunteers conducted as part of the Phase I clinical pharmacology program.

Objectives

The primary objectives of the population analyses were:

(a) To characterize the plasma pharmacokinetics and pharmacokinetic-pharmacodynamics of Drug-P and Metabolite-M in healthy volunteers after oral administration
(b) To quantify the PK and PKPD variability between subjects
(c) To predict dosing regimens that will achieve a long-term reduction in LDL of 50–90%. This information will be used to inform dose selection for phase II studies

Methods

Study Design and Data Description

Data for the population PKPD analyses was sourced from a phase I dose-ascending study. The study included 50 healthy volunteers who received either 1 mg, 3 mg, 10 mg, 30 mg, or 100 mg daily of Drug-P for 14 days. An additional 10 healthy volunteers were also included in a placebo group in each dosing arm (i.e., an extra two subjects were given placebo in each arm). Samples for Drug P and Metabolite-M plasma concentrations were collected pre-dose, and 0.5, 1, 1.5, 2, 3, 3.5, 4, 5, 6, 8, 10, 12, and, 24 h after the dose on study days 1, 7, and 14. LDL-cholesterol plasma concentrations were collected daily just prior to the dose of Drug-P. All analyses were conducted using untransformed data. Data from placebo groups are excluded from the analyses.

Handling Missing Data

Drug concentration data below the limit of quantification (BLQ) was analyzed using a likelihood based censoring method if $>10\%$ of the data were BLQ (Beal 2001). If $<10\%$ of the data were BLQ, the BLQ samples are omitted from the analysis.

Data with missing independent variables (e.g., dose, sampling times) were to be excluded.

Handling Data Outliers

Outlying data points were identified using the conditional weighted residuals (CWRES) where observations that were more than six standard deviations away from the null (based on CWRES) were excluded from the analysis. A sensitivity analysis was performed where outlying data were included, and model fits with and without outlying data were to be compared.

Computer Hardware and Software Platforms Used in the Analysis

The population analysis was performed using NONMEM (version 7.3) and the Intel visual fortran compiler. Runs were performed using first-order conditional estimation (FOCE) with INTERACTION. Pre- and post-processing was performed using SAS v.9.3. The computer processors used were Intel Xeon® CPU, 2.53 GHz.

Model Building Process

Model building was based on a two stage process: (1) development of the best base population model (the best model in the absence of covariates), including structural components and statistical models for random residual variability and between subject variability, and (2) final model refinement and evaluation. No covariate modeling was conducted.

Structural Models Tested

One, two, and three compartment structural models were considered for the disposition of Drug-P. Various absorption models, including transit compartment models, were also considered. A parent-metabolite model was constructed with a one compartment model for the metabolite, Metabolite-M. The conversion of Drug-P to Metabolite-M was fixed to 30%.

An inhibitory turn-over model was considered for the concentration-response relationship of Metabolite-M on plasma LDL-cholesterol concentrations as follows:

$$\frac{dLDL}{dt} = R_{in} \times \left(1 - E_{max} \times \frac{C(t)^{\gamma}}{C_{50}^{\gamma} + C(t)^{\gamma}}\right) - k_{out} \times LDL$$

where R_{in} is the rate of cholesterol synthesis, k_{out} is the rate of cholesterol loss, E_{max} is the maximum value of synthesis inhibition, $C(t)$ is the concentration of Metabolite-M, γ is the sigmoidicity parameter (also termed the Hill coefficient), and LDL is the plasma concentration of LDL-cholesterol. The pharmacokinetic and pharmacodynamic (PKPD) data were analyzed using a sequential modeling method (IPP, individual pharmacokinetic parameters) (Zhang et al. 2003a, b). Metabolite-M plasma concentrations were introduced into the PD models using linear, Emax, and sigmoid Emax models.

Variability and Error Models Tested

Between-subject variability was described using an exponential model in the following format:

$$\theta_i = \theta_{pop} e^{\eta_i}$$

where θ_i is the i^{th} individual's value of parameter θ, θ_{pop} is the typical value of θ in the population, and η_i is the difference between the i^{th} individual's value and the typical value. η_i is assumed to be independent and identically distributed with variance Ω.

Residual variability was described using both an additive and proportional error model in the following format:

$$y_{ij} = y_{ij}\left(1 + \varepsilon_{1ij}\right) + \varepsilon_{2ij}$$

where y_{ij} is the j^{th} observation for the i^{th} individual, y_{ij} is the model predicted observation and ε_1 and ε_2 are the proportional and additive residual errors, respectively.

Covariate Models to Be Tested

No covariates were tested as part of this analysis.

Model Selection

Model selection was guided by: (1) the likelihood ratio test where a decrease in the objective function value (OFV) of 3.84 units (χ^2, $p < 0.05$) with one degree of freedom for nested models was considered significant, (2) graphical goodness of fit plots, (3) visual predictive checks (VPCs), (4) parameter precision, and (5) the biological plausibility of parameter estimates.

Procedures Used for Model Evaluation

Model evaluation was performed by assessing the visual predictive check (VPC) and prediction-corrected VPC (pcVPC) plots. In both cases, 100 data sets were simulated under the final model. For pcVPC, the 95% confidence interval (CI) of

5th, 50th, and 95th percentiles were shown against the observed data.

Model Simulations to Inform Dose Selection

The final PKPD model was implemented in NONMEM. Drug-P, Metabolite-M, and LDL plasma concentrations were stochastically simulated for 1, 3, 10, 20, 30, and 100 mg daily dosing regimens and for 0.5, 1.5, 5, 10, 15, and 50 mg twice daily dosing. Simulations were conducted assuming 50 subjects in each dosing arm and for 12-week duration of Drug-P. Each dosing regimen was assessed by determining the percent reduction in LDL plasma concentrations for each simulation at week 12 of Drug-P. The dosing regimen was selected if the 95% CI for LDL reduction was within 50–90% of baseline.

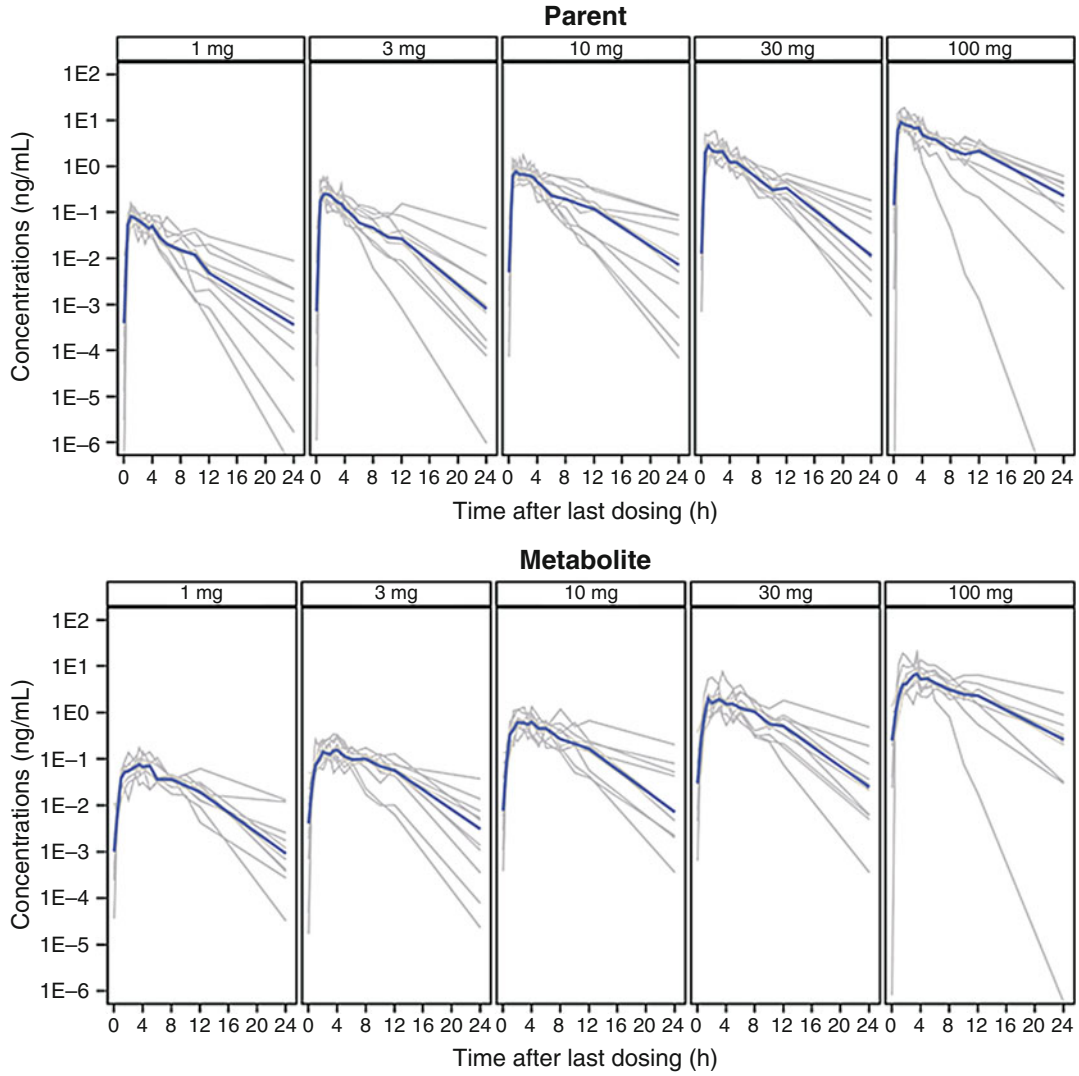


Fig. 8 Spaghetti plots of the raw data for the Parent (Drug P) and metabolite (Drug M) for each dose group. *Blue line* is the median

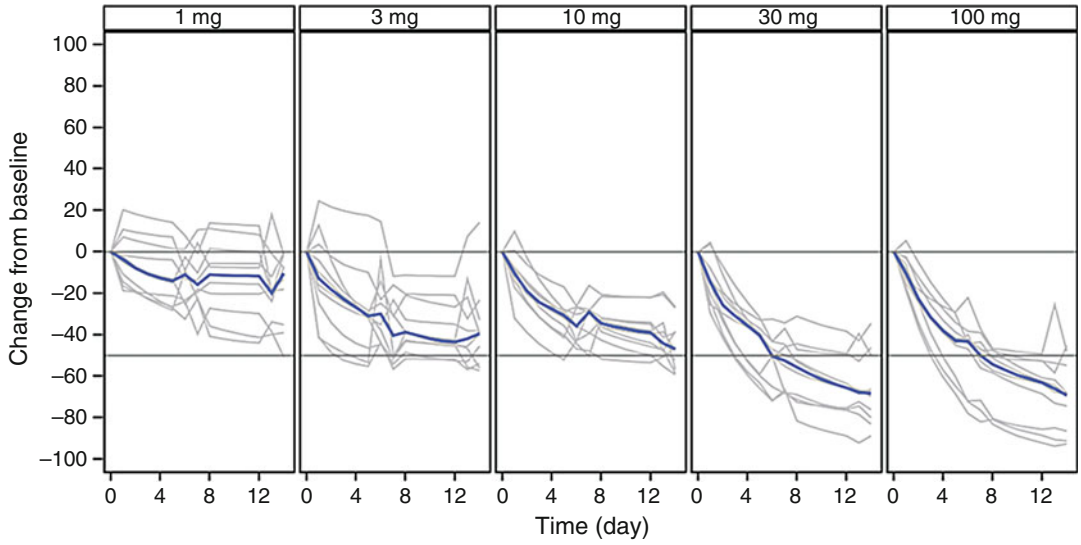


Fig. 9 Spaghetti plots of the percent change in LDL-cholesterol from baseline for each dose group

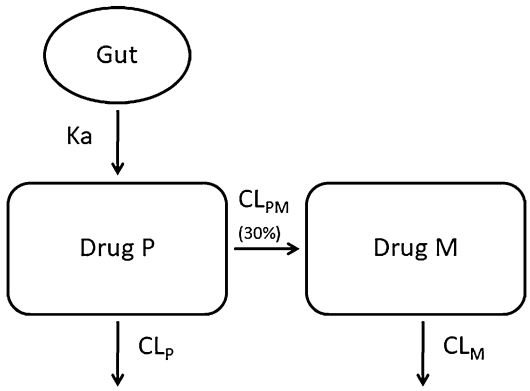


Fig. 10 Schematic of the final PK model for Drug-P and Metabolite-M

Table 2 Parameter estimates for the final PK model for Drug-P and Metabolite-M

Parameter	Population estimate (RSE%)	BSV as CV% (RSE%)
CL_P (L/hr)	2040 (5.6)	39.0 (8.7)
V_P (L)	9210 (4.9)	34.3 (10.1)
CL_M (L/hr)	535 (5.7)	40.1 (9.7)
V_M (L)	948 (6.8)	45.3 (9.0)
ka (hr^{-1})	2.81 (6.9)	38.8 (11.2)
LAG (hr)	0.203 (4.4)	0 FIX
$\sigma_{prop\ (Drug\text{-}P)\ \%}$	30.7 (2.0)	
$\sigma_{prop\ (Drug\text{-}M)\ \%}$	21.0 (1.5)	

CL_P Drug-P clearance, *V_P* Drug-P volume, *CL_M* Metabolite-M clearance, *V_M* Metabolite-M volume, *ka* absorption rate constant, *LAG* absorption time lag, *σ_{prop}* proportional residual error, *σ_{add}* additive residual error

Results

Data Analyzed

The dataset included 2100 Drug-P and Metabolite-M plasma concentrations and 700 LDL-cholesterol plasma concentrations from 50 subjects. All 50 subjects who took Drug-P completed the study and contributed measureable PK and PD samples. All Drug P and M plasma concentrations collected prior to the first dose of Drug Y were BLQ and were excluded from the analysis. No other BLQ samples were present. The final dataset therefore included 2050 Drug P and Metabolite-M plasma concentrations and 700 LDL-cholesterol plasma concentrations from five different dosing regimens. Spaghetti plots of the raw data are presented in Figs. 8 and 9.

Population Pharmacokinetic Model

The best fit to the data for Drug-P and Metabolite-M was a two compartment parent-metabolite model (one compartment for each compound) with first-order absorption and elimination. A schematic of the final PK model is shown in Fig. 10. The final model parameters are

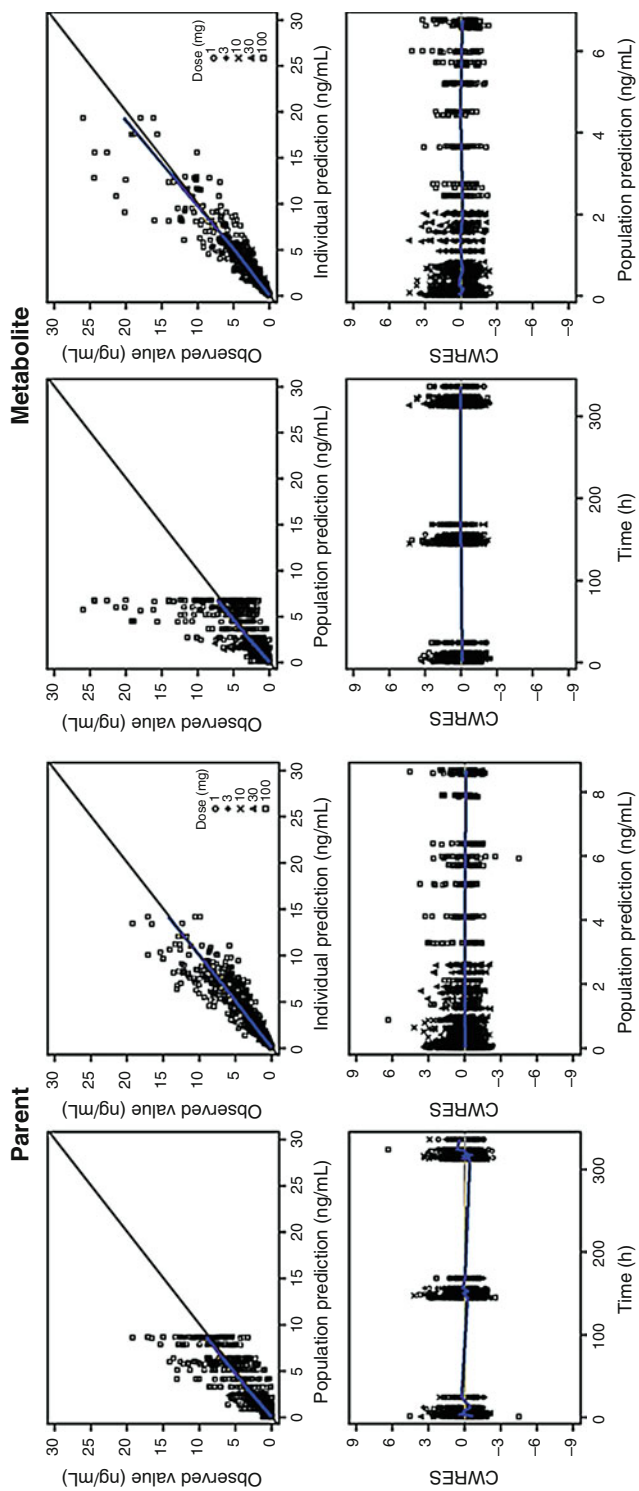


Fig. 11 Goodness of fit plots for Drug-P (parent) and Metabolite-M (metabolite)

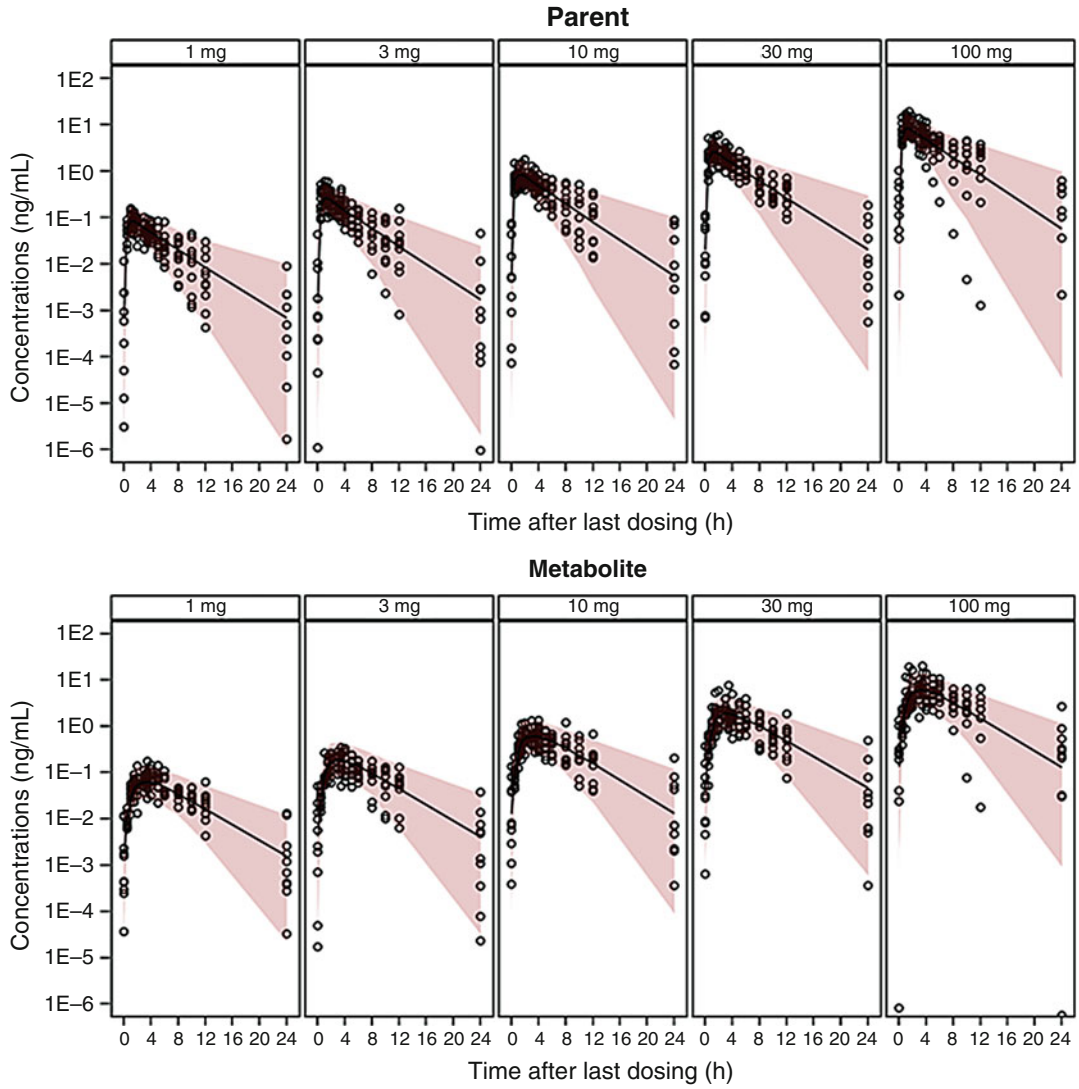


Fig. 12 Visual predictive check (VPC) for the final PK model for Drug-P and Metabolite-M. The median model-prediction (*solid line*) and the 90% prediction interval (*shaded area*) are displayed against the observed data (*black dots*)

summarized in Table 2. Standard goodness-of-fit plots are presented in Fig. 11. The conditional-weighted residual plots (bottom row) show no obvious bias across the model predictions and time. A scatter of under-predicted plasma concentrations for Metabolite B is noted for both the population predictions and individual predictions. These are few in number and are all at higher plasma concentrations. In addition, the predicted versus observed plots presented in Fig. 11 are of limited diagnostic value in the setting of repeated measures data. In this case, the under-predicted values predictions are from two individuals and therefore can be seen to clump together in the plots. Overall, there is no obvious model mis-specification evident from the goodness of fit plots.

Visual predictive checks (VPC) and prediction-corrected visual predictive checks (pcVPC) are presented in Figs. 12 and 13 for Drug P and Metabolite M. In Fig. 12, the median model-prediction and the 90% prediction interval suggest an acceptable model fit. Note that the confidence

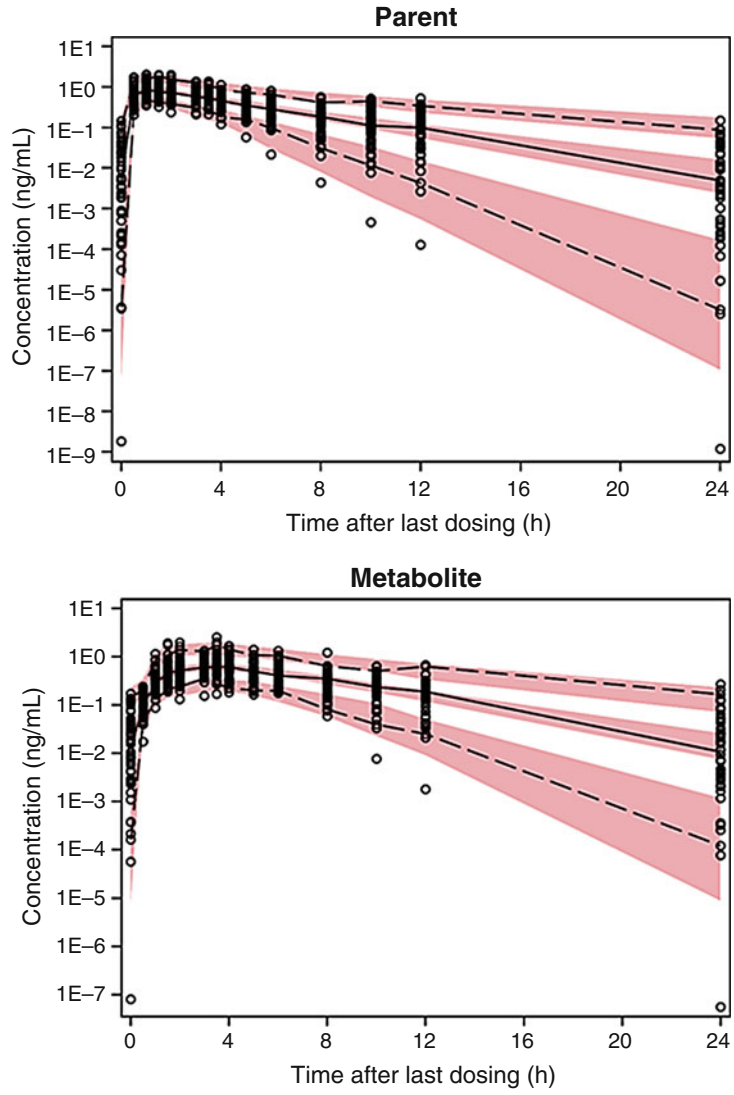


Fig. 13 Prediction-corrected visual predictive check (pcVPC) for the final PK model for Drug-P and Metabolite-M. The median model-predictions (*solid line*) and the 5th and 95th percentiles (*dashed lines*), as well as the 95% CI around the percentiles (*shaded area*) are displayed against the observed data (*black dots*)

intervals are not shown in Fig. 12 due to the small number of subjects. In Fig. 13, the median and 5th and 95th percentiles of the observed data were included within the confidence interval of each percentile for model predicted plasma concentrations, suggesting an acceptable model fit in both Drug-P and Metabolite-M.

Population Pharmacokinetic-Pharmacodynamic Model

The effect of Metabolite-M on LDL-cholesterol was modeled using an inhibitory turnover model. A schematic of the model is presented in Fig. 14. An IPP model was used so the individual PK parameters were fixed to the EBE (empirical Bayes estimate) values obtained from the final PK model. Metabolite-M plasma concentrations entered into the turnover model using an Emax model. The final model parameters are summarized in Table 3.

Standard goodness-of-fit plots are presented in Fig. 15. The conditional-weighted residual plots (bottom row) show no obvious bias for the model predictions. Overall, there is no obvious model

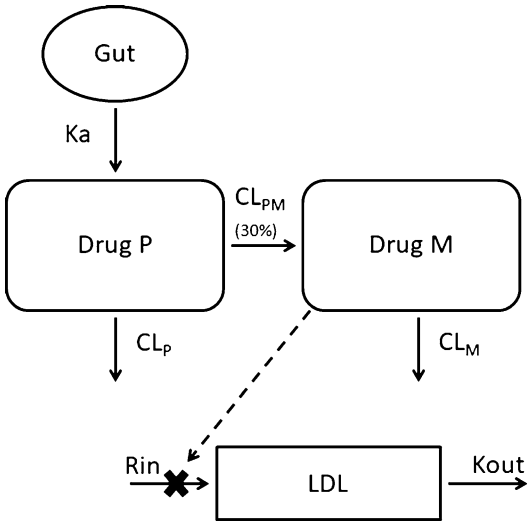


Fig. 14 Schematic of the final PKPD model

misspecification evident from the goodness of fit plots.

Visual predictive checks (VPC) and prediction-corrected visual predictive checks (pcVPC) are presented in Figs. 16 and 17. In Fig. 16, the median model-prediction and the 90% prediction interval suggest an acceptable model fit. Note that the confidence intervals are not shown in Fig. 12 due to the small number of subjects. In Fig. 17, the median and 5th and 95th percentiles of the observed data were included within the confidence interval of each percentile for model predicted LDL concentrations, suggesting an acceptable model fit in both cases.

Model Simulations to Inform Dose Selection

Model-predicted LDL plasma concentrations under once daily Drug-P dosing regimens are presented in Fig. 18. Doses of 10, 20, 30, and 100 mg daily resulted in median LDL concentrations that met the target of a 50% reduction in LDL concentration from baseline. A dose of 100 mg daily resulted in a median LDL reduction from baseline of about 85%; however, the 95% CI of the model-predicted LDL exceeded the 90% target (95% CI of approximately 75–95). Doses of 20 mg and 30 mg produced median LDL reductions of about 65% and 70% (95% CI of approximately 55–75 and 60–80), respectively.

Table 3 Parameter estimates for the final PD parameters

Parameter	Population estimate (RSE%)	BSV as CV% (RSE%)
Rin (mg/dL/hr)	1.04 (3.3)	53.1 (20.5)
Emax	1 Fix	–
C50 (ng/mL)	0.101 (3.1)	115.0 (11.8)
Baseline LDL (mg/dL)	102 (2.6)	17.9 (18.7)
Kout (hr^{-1})	Rin/baseline	–
$\sigma_{prop\ (LDL)\ \%}$	7.7 (7.0)	

Rin production rate of LDL, *C50* the plasma concentration at ½ maximum effect, *Kout* elimination rate constant for LDL, σ_{prop} proportional residual error, σ_{add} additive residual error

Model-predicted LDL plasma concentrations under twice daily Drug-P dosing regimens are presented in Fig. 19. Doses of 5, 10, 15, and 50 mg twice daily resulted in LDL concentrations that met the target of a 50% reduction in LDL concentration from baseline. Dosing at 50 mg and 15 mg twice daily resulted in LDL reductions that exceeded the predefined upper limit of 90% (95% CI for each dose includes 90%). Doses of 5 mg and 10 mg produced median LDL reductions of about 63% and 77% (95% CI of approximately 55–74, and 66–86), respectively.

Dose Selection for Phase II Studies

A dose of 20 mg once daily was selected as the dosing regimen to progress to Phase II trials based on the pharmacometric analyses. The median and 95% CI of the model-predicted LDL reduction were within the predefined target of 50–90% reduction from baseline (65%, 95% CI of approximately 55–75). Although the 5 mg twice daily produced LDL concentrations that met the LDL reduction target with lower daily dose compared to 20 mg once daily, the once daily dosing was preferred due to ease of administration in the clinic leading to higher compliance.

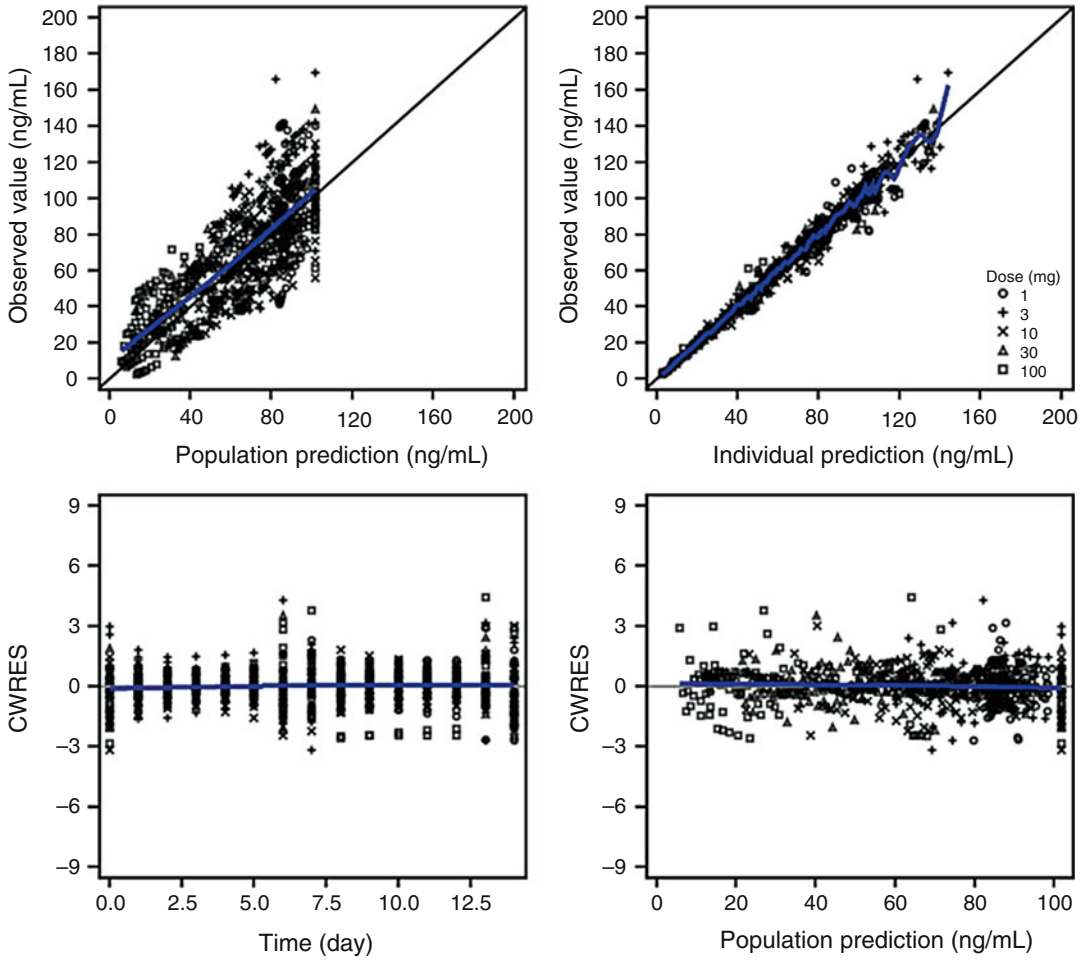


Fig. 15 Goodness of fit plots for the LDL data

Conclusion to Part 4

Population pharmacokinetic and pharmacokinetic-pharmacodynamic models were developed for Drug P, Metabolite M, and the impact on LDL-cholesterol reduction. The best fit to the data for Drug P and Metabolite-M was a two compartment parent-metabolite model (one compartment for each compound) with first-order absorption and elimination. The effect of Metabolite M on LDL-cholesterol was modeled using an inhibitory turnover model. A dose of 20 mg once daily was selected as the dosing regimen to progress to Phase II trials based on the model-predicted LDL reduction of 65% (95% CI of approximately 55–75) from baseline.

Conclusion

Clinical pharmacology and pharmacometrics share common goals and research themes, including pharmacokinetics and pharmacodynamics. A primary purpose of population pharmacokinetic and pharmacokinetic-pharmacodynamic analyses is to aid dosing decisions, whether in the industrial or clinical setting. The goal is to ensure that medicines are safe and effective and

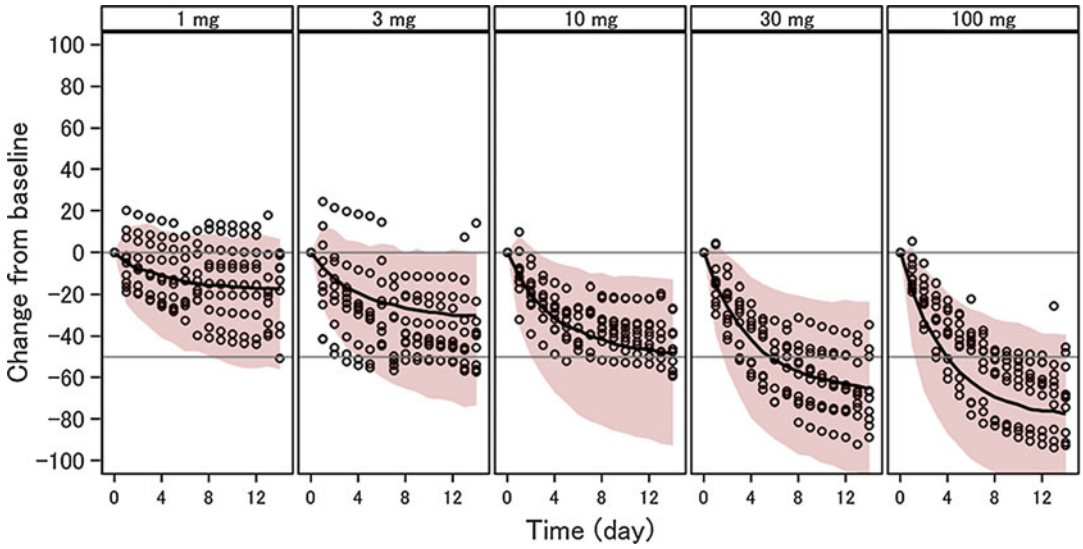


Fig. 16 Visual predictive check (VPC) for the final PKPD model (LDL). The median model-prediction (*solid line*) and the 90% prediction interval (*shaded area*) are displayed against the observed data (*black dots*)

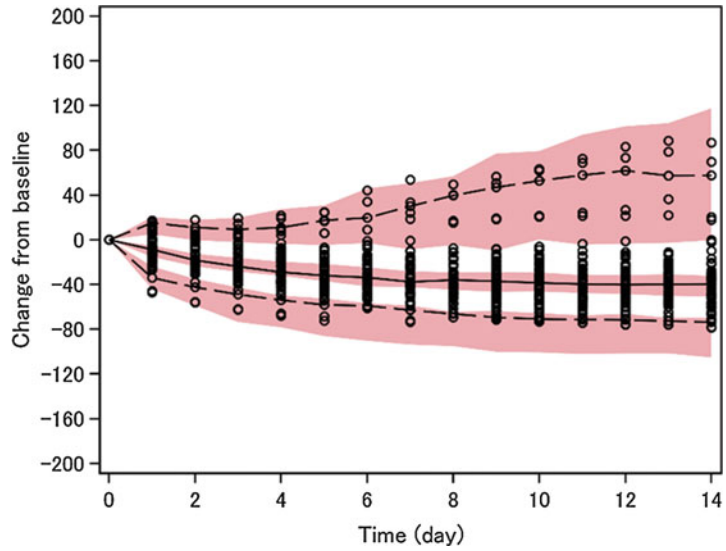


Fig. 17 Prediction-corrected visual predictive check (pcVPC) for the final PKPD model (LDL). The median model-predictions (*solid line*) and the 5th and 95th percentiles (*dashed lines*), as well as the 95% CI around the percentiles (*shaded area*) are displayed against the observed data (*black dots*)

that dosing guidance is underpinned by a scientific understanding of drug behavior and pharmacological response. In addition, variability in drug response between and within individuals can be quantified, and factors which predict this variability, such as body weight, can be accounted for. The resulting models have utility for predicting drug response into new settings, such as the simulation of Phase II studies using a model developed from Phase I data presented in this chapter. By optimizing dose selection, population pharmacokinetic and pharmacokinetic-pharmacodynamic analyses help to improve the chances of success in confirmatory Phase III trials for new drugs and to aid dose individualization in the clinical setting.

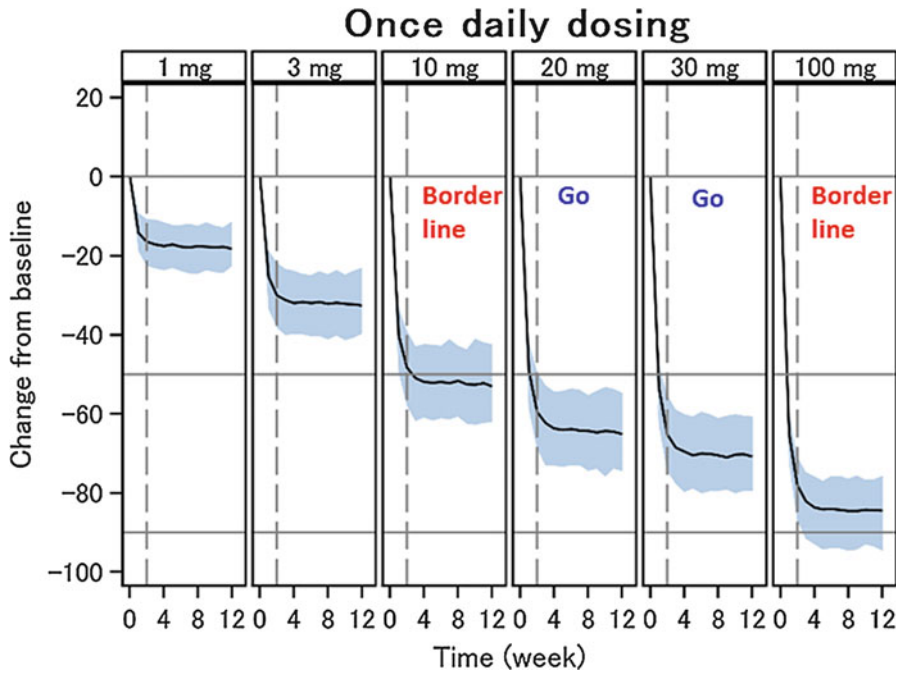


Fig. 18 Model-predicted percent change from baseline for LDL concentrations after once daily dosing for 12 weeks. The *black line* is the median prediction and the shaped *blue areas* are the 95% CI

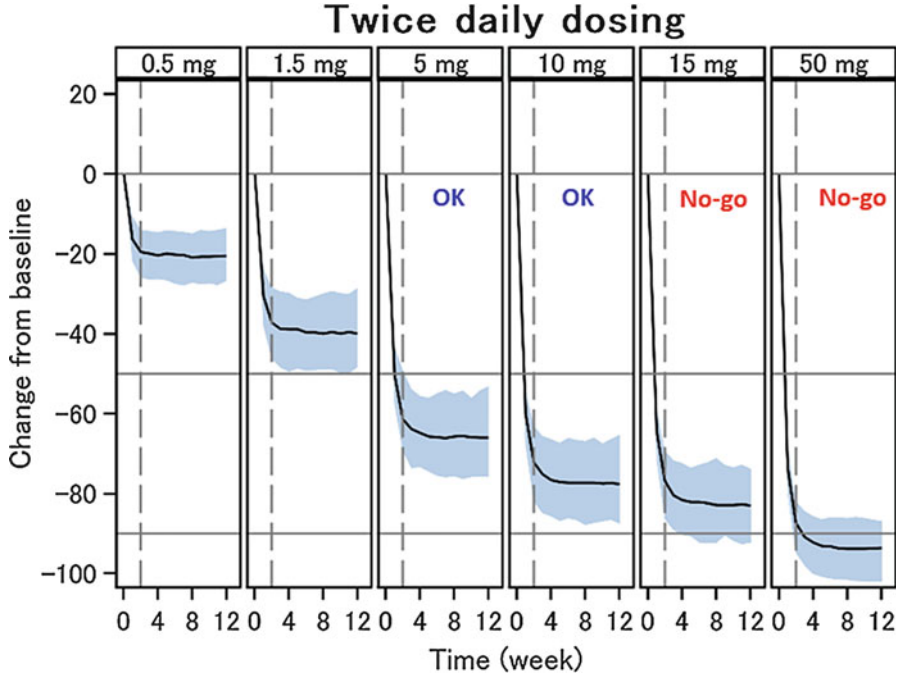


Fig. 19 Model-predicted percent change from baseline for LDL concentrations after twice daily dosing for 12 weeks. The *black line* is the median prediction and the shaped *blue areas* are the 95% CI

References and Further Reading

Al-Sallami HS, Cheah SL, Han SY et al (2014) Between-subject variability: should high be the new normal? Eur J Clin Pharmacol 70:1403–1404

Ariens EJ, Van Rossum JM, Simonis AM (1957) Affinity, intrinsic activity and drug interactions. Pharmacol Rev 9:218–236

Beal SL (2001) Ways to fit a PK model with some data below the quantification limit. J Pharmacokinet Pharmacodyn 28:481–504

Beal S, Sheiner LB, Boeckmann A, Bauer RJ (2011) NONMEM user's guides. (1989–2011). Icon Development Solutions, Ellicott City

Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69:89–95

Bonate PL (2011) Pharmacokinetic-pharmacodynamic modeling and simulation, 2nd edn. Springer, New York

Davidian M, Giltinan DM (1995) Nonlinear models for repeated measurement data. Chapman & Hill, London

Dayneka NL, Garg V, Jusko WJ (1993) Comparison of four basic models of indirect pharmacodynamic responses. J Pharmacokinet Biopharm 21:457–478

Duffull SB (2012) Is the ideal anticoagulant a myth? Expert Rev Clin Pharmacol 5:231–236

Holford NH, Buclin T (2012) Safe and effective variability – a criterion for dose individualization. Ther Drug Monit 34:565–568

Holford NH, Sheiner LB (1982) Kinetics of pharmacologic response. Pharmacol Ther 16:143–166

Jusko WJ, Ko HC (1994) Physiologic indirect response models characterize diverse types of pharmacodynamic effects. Clin Pharmacol Ther 56:406–419

Leake CD (1961) The scientific status of pharmacology. Science 134:2069–2079

Rowland M, Tozer TN (2011) Variability. Clinical pharmacokinetics and pharmacodynamics: concepts and applications, 4th edn. Lippincott Williams & Wilkins, Philadelphia

Sharma A, Jusko WJ (1998) Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. Br J Clin Pharmacol 45:229–239

Sheiner LB, Beal SL (1980) Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm 8:553–571

Sheiner LB, Rosenberg B, Marathe VV (1977) Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokinet Biopharm 5:445–479

Siripuram VK, Wright DFB, Barclay ML et al (2017) Deterministic identifiability of population pharmacokinetic and pharmacodynamic models. J Pharmacokinet Pharmacodyn 44:415–423

Vonesh EF, ChinChilli VM (1997) Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York

Wright DFB, Winter HR, Duffull SB (2011) Understanding the time course of pharmacological effect: a PKPD approach. Br J Clin Pharmacol 71:815–823

Zhang L, Beal SL, Sheiner LB (2003a) Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. J Pharmacokinet Pharmacodyn 30:387–404

Zhang L, Beal SL, Sheinerz LB (2003b) Simultaneous vs. sequential analysis for population PK/PD data II: robustness of methods. J Pharmacokinet Pharmacodyn 30:405–416

In Silico Drug Repositioning Using Omics Data: The Potential and Pitfalls

51

Enya E. Scanlon and Jaine K. Blayney

Contents

E. E. Scanlon · J. K. Blayney (✉)
Stratified Medicine Group, Centre for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast, UK
e-mail: j.blayney@qub.ac.uk

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_20

Abstract

With the recognition of the heterogeneity within complex diseases, such as cancer, there is an accompanying understanding of the need for a stratified approach to treatment. Patients with different underlying biologies originating at the genomic, epigenetic, or transcriptomics levels may present with similar phenotypes at diagnosis. The same treatment may thus result in different outcomes. Using the wealth of public information that is available, particularly from high-throughput experiments, regarding the behavior of approved drugs may facilitate the discovery of novel treatments for subgroups of patients. In silico approaches to drug repositioning have been developed over the past 15 years with a view to enabling this process, with a focus on mapping compounds to patient phenotypes and uncovering novel mechanisms of action. An understanding of the core structure and design of each of these tools, possible applications, and how different inputs can influence results is essential in order that users can maximize the potential of such in silico analyses. This in turn will accelerate the preclinical stage of the biomarker translational pipeline, often perceived as a key bottleneck.

Application of -Omics Technologies in Biomedical and Clinical Research

In the era of digital medicine, high-throughput sequencing technologies have generated exponential amounts of data allowing scientists to examine details of disease pathogenesis in silico at the single-omic and multi-omic level (Ritchie et al. 2015). Multi-omics combines multiple biological data types inclusive of genomic, epigenetic, transcriptomic, proteomic, metabolomic, microbiome, and single-cell data all of which contribute to the genetic profile of an individual. Compared to a single-omics approach, multi-omics promises to gather information from multiple data types to better inform the biology of complex diseases such as autoimmune disorders and cancer (Holzinger and Ritchie 2012). By profiling disease at the molecular level, we can quantify different aspects of genomic activity from biological tissue reflecting fluctuations in physiology across disease states, unveiling patterns in disease progression. This type of analysis can lead to identification of novel drug targets, more effective therapeutics, and improved patient survival, yet the approval process in novel drug development can often take on average 14 years from discovery to distribution (Kaitin and Dimasi 2011).

Rationale for Drug Repositioning

The past 20 years have seen a rise in infectious diseases caused by multidrug-resistant pathogens, along with high resistance rates in patients undergoing long-term therapy for cancer and chronic illness (Aslam et al. 2018; Camidge et al. 2014).

This, together with the continued growth of the global population, has resulted in a distinct need for effective and novel pharmaceuticals for the treatment of disease. Yet, recent years have seen a downturn in the growth and sales of pharmaceutical companies. This is mainly as a result of the investment of billions of dollars into research and development each year for de novo drug

development, which in turn is due to the resource-intensive process of high-throughput screening in vitro. Unfortunately, due to high attrition rates in pharmaceutical development, research costs vastly outweigh productivity (Alex et al. 2015).

To overcome the issue surrounding research expense and time to approval, researchers in academia and industry are collaborating in order to identify drugs, with existing approvals, for novel therapeutic purposes in a process known as drug repositioning or drug repurposing. Advantage can be taken of previously approved drugs having undergone rigorous dosing regimens to ensure limited side effects and optimal pharmacokinetic (PK) and pharmacodynamic (PD) properties. The prediction, and subsequent validation, of new clinical indications in previously approved drugs can circumvent the need for extended dosage and toxicity trials while enabling a faster turnaround for disease therapeutics reaching the patient population (Laifenfeld et al. 2017).

Disease Pathology Underpins Repurposing Potential

Enhanced understanding of disease etiology at the molecular level enables the potential for drug repositioning in clinical practice due to similarities in pathology across disease. A well-established example of drug repositioning is sildenafil (Viagra), originally used to treat coronary heart disease, the side effects of the drug led to approval for the treatment of erectile dysfunction (Boolell et al. 1996). What may not be as well-known are the further applications of drug repositioning across a range of clinical disciplines, from oncology to immunology to psychiatry, distinct biological similarities in disease phenotypes and known mechanisms of action in pharmaceuticals inform repositioning potential.

Repositioning: Prostate Cancer to Gynecology

One example of known mechanisms of action in a therapeutic is leuprorelin, a gonadotrophin-releasing hormone (GnRH) agonist which is widely used to treat advanced stage hormone-responsive prostate and breast cancers, with secondary application in gynecology in the treatment of chronic pain in endometriosis and controlled ovarian stimulation with in vitro fertilization (IVF) (Gerhard et al. 1992). Endometriosis is a common condition affecting approximately 10% of women; it is estrogen-dependent and occurs when endometrial tissue grows on organs outside of the uterus. Although there is no cure for the disease, the use of GnRH agonists has been proven to drastically reduce the severity of the chronic pain experienced by these patients (Gerhard et al. 1992). Leuprorelin acts as an agonist at the GnRH receptors on the pituitary gland, firstly resulting in increased secretion of both follicle stimulating hormone (FSH) and luteinizing hormone (LH) form the anterior pituitary (Schally et al. 1976). This results in increased serum levels of testosterone and estrogen via the hypothalamic pituitary gonadal axis, yet this process becomes desensitized after several weeks of GnRH agonist therapy resulting in the downregulated activity of GnRH receptors and decreased secretion of FSH and LH (Schally et al. 1976). The result is a consistent downregulation of serum testosterone and estrogen levels, limiting the pathogenesis of hormone-dependent disease regardless of gender.

Repositioning: Diabetes to Fertility, Cancer, and Beyond

A second example of drug repositioning is metformin, a widely available, low-cost medication given to patients with type II diabetes, used to help regulate blood sugar levels (Kim et al. 2013). From its original use in diabetes, the drug has found further clinical use in a variety of pathological diseases. Those with either proven or potential benefits with metformin include gestational diabetes, polycystic ovarian syndrome, nonalcoholic fatty liver disease, HIV-related metabolic abnormalities, and weight gain caused by antipsychotic medications (Chae et al. 2016). A huge area of current investigation is the use of metformin in helping to treat various cancers.

Due to reduced cancer risk noted in patients treated with metformin in addition to benefits associated with treating ovarian cancer, and decreased risk of colorectal cancer, clinical trials continue to assess the efficacy of metformin in treating pancreatic, breast, cervical, uterine, endometrial, thyroid, prostate, multiple myeloma, leukemia, and non-small cell lung cancers (Chae et al. 2016; Evans et al. 2005; Liu et al. 2017). Metformin reduces serum glucose levels by increasing uptake throughout the body, decreasing absorption from the intestine, and decreasing hepatic glucose production, resulting in increased insulin sensitivity. Metformin is known to achieve this outcome through inhibiting the activity of mitochondrial complex 1 which converts energy from food into a form that can be readily utilized by cells, yet that is only one way in which the drug is known to work, with other mechanisms of action still being elucidated (Rena et al. 2017).

Although drug repositioning can occur due to known similarities across disease pathologies (leuprorelin), most often opportunities arise through random drug screens. Often cases for drug repositioning arise sporadically such as: side effects of a drug in clinical trials (Viagra), observations during clinical trials (metformin), or by serendipity in the case of thalidomide for treating erythema nodosum leprosum (ENL) – a painful, inflammatory, immune complication of leprosy patients effecting fat cells under the skin. This instance occurred as a last-ditch attempt of a practitioner to help a patient in agonizing pain get to sleep, due to its strong sedative properties and the limited drug supply in a rural hospital, the practitioner gave thalidomide. The patient awoke after a night of sleep with no remaining pain leading to further testing of the drug for ENL and eventual approval at clinical trials; since then, this drug has seen further approval for treatment of multiple myeloma and in clinical trials for inflammatory bowel disease (Ashburn and Thor 2004).

In addition to the examples explained above, there are currently drugs being repositioned across a range of diseases. Some examples of drugs currently in trials for repositioning and those which have been repositioned successfully with further potential are described in Table 1.

In Vitro Versus In Silico Applications in Drug Repurposing

Both in vitro and in silico approaches to drug repositioning have their advantages and disadvantages. In vitro refers to the direct application of pharmaceutical product in drug screening to, for example, human cell lines, whereas in silico refers to bioinformatic systematic analysis of interactions between drugs and their respective targets derived from -omics data. Some advantages of in silico approaches over in vitro are reduced cost and faster results, assay development is not required, and does not involve working with the drug itself. Although the in silico approach identifies large numbers of results quickly, this method is prone to high rates of false positive testing during the drug screening process. In vitro approaches benefit from lower rates of false positives and easy validation of potential drug candidates. Therefore, by combining the benefits of both methods, we can use the results of in silico analysis to inform drug screens in vitro ensuring robust and efficient validation before progressing to preclinical and clinical testing.

Potentials and Pitfalls of Single-omics and Multi-omics Analytics

Multi-omics and single-omics for in silico analysis in drug repositioning both have their own respective advantages and disadvantages. Single-omics approaches can be limited to correlations between disease and control groups, often reflecting reactive or symptomatic processes rather than determining the root cause of disease. Multi-omics approaches may draw data from different sources, for example, incorporating patient data with curated or predicted protein–protein interactions, derived in turn multiple external sources (profiling the same patient sample at different –omics levels is both tissue and resource intensive). Furthermore by increasing the data complexity, more complex

Table 1 Drugs with proven and potential use in drug repurposing. *PTSD* posttraumatic stress disorder, *ADHD* attention deficit hyperactivity disorder, *RA* rheumatoid arthritis. (Table adapted from ClinicalTrials.gov available at: https://clinicaltrials.gov)

Drug	Original indication	Repurposed for (FDA approved)	Potential use (clinical trials)
Amantadine	Influenza		Parkinson's, brain injury, multiple sclerosis
Amphetamine	Decongestant		Stroke recovery, ADHD
Arsenic	Syphilis	Leukemia	HIV/AIDS, neuroblastoma
Aspirin	Antipyretic		Asthma
Ceftriaxone	Bacterial infection		Amyotrophic lateral sclerosis, psychosis
Gabapentin	Epilepsy	Nerve pain	Oropharyngeal cancer
Infliximab	Ulcerative colitis, rheumatoid arthritis	Crohn's disease	Bipolar depression
Methadone	Opioid addiction		Chronic pain
Methotrexate	Cancer	Psoriasis, rheumatoid arthritis	Asthma, schizophrenia
Modafinil	Narcolepsy		Major depressive disorder, PTSD
Naproxen	Inflammation		Alzheimer's disease prevention
Retinoic acid	Acne		Melanoma, Alzheimer's, Schizophrenia
Tamoxifen	Breast cancer	Leishmaniosis	Meningitis, muscular dystrophy
Thalidomide	Sedative	Multiple myeloma	Crohn's disease, Lung cancer
Zidovudine	Cancer		HIV/AIDS

analytical methods are often required to answer the relevant clinical and biological questions. Single-omics approaches may have advantages, depending on the type selected, for example, the biological signals within the higher levels, such as downregulation of genes (transcriptomics), encapsulate those found at lower levels, that is genomic (mutations) and epigenetic (hyper-methylation). Furthermore, transcriptional data is the first choice for measuring drug response in a preclinical setting.

Enhanced Understanding of Disease Etiology

The true potential of -omics analyses comes to light when considering the multiple factors influencing the etiology of complex diseases, defined as having genetic, environmental, and immunological factors potentially contributing to disease progression. Examples of complex diseases include cancer, diabetes, multiple sclerosis, asthma, epilepsy, inflammatory bowel disease, and depressive disorders. The integration of -omics data can be classed broadly into three factor-based approaches, namely phenotypic, environmental, and genetic. Each approach can inform a different aspect of complex disease progression with phenotypic analysis informing associations between disease states, molecular data, and clinical observations, typically using transcriptomic or proteomic data. Combined with microbiome and metabolomic data, we can deduce environmental associations with diet, exercise, smoking, and alcohol, and finally genomic data assessing chromosomal locations, mutations, and genes associated with disease risk.

Through analyzing -omics data types in parallel, it is possible to unveil patterns, highlighting the intricate details of a biological network, for example, connecting transcriptomic phenotypes to epigenetic silencing or a genetic mutation leading to loss of protein production. These small but significant details are what enable identification of drug targets or potential for drug repositioning, in efforts to restore the lost function or to suppress a molecule being overexpressed. Through

analyzing these details in large patient populations, we can identify patterns associated with disease phenotypes giving rise to molecular stratification of disease. Such patient stratification can facilitate an enhanced understanding of disease progression routes and a more personalized approach to treatment for patients. A prime example of this type of analysis was carried out through molecular profiling in colorectal cancer subgroups (Guinney et al. 2015).

Stratified Medicine

Using –omics data, it is possible to classify patient populations into disease subtypes, each with their own respective prognosis, mechanisms of action, clinical presentation, response to therapeutics, risk of disease recurrence, and molecular patterns reflecting variation in biological pathways involved in disease progression. The stratified medicine approach ensures that patients receive the most effective treatment for their distinct molecular profile at the correct time, rather than a one-drug-fits-all approach, resulting in ineffective drug therapy with adverse effects on subsets of the patient population.

Through analyzing -omics data types, researchers can identify biomarkers, either single or multiple pathologically driven disease characteristics that can be associated with a molecular subtype, diagnosis, prognosis, or predictive of response to treatments, further elucidating the etiology of complex disease. Predictive biomarkers in this case, also known as companion diagnostics, are diagnostics tests enabling a clinician to make clear-cut treatment decisions based on robust biostatistical analytics of disease markers derived from -omics data. The markers of response in disease subtypes determine the efficacy of a therapy for an individual based on their distinct molecular profile. One such example of a companion diagnostic is testing for a KRAS mutation, shown to confer resistance to cetuximab, in the treatment of advanced CRC (Gatell et al. 2008). Biomarkers of disease progression can also be repurposed for therapeutic drug monitoring, reflecting changes in PK and PD as a treatment continues.

Big Data, Big Problems

Although large-scale analytics of single-omics and multi-omics data holds vast potential, the variable dimensions and properties of multiple data types introduce difficulties during integration required for robust biostatistical analysis (Huang et al. 2017). When integrating various data types, we must consider preprocessing steps that vary across -omics types due to being developed on different platforms with subsequent non-standardized quality control protocols for removal of technical effects. These steps are vital in ensuring that patterns observed in data are reflective of biological variation between disease states rather than associations with technical artefacts. Within the context of drug repositioning, if these steps are not carried out appropriately, there will be high numbers of false positive results, highlighting the importance of in vitro validation.

For these reasons, the process of -omics data mining is heavily reliant on the field of bioinformatics, integrating, analyzing, and interpreting the statistical behaviors and subtle relationships relevant to each -omics platform and their original biological context. Therefore, prior knowledge of statistical programming and mathematical modelling are required for robust analysis, creating a bottleneck due to large amounts of biological data being produced and a small population of trained bioinformaticians for analysis and interpretation.

In Silico Approaches to Drug Repositioning

Categories of In Silico Approaches

In silico approaches can be grouped into several categories depending on approach, data input, statistical approach, and inferences on results obtained (Fig. 1).

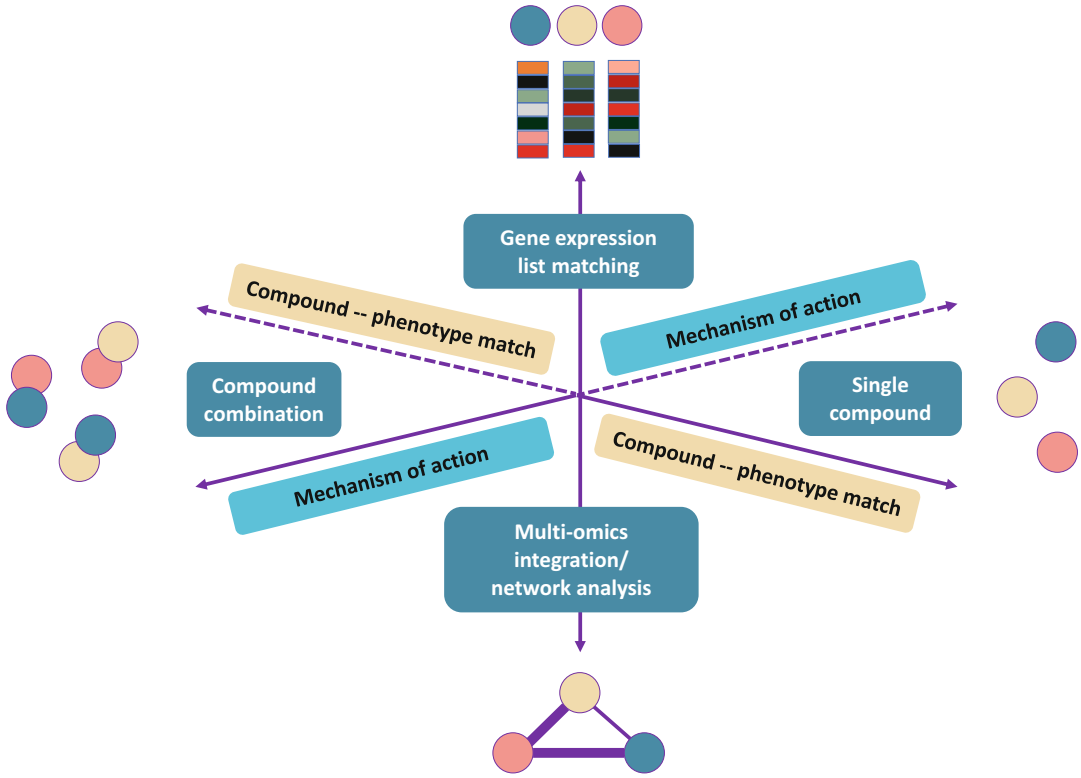


Fig. 1 Schema grouping in silico methods of drug repositioning into key classes based on methods (gene list similarity, network analysis), inputs (gene lists, reference libraries), and inferences (compound-phenotype and compound-mechanism of action) drawn

Compound-Phenotype Match

The simplest methods, in terms of output and interpretation, are those which are gene-list (transcriptomics) based. In this context, a set of genes which encapsulate phenotypic and thus, biological, significance are used as input to methods which are, in turn, based on the concept of compounds being applied to "reverse" the activity of genes in a disease or phenotypic state (this reversal will depend on the order of the comparison, so similarities in activity may be more appropriate). For example, if in patients with disease X, genes A, B, and C were upregulated with respect to healthy patients, then drug Y whose activity was found to reduce genes A, B, and C would be a possible therapeutic candidate. Examples of the compound to phenotype gene-list approach include CMAP (Lamb et al. 2006) and sscMAP (Zhang and Gant 2008) and LINCS (Subramanian et al. 2017). Others, e.g., CDA (Lee et al. 2012), have extended this approach to consider synergistic/antagonistic effects of compound pairs.

Compound-Mechanism of Action Match

Characterizing a compound's mechanism of action includes identification of the targets (genes or biological pathways) and the subsequent action on those targets. Methods that elucidate a compound's mechanism of action tend to be based on a "guilt by association" hypothesis and use multiple data types, including multi-omics, and complex modelling techniques, particularly network analysis. A simple example being that if the action of compound Y_a results in upregulation of genes A, B, and C and downregulation of D, E, and F in disease X, if compound Y_b results in the same behavior, both compounds are considered as having a similar mechanism of action. Within a network, these compounds will be represented as nodes with connections between them (edges).

The closer the nodes and/or thicker the edges, the more mechanistically similar. Examples include: Mantra (Iorio et al. 2010), MD-Miner (Wu et al. 2017), and DIGREM (Zhang et al. 2019).

Compound Datasets

Regardless of statistical methods applied or intended output, at the core of most in silico repositioning tools lies one of two datasets.

The first, the Connectivity Map (CMAP) build 2, was developed by the Broad Institute, (Lamb et al. 2006). This version of the dataset consists of transcriptional profiles (Affymetrix U133a platform) for five cell lines, including cancer cell lines MCF7 and ssMCF7 (both breast), HL60 (leukemia), PC3 (prostate), and SKMEL5 (melanoma). The full dataset includes 7,000 expression profiles including untreated states and treated states (by approximately 1,300 compounds). No longer updated nor supported, the data is still available via accession number GSE5258 from the Gene Expression Omnibus (GEO) (Edgar and Domrachev 2002). Various external tools may carry out further processing or transformation of the data before using it as a reference, e.g., converting expression levels into ranked form between treated and untreated cell lines. Despite its size, this version of the CMAP dataset is limited in that it considers four cancer types, i.e., three solid tumors and one blood. There are a number of recognized caveats, including the limited ability of four cancer types to fully characterize compound-induced behavior across multiple disease sites or even heterogeneity within the disease itself.

An update to CMAP, the Library of Integrated Network-Based Cellular Signatures (LINCS), was released in 2017 (Subramanian et al. 2017). Instead of the original microarray platform, transcription was quantified using the L1000 assay, using ligation-mediated amplification. This new platform focuses on the expression level of approximately key 1,000 genes, which were selected after a meta-analysis of publicly available datasets. This meta-analysis demonstrated that the remainder of the transcriptome could be estimated reliably by imputation techniques. Additional cancer cell lines in untreated and untreated states are included with this release, including: breast (MCF7), prostate (PC3, VCaP), lung (A549, HCC515), melanoma (A375), liver (HePG2), and colorectal (HT29). In total, 1.3 million L1000 profiles are available corresponding to treatment with 42,080 perturbagens, including 19,811 small molecule compounds. This dataset is available from the GEO (accession number GSE92742). An online tool, the CMAP-linked user environment (CLUE) analysis environment is also available (Broad Institute 2019).

Recently, the Genomics of Drug Sensitivity Cancer (GDSC) Project (Yang et al. 2013) published the results of their 4-year large-scale drug screen, which promises to power the future of drug repositioning. As of the version 8 Release (July 2019), the GDSC drug dataset consists of 1,000 cancer-specific cell lines in the context of 453 treatment compounds, which in turn target 24 biological pathways. Previous versions of the GDCSC dataset facilitated the pharmaceutical development and use of PARP inhibitors in childhood bone cancer, in addition to informing over 70 research studies globally. The GDSC project encompasses over 494,973 genomic associations across 386,293 dose response curves, with the aim of supporting the systematic identification of biomarkers of drug response. Cell line sensitivity to a compound is correlated with vast amounts of genomic associations to unveil genetic features associated with drug sensitivity. The large-scale analysis of cancer-specific cell lines allows for a high-level view of disease heterogeneity underpinning therapeutic response to treatment, thereby connecting genotypes to phenotypes. The GDSC is now the worlds' largest repository of -omics data regarding the genetic influences surrounding cancer and how tumors respond to treatment. All results have been made publicly available and incorporated into the Cancer Dependency Map Project (Wellcome Sanger Institute 2019).

Key Tools

Popular compound-phenotype matching tools include including CMAP, sscMAP, and CLUE. Each follows a similar overall approach, in terms of matching gene list similarity/dissimilarity,

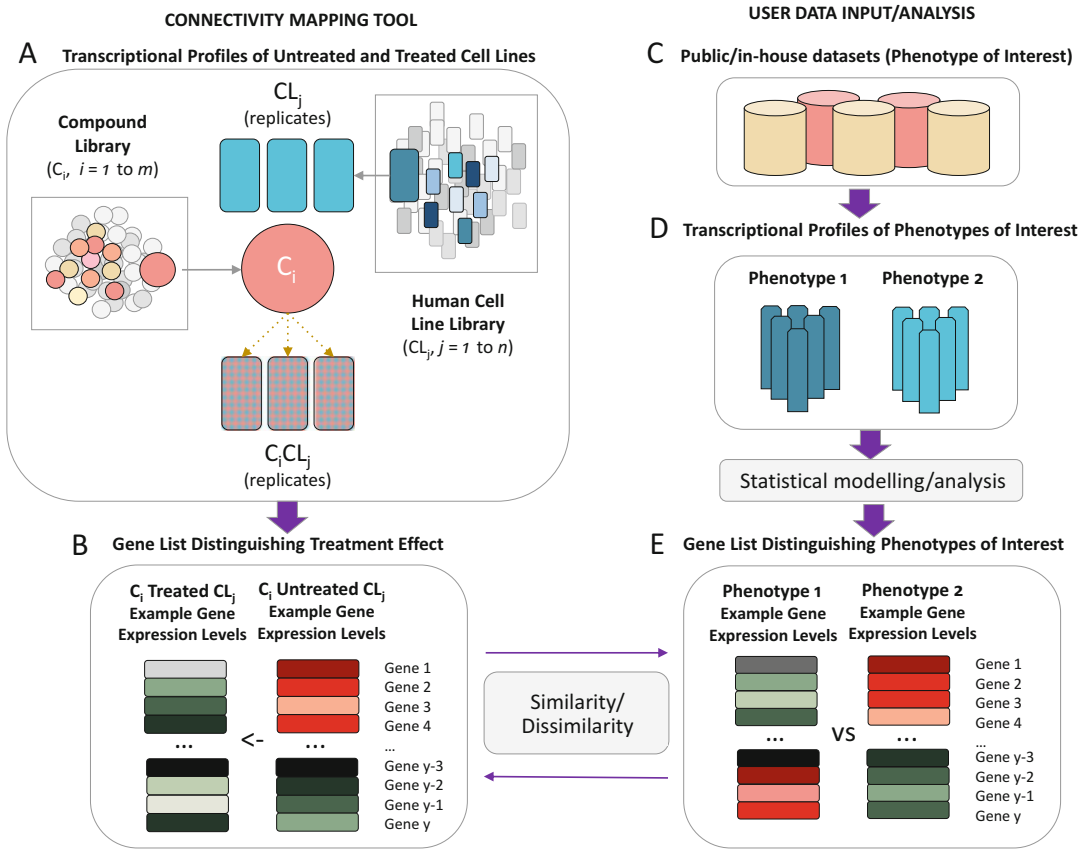


Fig. 2 Schema of generic connectivity mapping approach. (**a**) Reference library of gene expression profiles from multiple human cancer cell lines in both treated states (by different compounds) and untreated states. (**b**) Derivation of gene list corresponding to effect of compound treatment (reference phenotype – treated or untreated – may vary according to in silico tool). (**c**) Curation of public (or in-house) datasets containing phenotypes of interest. (**d**) Selection of gene expression profiles corresponding to phenotypes to be compared. (**e**) Derivation of gene list (various methods) comparing user-selected phenotypes to be compared (various methods) to the compound effect

though each may differ in their statistical implementation (Fig. 2a and b).

The original CMAP and LINCS/CLUE takes a ranked user's gene list, from the analysis of two phenotypes (Lamb et al. 2006; Subramanian et al. 2017). This is ranked by fold change (or equivalent), e.g., between a disease (prostate cancer, diabetes, cystic fibrosis) phenotype and a healthy (or healthier) phenotype, genes are ordered from most upregulated to most downregulated in patients with the disease. This gene list (usually between 10 and 150 in length with a focus on upregulation) is then compared to the transcriptional effect of compounds, with an emphasis on genes at the top and bottom ends of both lists. The comparison is made using a nonparametric test probability distribution comparison test (Kolmogorov–Smirnov). Random permutations are used to estimate the robustness of the resultant connectivity or connection score.

Another connectivity mapping program, statistically significant connections' map (sscMAP) (Zhang and Gant 2008) uses a nonparametric association (Spearman's correlation) test to determine matches between the behavior, at the transcriptional level, of treated cell lines and a user's gene list of interest. In sscMAP, the reference library is based on the second edition of the Broad Institute Connectivity Map (Lamb et al. 2006). The user enters a ranked gene list, i.e., genes ordered by their relative importance that differentiates between two phenotypes. A

connectivity score (similar to CMAP/LINCS) identifies those compounds whose action on cell lines is similar/dissimilar to the difference between phenotypes. The robustness (p-value) of the observed matching compound results is estimated using the results generated by the repeated input of random gene lists. The effect of gene list perturbation is also considered, that is determining if the same compounds are returned when one gene is removed from the list. A further update of sscMAP based on the LINCS/CLUE dataset (Subramanian et al. 2017) was released in 2016, which offered improved scalability and speed (O'Reilly 2016).

In both cases similar and dissimilar (in terms of the gene list used as input) compounds are returned. These require interpretation within the context of how a user derived their gene list (see "Deriving the Gene List"). For example, if a gene list was derived comparing the profiles of poor prognosis against good prognosis esophageal cancer patients, genes of interest would be those that were upregulated in the poor (with respect to good) prognosis patients. This is due mainly to the fact that it is easier to reduce, rather than induce, gene activity through the action of compounds.

The Drug-Induced Genomic Response Model (DIGREM) is a method used in predicting the synergistic effects of compound pairs for identification of novel, multidrug combination therapies. Based on drug-treated transcriptomic data (Subramanian et al. 2017), DIGREM models changes in gene expression and the dynamics of drug response following treatment of individual drugs. The output is a predicted score detailing the synergistic effect of combinational drug therapies in a ranked list prioritizing efficacy of novel drug combinations with the potential for further in vitro validation (Zhang et al. 2019).

In the Combinatorial Drug Assembler (CDA), similarities between a user-defined gene list is compared to the action of single compounds and predicted action (synergistic/antagonistic) of compound pairs. Similarity between query expression profiles and compound action profiles are considered with respect to genes enriched in signaling pathways (Lee et al. 2012).

In a disease-drug network-based approach, three key inputs are used: (a) signaling pathways (sets of genes) together with (b) information on genes targeted by compounds and (c) those genes associated with disease progression (Peyvandipour et al. 2018). This disease-drug network is integrated with gene expression profiles associated with drug/compound effects (CMAP/LINCS) and those associated with disease. Disease-specific and compound effect gene lists are perturbed and compared resulting in a matching score. Using publicly available gene expression datasets, disease-drug network were inferred for idiopathic pulmonary fibrosis, non-small cell lung, prostate, and breast cancers. This network approach was shown to outperform, in terms of known treatments identified, existing gene list-based methods.

NFFinder was originally developed to identify novel compounds for orphan diseases such as Neurofibromatosis (NF) (Setoain et al. 2015). The underlying data including compound effect-associated ranked gene lists from CMAP (Lamb et al. 2006) and disease and/or drug-associated ranked gene lists from public datasets curated from GEO (Edgar and Domrachev 2002). A user's gene list (or an miRNA list converted internally to a gene list) is compared to the disease and/or drug-associated gene lists using Kolmogorov–Smirnov-like statistic (as in CMAP/LINCS).

gene2drug uses the CMAP/LINCS dataset, using a pathway-based version of gene-set enrichment. Instead of ranking differentially expressed genes between treated and untreated cell lines, the effect of each compound on pathways is evaluated – the equivalent of differentially expressed pathways (Napolitano et al. 2018). Pathways are ranked and each drug is allocated an associated enrichment score. gene2drug's predictions were validated experimentally, with potential application to non-cancer conditions including Bird's disease (primary hyperoxaluria) and neurodegenerative disorders.

Applications

In a connectivity mapping (CMAP) study of the most common form of epithelial ovarian cancer, high-grade serous, survival analysis (Cox proportional hazards (PH) regression), was used to

identify genes from a number of publicly accessible clinical datasets, that were associated with recurrence (Raghavan et al. 2016). Hazard ratios were used as the equivalent of fold changes (as in differential expression analysis). Lists derived from univariate and multivariate analyses, adjusted for recognized prognostic clinical factors, were used as input and compared. Interestingly, gene lists derived from two of the datasets did not produce matches with structurally or mechanistically similar compounds currently used to treat EOC, though topoisomerase II inhibitors (as used in breast and testicular cancers) were identified. Also, the high recurrence rate in EOC may be attributed, in part, to the lack of effective treatments. This may in turn explain the lack of platinum-based compounds appearing in this study as matches. Finally, the intrinsic gene lists associated with the characterization of molecular subtypes within high-grade serous EOC may prove more informative.

A similar approach was used in a colorectal cancer connectivity mapping study (sscMAP). Here, prior biological knowledge in terms of driver genes from target pathways was combined with co-expression and survival analysis to derive an input gene list for treatment of colorectal cancers associated with aggressive progression (Yuen et al. 2013). Using publicly available colorectal cancer datasets, patients were stratified into four groups based on the relative expression levels (high/low) of three genes, TAZ, AXL, and CTGF. From these, two patient groups, one good prognosis and one poor prognosis, were characterized via survival analysis (log-rank and Cox PH regression) and differentially expressed genes between these two groups identified using the nonparametric Mann-Whitney test. The connectivity mapping analysis was able to highlight two compounds (amiloride and tretinoin), which had been previously shown to be effective in the preclinical setting.

The use of human cell lines as a reference library has limitations with respect to immuno-oncology discovery, due to the lack of a functional immune system. Reference libraries based on syngeneic transplantable and genetically engineered mouse models would be of more relevance, but resource-wise, the costs would be prohibitive. The use of cancer-based cell lines would appear, at first instance, to be a further limitation. However although CMAP, LINCS, and sscMAP were developed in the context of cancer research, there have been applications in non-cancer fields, including inflammatory disease (Malcomson et al. 2016). Connectivity mapping (sscMAP) was applied to inflammation reduction in cystic fibrosis, particularly through activation of the gene TNFAIP3. Using correlation analysis of public datasets, genes were identified whose expression levels rose or fell with the target gene TNFAIP3. From this gene list, the authors next considered genes associated with specific biological processes, in particular the NF-κB signaling pathway. Although five compounds were identified which induced the target gene (TNFAIP3), counter-productive field effects were also observed.

Also within inflammatory disease, the CMAP-version of connectivity mapping was applied to inflammatory bowel disease (IBD), in order to identify novel treatments for Crohn's disease and ulcerative colitis (Chiang et al. 2011). Differentially expressed genes in publicly available, transcriptional data were identified using a moderated t-test between IBD and control samples. Functional enrichment of biological processes and diseases were used to further filter the gene list. The anticonvulsant topiramate was identified and validated in an in vivo model.

Compound-Mechanism of Action Match

Iorio et al. (2010) developed a drug similarity network, MANTRA, based on transcriptional cell lines profiles from CMAP (Lamb et al. 2006) which facilitates the prediction of the mechanism of actions of query compounds. Available as an online tool, the network is constructed by first determining ranked gene lists corresponding to each compound's generic cell line behavior. The amount by which gene lists for two compounds overlapped at the extremes (top upregulated and downregulated) equated to a distance. Compounds equated to nodes in the network, while edges represented the similarity/dissimilarity between each compound pair's

actions. Functional enrichment allowed for the characterization of key clusters of nodes by biological pathways. The authors used paclitaxel as a proof of concept of the network's function and were able to stratify the compound's primary and secondary effects. A differential gene list corresponding to the behavior of query compounds can then be evaluated within the established network.

Mechanism and Drug Miner (MD-Miner) (Wu et al. 2017) also focuses on signaling pathways within a network-based approach to predict compounds and characterize mechanisms of actions. Two networks are constructed, one comprising drug targets and gene expressional profiles associated with treatment by compounds (Lamb et al. 2006; Subramanian et al. 2017) forms a mechanism of action network, the other a patient-specific disease network consists of genes and gene expression profiles associated with disease. Additional data via transcription factors interactions and protein–protein interactions are included in both networks. Associations between the mechanism network and disease network are assessed by the number of shared genes.

DrugComboRanker is similar in structure to MD-Miner in that two networks are created and an intersection sought between the two. The drug functional network derived from the CMAP/LINCS dataset (Subramanian et al. 2017) incorporates smaller sub-targets based around common drug targets. Disease networks are then inferred using protein–protein interactions and signaling pathways. The targets from drug pairs are next mapped onto the disease network to determine synergy in functionality.

Drug-set enrichment analysis (DSEA) predicts shared mechanisms of actions of compounds, based on similar gene activity (Napolitano et al. 2016). This behavior is considered within the context of pathways, i.e., sets of genes which are functionally related. DSEA uses the CMAP/LINCS dataset to determine common pathway targets among compounds via the enrichment analysis of gene expression profiles. Each compound-pathway pair is allocated an enrichment score and ranked.

Deriving the Gene List

Although the above studies use similar reference datasets and phenotype-compound matching approaches, the methods used to derive phenotype-associated gene lists varied from correlation, differential expression, survival analysis, or a mixture of various techniques, alongside the incorporation of prior knowledge. As with the reference dataset, the robustness and relevance of the gene list to be used as input is key. The choice of method by which to obtain a list of representative genes may depend on a number of issues, including the study's hypothesis and power, which in turn can often be limited by the datasets available.

Data Sources

Researchers may have limited resources by which to curate large patient datasets and/or carry out profiling or sequencing runs. The reanalysis of publicly accessible data therefore presents a viable, cost-effective alternative (Fig. 2c).

Two key open-access repositories are available: the GEO (Edgar and Domrachev 2002) and ArrayExpress (Athar et al. 2019). Both contain datasets (often overlapping) submitted by researchers, usually in accompaniment to published papers. Datasets consist mainly of high-throughout transcriptional profiling or sequencing data, methylation and genomic data. In the case of patient data, clinicopathological data, e.g., survival, staging, sex, and age may be included. Although the majority of datasets focus on patient datasets across a range of diseases, there are additional datasets corresponding to in vitro and in vivo experiments, including human cell lines, patient-derived xenografts, and genetically engineered mouse models. Omics data is usually available in either raw form (which in turn required processing and normalization) and/or processed form. In the case of the GEO, there is an accompanying tool, GEO2R (NCBI 2019), which facilitates differential expression analysis online. For users that wish to download data and analyze the data locally, tools such as

GEOQuery, available as a package within the statistical programming framework, R (suitable for use on Windows, Linux, and Mac, within the BioConductor suite) (The Comprehensive R Archive Network 2019).

Additionally, The Cancer Genome Atlas (TCGA) also offers a range of datasets across multiple -omics levels for those working in cancer research. The most recent pan-cancer set of studies brings together data from 33 cancers (both tumor and normal) across genomic, epigenetic, transcriptomic, proteomic levels (Hoadley et al. 2018). Hosted by the University of California's Xena platform, the accompanying online tools allows users to perform basic statistical analysis on selected datasets (University of California 2016).

There are considerations in the reanalysis of public data. Data may be mislabeled, also batch or technical effects may be undetected/unaccounted for, which in turn can skew the true biological signal. Furthermore, each dataset will have been constructed centered around testing a particular set of hypotheses, e.g., a dataset constructed to investigate the dysregulation of biological pathway X in patients with disease Y under condition Z may not be representative of the full population of patients with disease Z.

Methods

Power

In the derivation of gene lists, to avoid the propagation of errors, it is advisable to minimize type I/II errors through power and sample size estimations. Power is related to sample size, minimum effect size to be detected, and significance levels. In the majority of studies, the minimum power is set at 80%, with significance at 0.05. The same dataset can produce small, medium, and large effect sizes depending on the proposed statistical method, which itself is dependent on the study's hypothesis. The choice of phenotypes to compare may also be limited by data availability (Fig. 2d).

Phenotype Comparison

In some cases, researchers will be interested in the differences between a disease state and healthy controls. In other cases, the pretreatment transcription profiles of EOC patients who have received platinum-based chemotherapy post-surgery and who have not responded (innate resistance) could be compared to patients who have not recurred within 12 months (classed as responders). Likewise for patients with acquired resistance (recurrence within 6–12 months) could be compared to responders. Theoretically, the addition of compounds associated with such gene lists would allow the resistant phenotypes become more similar to the responding phenotypes and therefore benefit from the chemotherapy (Fig. 2d, e). By comparing phenotype class labels, e.g., response and nonresponse, there is the assumption that the biology within each phenotype is the same. This may not be correct as molecular subtyping of disease, e.g., in colorectal cancer (Guinney et al. 2015) has demonstrated. It is quite possible that two (or more) groups of patients present with the same phenotype, e.g., response, yet have different underlying biologies. Unsupervised or semi-supervised clustering, the latter using the top 1,000–5,000 genes with the largest standard deviation, can help with characterization of phenotypes. This approach will only be suitable for larger datasets, though the actual size will depend on the heterogeneity of the phenotype itself.

Differential Expression

Where class labels, e.g., treatment responder/nonresponders, good/poor prognosis, healthy/diseased are available, the choice of statistical modelling methods could include differential expression or regression analysis (Table 2). In parametric versions of either assumption testing is a prerequisite. Differential expression involves two classes of patient (or preclinical models) groups and via multiple t-tests will identify those genes whose expression levels are consistently different across each group. As with any

Table 2 Statistical methods by which to derive gene lists for input into compound mapping tools including: differential expression, regression, functional enrichment, gene-set enrichment, and survival analysis

Method Family	Output	Example methods/tools
Differential expression analysis	Genes (gene list) up-regulated or down-regulated in one phenotype relative to another, e.g., treatment responders versus nonresponders. Fold changes can be used to determine relative "strengths"	T-test, Mann-Whitney, moderated t-test
Regression analysis	Genes whose expression levels can predict class labels (usually binary). The contribution of each gene's expression levels to the outcome variable can be extracted from coefficients	Logistic, ordinal, or linear regression
Functional enrichment analysis	Over-representation of defined biological processes and pathways within user-defined defined phenotype-associated gene list. Allows for the selection of pathways (enrichment scores), genes require ranking by, e.g., fold change within each pathway	Database for Annotation, Visualization and Integrated Discovery (DAVID)
Gene-set enrichment analysis	Over-representation of defined biological processes and associated gene lists within transcriptional profiles in one phenotype with respect to another. Genes may require further ranking by, e.g., fold change within each pathway or process	Gene-Set Enrichment Analysis (GSEA)
Survival analysis	Genes associated with time to event, e.g., recurrence, death. Hazard ratios can be used to rank the contribution of genes to outcome	Log-rank, Cox proportional hazards regression

form of multiple testing, adjustment of p-values is required, e.g., Bonferroni correction or Benjamini and Hochberg false discovery rate (Table 2).

Regression

Although categorization of a continuous variable such as response may be simpler to model and more intuitive to understand, there may be an accompanying loss of information. By using classes or categories, e.g., chemotherapy responders and nonresponders, we are assuming that all responding patient have the same "strength" of response. If tumor size is measured posttreatment relative to pretreatment, it may be found that some patients have a 5% reduction, whereas others have a 50% reduction in tumor volume. Using a form of regression to model tumor size posttreatment may result in a more informative gene list. Univariate regression (with multiple comparison adjustment) can also be used to model class labels (logistic or ordinal regression), e.g., response and nonresponse, as an alternative to differential expression. Both differential expression and univariate regression can return hundreds and sometimes thousands of genes, even with multiple comparison adjustment. Certain connectivity mapping tools may limit the number of genes that can be used as input. The magnitude of fold change or regression coefficient associated with each gene could be used to reduce any list. Additionally, further stratification of patients with respect to phenotype or class labels may help with the development of informative gene lists (Table 2).

Enrichment

Further filtering of gene lists obtained via differential expression analysis can be achieved using functional enrichment. Online tools such as Database for Annotation, Visualization and Integrated Discovery (DAVID) (Immunoinformatics 2019) take a user's gene list and returns biologically associated groups of genes that are estimated to be overrepresented (Huang et al. 2009). For example if Pathway E is associated with 45 genes and a user's gene list ($n = 60$) has 40 of those genes, then Pathway E would be considered as being enriched. A gene list can be analyzed with respect to gene ontology categories,

pathways, and associated diseases, with an enrichment score for each group assigned. Using a form of "fuzzy" clustering, it is possible for a gene to belong to several groups, both enriched and non-enriched. Within each gene group returned by DAVID, it will be necessary for the genes to be ranked by, for example, fold change (Table 2).

An alternative would be to consider gene-set enrichment analysis (GSEA), a method which can evaluate transcriptomic data at the gene-set level (Subramanian et al. 2005). The gene sets used are based on prior biological knowledge derived from published works in co-expression experiments and biochemical pathways. Thousands of genes across multiple samples are assigned to one of two classes (e.g., responders and nonresponders to treatment) creating phenotypic gene-expression profiles. Differential expression between phenotypic classes determines the order of a gene in a ranked list. GSEA aims to determine if the genes within a gene list occur near the top or bottom of the ranked gene list. Positive and negative enrichment to the phenotypic class depends on how high or low the genes within gene-sets appear in the ranked list. As all member genes of a gene set are typically not involved in a biological processes, the accompanying leading-edge analysis can be useful in determining the core genes driving the enrichment score in highly scoring gene-sets.

Survival Analysis

Survival analysis, in particular Cox proportional hazards (PH) regression, can help identify genes that are associated outcomes such as recurrence or death, i.e., time to event analysis. Good and poor prognosis labels can be used, but as with response/nonresponse, there will be variations of strength with respect to "good" and "poor." Univariate Cox PH regression (with multiple comparison adjustment and assumptions testing) can select those genes whose expression levels are associated with the selected event/nonevent type. If used in continuous form (preferred), the associated hazard ratios will be with respect to each unit increase in a gene's expression level (the nonparametric equivalent, log-rank, can be used for categorical variables but does not return a hazard ratio, only significance) (Table 2).

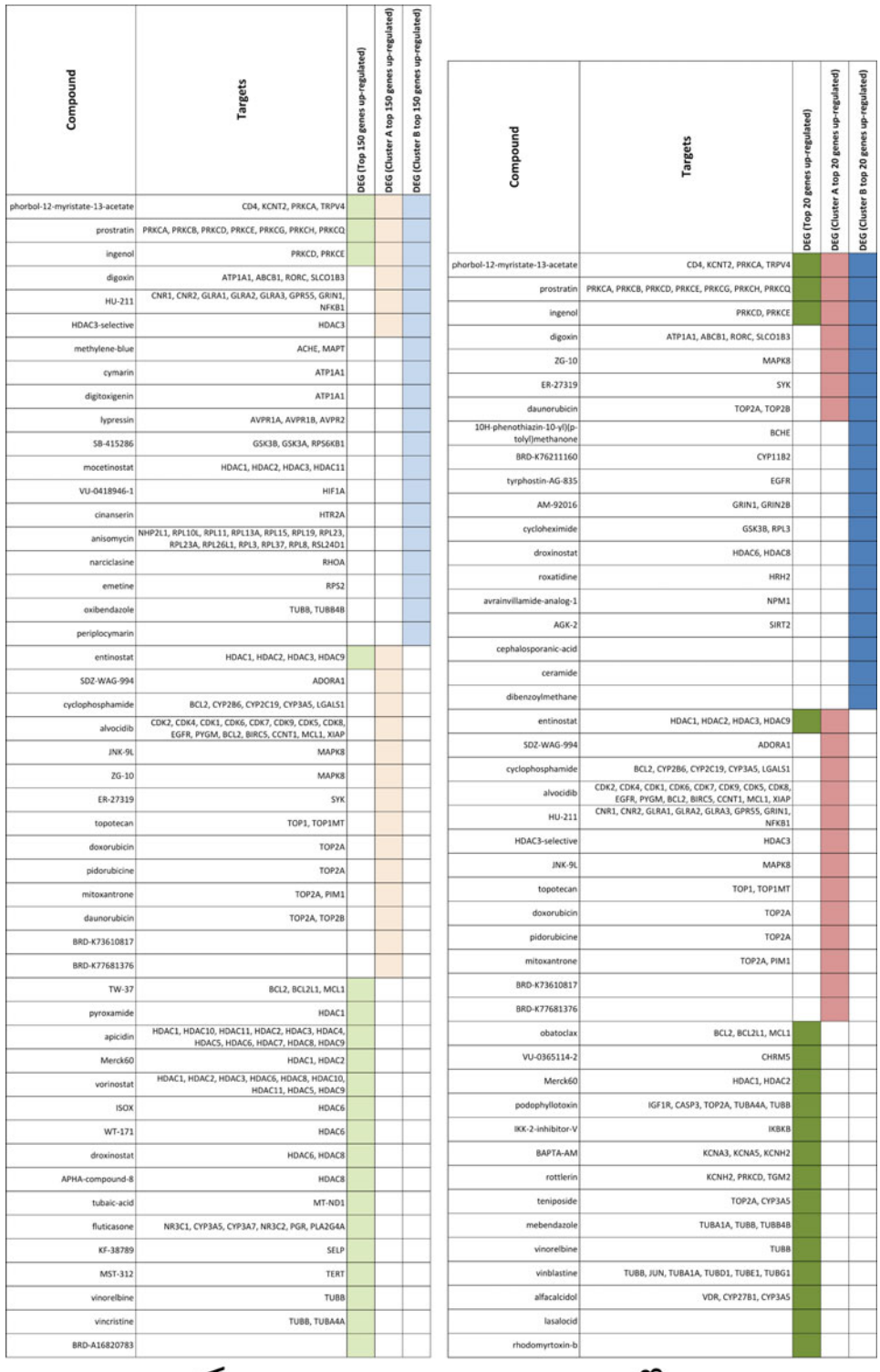
Case Study

A dataset consisting of gene expression profiles of eight patients with HPV-negative head and neck cancer and four normal controls (Tomar et al. 2016) (GEO accession number: GSE55549). This was selected for ease of replicability due to small sample size. Using GEO2R, a differential expression analysis was carried out via a moderated t-test. Applying a false discovery rate of 0.05, the top 150 and 20 (by fold change size) upregulated genes were selected and used as input into the CLUE connectivity mapping online tool. Next, to mimic molecular subtyping, the 5,000 most variable genes were identified in the dataset and hierarchical clustering performed using Euclidean/Ward.D2. Two clusters within the cancer samples were identified and each cluster compared to control samples using differential expression analysis. The top 150 and 20 upregulated genes were used as input into the CLUE connectivity mapping online tool as before. The top 20 compounds were returned for each of the connectivity mapping analyses and compared for overlap, considering gene list size and molecular subtype (Fig. 3).

Gene List Size

Considering the connectivity mapping analyses resulting from the top 150 and top 20 genes from the full differential expression analysis (comparing eight cancer to four normal samples), six compounds, phorbol-12-myristate-13-acetate, Merck60, entinostat, prostratin, ingenol, and vinorelbine, were found to be in common (Fig. 3a and b).

Using the differential expression analysis from the larger cluster ($n = 5$) versus control samples as input, both the larger and smaller versions of the gene lists returned the same compounds: phorbol-12-myristate-13-acetate,



A

Compound	Targets	DEG (Top 150 genes up-regulated)	DEG (Cluster A top 150 genes up-regulated)	DEG (Cluster B top 150 genes up-regulated)
phorbol-12-myristate-13-acetate	CD4, KCNT2, PRKCA, TRPV4			
prostratin	PRKCA, PRKCB, PRKCD, PRKCE, PRKCG, PRKCH, PRKCQ			
ingenol	PRKCD, PRKCE			
digoxin	ATP1A1, ABCB1, RORC, SLCO1B3			
HU-211	CNR1, CNR2, GLRA1, GLRA2, GLRA3, GPR55, GRIN1, NFKB1			
HDAC3-selective	HDAC3			
methylene-blue	ACHE, MAPT			
cymarin	ATP1A1			
digitoxigenin	ATP1A1			
lypressin	AVPR1A, AVPR1B, AVPR2			
SB-415286	GSK3B, GSK3A, RPS6KB1			
mocetinostat	HDAC1, HDAC2, HDAC3, HDAC11			
VU-0418946-1	HIF1A			
cinanserin	HTR2A			
anisomycin	NHP2L1, RPL10L, RPL11, RPL13A, RPL15, RPL19, RPL23, RPL23A, RPL26L1, RPL3, RPL37, RPL8, RSL24D1			
narciclasine	RHOA			
emetine	RPS2			
oxibendazole	TUBB, TUBB4B			
periplocymarin				
entinostat	HDAC1, HDAC2, HDAC3, HDAC9			
SDZ-WAG-994	ADORA1			
cyclophosphamide	BCL2, CYP2B6, CYP2C19, CYP3A5, LGALS1			
alvocidib	CDK2, CDK4, CDK1, CDK6, CDK7, CDK9, CDK5, CDK8, EGFR, PYGM, BCL2, BIRC5, CCNT1, MCL1, XIAP			
JNK-9L	MAPK8			
ZG-10	MAPK8			
ER-27319	SYK			
topotecan	TOP1, TOP1MT			
doxorubicin	TOP2A			
pidorubicine	TOP2A			
mitoxantrone	TOP2A, PIM1			
daunorubicin	TOP2A, TOP2B			
BRD-K73610817				
BRD-K77681376				
TW-37	BCL2, BCL2L1, MCL1			
pyroxamide	HDAC1			
apicidin	HDAC1, HDAC10, HDAC11, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9			
Merck60	HDAC1, HDAC2			
vorinostat	HDAC1, HDAC2, HDAC3, HDAC6, HDAC8, HDAC10, HDAC11, HDAC5, HDAC9			
ISOX	HDAC6			
WT-171	HDAC6			
droxinostat	HDAC6, HDAC8			
APHA-compound-8	HDAC8			
tubaic-acid	MT-ND1			
fluticasone	NR3C1, CYP3A5, CYP3A7, NR3C2, PGR, PLA2G4A			
KF-38789	SELP			
MST-312	TERT			
vinorelbine	TUBB			
vincristine	TUBB, TUBA4A			
BRD-A16820783				

B

Compound	Targets	DEG (Top 20 genes up-regulated)	DEG (Cluster A top 20 genes up-regulated)	DEG (Cluster B top 20 genes up-regulated)
phorbol-12-myristate-13-acetate	CD4, KCNT2, PRKCA, TRPV4			
prostratin	PRKCA, PRKCB, PRKCD, PRKCE, PRKCG, PRKCH, PRKCQ			
ingenol	PRKCD, PRKCE			
digoxin	ATP1A1, ABCB1, RORC, SLCO1B3			
ZG-10	MAPK8			
ER-27319	SYK			
daunorubicin	TOP2A, TOP2B			
10H-phenothiazin-10-yl)(p-tolyl)methanone	BCHE			
BRD-K76211160	CYP11B2			
tyrphostin-AG-835	EGFR			
AM-92016	GRIN1, GRIN2B			
cycloheximide	GSK3B, RPL3			
droxinostat	HDAC6, HDAC8			
roxatidine	HRH2			
avrainvillamide-analog-1	NPM1			
AGK-2	SIRT2			
cephalosporanic-acid				
ceramide				
dibenzoylmethane				
entinostat	HDAC1, HDAC2, HDAC3, HDAC9			
SDZ-WAG-994	ADORA1			
cyclophosphamide	BCL2, CYP2B6, CYP2C19, CYP3A5, LGALS1			
alvocidib	CDK2, CDK4, CDK1, CDK6, CDK7, CDK9, CDK5, CDK8, EGFR, PYGM, BCL2, BIRC5, CCNT1, MCL1, XIAP			
HU-211	CNR1, CNR2, GLRA1, GLRA2, GLRA3, GPR55, GRIN1, NFKB1			
HDAC3-selective	HDAC3			
JNK-9L	MAPK8			
topotecan	TOP1, TOP1MT			
doxorubicin	TOP2A			
pidorubicine	TOP2A			
mitoxantrone	TOP2A, PIM1			
BRD-K73610817				
BRD-K77681376				
obatoclax	BCL2, BCL2L1, MCL1			
VU-0365114-2	CHRM5			
Merck60	HDAC1, HDAC2			
podophyllotoxin	IGF1R, CASP3, TOP2A, TUBA4A, TUBB			
IKK-2-inhibitor-V	IKBKB			
BAPTA-AM	KCNA3, KCNA5, KCNH2			
rottlerin	KCNH2, PRKCD, TGM2			
teniposide	TOP2A, CYP3A5			
mebendazole	TUBA1A, TUBB, TUBB4B			
vinorelbine	TUBB			
vinblastine	TUBB, JUN, TUBA1A, TUBD1, TUBE1, TUBG1			
alfacalcidol	VDR, CYP27B1, CYP3A5			
lasalocid				
rhodomyrtoxin-b				

Fig. 3 Compounds returned from CLUE/LINCS connectivity analysis illustrating the effect of different gene list derivation methods on an example dataset of head and neck cancers versus normal samples including (i) differential expression analysis of full cancer cohort versus normal samples and (ii) two clusters (representing molecular subtypes) within cancer cohort and comparison of each to normal samples. In both cases, the top (**a**) 150 and top (**b**) 20 upregulated genes were chosen as input

entinostat, prostratin, ingenol, SDZ-WAG-994, digoxin cyclophosphamide, alvocidib, HU-211, HDAC3-selective, JNK-9L, ZG-10, ER-27319, topotecan, doxorubicin, epirubicin, mitoxantrone, daunorubicin, BRD-K73610817, and BRD-K77681376. With respect to the connectivity mapping results obtained using different lengths of gene lists from the differential analysis of the smaller cluster ($n = 3$) versus the control samples, four compounds were found to be overlapping: phorbol-12-myristate-13-acetate, prostratin, ingenol, and digoxin (Fig. 3a and b).

Molecular Subtype

For both the 20 gene and 150 gene lists as input, three compounds overlapped between the three connectivity mapping analyses, full cohort, larger cluster, and smaller cluster (all vs. controls): phorbol-12-myristate-13-acetate, prostratin, and ingenol. Considering the overlap between cluster results, when using the 150 (20) gene list, there were six (seven) overlapping compounds between the connectivity mapping analyses of the larger and smaller clusters, both with respect to control samples. Next considering which, if any, of the clusters drove the results for the full cohort, only entinostat was returned in the analyses of the largest cluster (vs. control samples), but not in the smaller cluster version, using both the top 150 and 20 gene lists (Fig. 3a and b).

There are limitations of this example study, i.e., it was only sufficiently powered to detect relatively large (0.5) effect sizes (0.11, 0.10, and 0.08 at 0.05 significance for differential expression using 8 vs. 4 samples, 5 vs. 4 samples, and 3 vs. 4 samples respectively). Furthermore, the clusters identified to represent molecular subtypes may not carry any prognostic weight (due to lack of associated clinic-pathological data). This example was designed to demonstrate that gene list derivation (full cohort versus biologically different sub-cohorts) and gene list size (top 150 vs. top 20 upregulated) could alter the compounds returned. Although the nature of the compounds was of less relevance, as this was an illustrative case, it was interesting to note the number of HDAC inhibitors ($n = 11$) returned (He et al. 2019). Although EGFR is upregulated in HPV-negative head and neck cancer, only one EGFR-targeting compound, tyrphostin-AG-835 was selected. PKC inhibition (returned in all six analyses) has been associated with increased response rates (Adams et al. 2015).

Summary

The case study was chosen to act as an illustrative example demonstrating how different approaches to gene list derivation can impact on compounds returned using a compound-phenotype match tool. While the previous drug repurposing cases have been largely due to happenstance, in silico tools matching compound to phenotype and compound to mechanism of action can help reduce the search space from among thousands of compounds or small molecules. Despite extensive reference library and robust statistical approaches, these tools can still be considered as exploratory in nature. Each method may return a different set of compounds for the same query data, determining the accuracy of predictions can be only be determined in in vivo and in vitro settings. The input data used for such tools is pivotal, particularly from the user's perspective. Spending time on developing a representative gene list or lists may well save time and precious resources later on. The development of a gene list may also depend on the underlying structure and statistical modelling used by each compound mapping method, in particular the direction of comparison used – treated versus untreated or untreated versus treated cell lines. A multidisciplinary approach to using such tools is therefore key, incorporating input from clinicians, biologists, and bioinformaticians in the curation of datasets, developing a gene list and the interpretation of results. Likewise, the choice of compounds in the subsequent validation within in vivo and in vitro models may depend on a number of factors, including prior knowledge. The answers to future drug repurposing questions may well be found within such tools; however, knowing what questions to ask of the data is the rate-limiting step.

References and Further Reading

Adams RA, D'Souza MMA, Pierce CJ et al (2015) Ectopic expression of protein kinase C-β sensitizes head and neck squamous cell carcinoma to diterpene esters. Anticancer Res 35:1291–1296

Alex A, Harris J, Smith DA (2015) Attrition in the pharmaceutical industry: reasons, implications, and pathways forward. Wiley, Hoboken, pp 1–356. https://doi.org/10.1002/9781118819586

Ashburn TT, Thor KB (2004) Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 3:673–683

Aslam B, Wang W, Arshad MI et al (2018) Antibiotic resistance: a rundown of a global crisis. Infect Drug Resist 11:1645–1658. https://doi.org/10.2147/IDR.S173867

Athar A, Füllgrabe A, George N et al (2019) ArrayExpress update – from bulk to single-cell expression data. Nucleic Acids Res 47:D711–D715. https://doi.org/10.1093/nar/gky964

Boolell M, Allen MJ, Ballard SA et al (1996) Sildenafil: an orally active type 5 cyclic GMP-specific phosphodiesterase inhibitor for the treatment of penile erectile dysfunction. Int J Impot Res 8:47–52

Broad Institute (2019) CLUE (Connectivity Map). https://clue.io/

Camidge DR, Pao W, Sequist LV (2014) Acquired resistance to TKIs in solid tumours: learning from lung cancer. Nat Rev Clin Oncol 11:473–481. https://doi.org/10.1038/nrclinonc.2014.104

Chae YK, Arya A, Malecek M-K et al (2016) Repurposing metformin for cancer treatment: current clinical studies. Oncotarget 7. https://doi.org/10.18632/oncotarget.8194

Chiang AP, Dudley JT, Shenoy M et al (2011) Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. Sci Transl Med 3:96ra76–96ra76. https://doi.org/10.1126/scitranslmed.3002648

Edgar R, Domrachev MLA (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207–210

Evans JMM, Donnelly LA, Emslie-Smith AM et al (2005) Metformin and reduced risk of cancer in diabetic patients. Br Med J 330:1304–1305. https://doi.org/10.1136/bmj.38415.708634.F7

Karapetis CS, Khambata-Ford S, Jonker DJ et al (2008) K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med 359(17):1757–65. https://doi.org/10.1056/NEJMoa0804385

Gerhard I, Schindler AE, Buhler K et al (1992) Treatment of endometriosis with leuprorelin acetate depot: a German multicentre study. Clin Ther 14:3–16

Guinney J, Dienstmann R, Wang X et al (2015) The consensus molecular subtypes of colorectal cancer. Nat Med 21:1350–1356. https://doi.org/10.1038/nm.3967

He L, Gao L, Shay C et al (2019) Histone deacetylase inhibitors suppress aggressiveness of head and neck squamous cell carcinoma via histone acetylation-independent blockade of the EGFR-Arf1 axis 06 Biological Sciences 0601 Biochemistry and Cell Biology. J Exp Clin Cancer Res 38. https://doi.org/10.1186/s13046-019-1080-8

Hoadley KA, Yau C, Hinoue T et al (2018) Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell 173:291–304.e6. https://doi.org/10.1016/j.cell.2018.03.022

Holzinger ER, Ritchie MD (2012) Integrating heterogeneous high-throughput data for meta-dimensional pharmacogenomics and disease-related studies. Pharmacogenomics 13:213–222. https://doi.org/10.2217/pgs.11.145

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57. https://doi.org/10.1038/nprot.2008.211

Huang S, Chaudhary K, Garmire LX (2017) More is better: recent progress in multi-omics data integration methods. Front Genet 8. https://doi.org/10.3389/fgene.2017.00084

Iorio F, Bosotti R, Scacheri E et al (2010) Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci 107:14621–14626. https://doi.org/10.1073/pnas.1000138107

Kaitin KI, Dimasi JA (2011) Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000–2009. Clin Pharmacol Ther 89:183–188. https://doi.org/10.1038/clpt.2010.286

Kim W-J, Shim M-S, Kim J-Y (2013) Management of hyperglycemia in type 2 diabetes: a patient-centered approach developed by the American Diabetes Association and the European Association for the Study of Diabetes. J Korean Diab 13:172. https://doi.org/10.4093/jkd.2012.13.4.172

Laboratory of Human Retrovirology and Immunoinformatics (LHRI) (2019) DAVID. https://david.ncifcrf.gov/. Accessed 26 Jul 2019

Laifenfeld D, Cha Y, Erez T et al (2017) Drug repurposing from the perspective of pharmaceutical companies. Br J Pharmacol. https://doi.org/10.1111/bph.13798

Lamb J, Crawford ED, Peck D et al (2006) The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science (80-) 313:1929–1935. https://doi.org/10.1126/science.1132939

Lee JH, Kim DG, Bae TJ et al (2012) CDA: combinatorial drug discovery using transcriptional response modules. PLoS One 7. https://doi.org/10.1371/journal.pone.0042573

Liu F, Yan L, Wang Z et al (2017) Metformin therapy and risk of colorectal adenomas and colorectal cancer in type 2 diabetes mellitus patients: a systematic review and meta-analysis. Oncotarget 8. https://doi.org/10.18632/oncotarget.13762

Malcomson B, Wilson H, Veglia E et al (2016) Connectivity mapping (ssCMap) to predict A20-inducing drugs and their antiinflammatory action in cystic fibrosis.

Proc Natl Acad Sci 113:E3725–E3734. https://doi.org/10.1073/pnas.1520289113

Napolitano F, Sirci F, Carrella D, Di Bernardo D (2016) Drug-set enrichment analysis: a novel tool to investigate drug mode of action. Bioinformatics 32:235–241. https://doi.org/10.1093/bioinformatics/btv536

Napolitano F, Carrella D, Mandriani B et al (2018) Gene2drug: a computational tool for pathway-based rational drug repositioning. Bioinformatics 34:1498–1505. https://doi.org/10.1093/bioinformatics/btx800

NCBI (2019) GEO2R. https://www.ncbi.nlm.nih.gov/geo/geo2r/. Accessed 26 Jul 2019

O'Reilly P (2016) QUADrATiC: software. http://go.qub.ac.uk/QUADrATiC

Peyvandipour A, Saberian N, Shafi A et al (2018) Systems biology: a novel computational approach for drug repurposing using systems biology. Bioinformatics 34:2817–2825. https://doi.org/10.1093/bioinformatics/bty133

Raghavan R, Hyter S, Pathak HB et al (2016) Drug discovery using clinical outcome-based connectivity mapping: application to ovarian cancer. BMC Genomics 17. https://doi.org/10.1186/s12864-016-3149-5

Rena G, Hardie DG, Pearson ER (2017) The mechanisms of action of metformin. Diabetologia 60:1577–1585. https://doi.org/10.1007/s00125-017-4342-z

Ritchie MD, Holzinger ER, Li R et al (2015) Methods of integrating data to uncover genotype-phenotype interactions. Nat Rev Genet 16:85–97. https://doi.org/10.1038/nrg3868

Schally A, Kastin A, Coy D (1976) Edward T. Tyler Prize Oration: LH-releasing hormone and its analogues: recent basic and clinical investigations. Int J Fertil 21:1–30

Setoain J, Franch M, Martínez M et al (2015) NFFinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. Nucleic Acids Res 43:W193–W199. https://doi.org/10.1093/nar/gkv445

Subramanian A, Tamayo P, Mootha VK et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci 102:15545–15550. https://doi.org/10.1073/pnas.0506580102

Subramanian A, Narayan R, Corsello SM et al (2017) A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell 171:1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049

The Comprehensive R Network (2019) The Comprehensive R Archive Network. https://cran.r-project.org/. Accessed 26 Jul 2019

Tomar S, Graves C, Altomare D et al (2016) Human papillomavirus status and gene expression profiles of oropharyngeal and oral cancers from European American and African American patients. Head Neck 38:E694–E704. https://doi.org/10.1002/hed.24072

University of California SC (2016) Xena. https://tcga.xenahubs.net. Accessed 26 Jul 2019

Wellcome Sanger Institute (2019) GDSC. http://www.cancerrxgene.org. Accessed 27 Jul 2019

Wu H, Miller E, Wijegunawardana D et al (2017) MD-Miner: a network-based approach for personalized drug repositioning. BMC Syst Biol 11. https://doi.org/10.1186/s12918-017-0462-9

Yang W, Soares J, Greninger P et al (2013) Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res 41. https://doi.org/10.1093/nar/gks1111

Yuen HF, McCrudden CM, Huang YH et al (2013) TAZ expression as a prognostic indicator in colorectal cancer. PLoS One 8. https://doi.org/10.1371/journal.pone.0054211

Zhang S-D, Gant TW (2008) A simple and robust method for connecting small-molecule drugs using gene-expression signatures. BMC Bioinf 9:258. https://doi.org/10.1186/1471-2105-9-258

Zhang M, Lee S, Yao B et al (2019) DIGREM: an integrated web-based platform for detecting effective multi-drug combinations. Bioinformatics 35:1792–1794. https://doi.org/10.1093/bioinformatics/bty860

Toxicokinetics and Safety Ratios

52

Ferdi Sombogaard

Contents

Abstract

Pharmacokinetics is essentially the study of "how a substance gets into the body and what happens to it in the body." Similarly, toxicokinetics deals with what the body does with a drug or other substance when given a relatively high dose relative to the therapeutic dose. Toxicokinetic studies are generally carried out at much higher doses than those used in pharmacokinetic studies, and this dose information is critical for predicting the safety of substances. The primary objective of toxicokinetics is to describe the systemic exposure achieved in animals and its

F. Sombogaard (✉)
Department of Clinical Pharmacology and Pharmacy, Amsterdam University Medical Centers – VUmc, Amsterdam, The Netherlands
e-mail: F.Sombogaard@amsterdamumc.nl

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_21

relationship to dose level and the time course of the toxicity study. The ratio of drug exposure in animals at the no observed adverse effect level (NOAEL) and in humans at the expected therapeutic dose is one of the precautionary principles to determine the risk benefit profile of pharmaceuticals.

Purpose and Rationale

Pharmacokinetics is essentially the study of "how a substance gets into the body and what happens to it in the body." Similarly, toxicokinetics deals with what the body does with a drug or other substance when given a relatively high dose relative to the therapeutic dose. Toxicokinetic studies are generally carried out at much higher doses than those used in pharmacokinetic studies, and this dose information is critical for predicting the safety of substances (Welling 1995).

The primary objective of toxicokinetics is to describe the systemic exposure achieved in animals and its relationship to dose level and the time course of the toxicity study. Secondary objectives are:

- To relate the exposure achieved in toxicity studies to toxicological findings and contribute to the assessment of the relevance of these findings to clinical safety
- To support the choice of species and treatment regimen in nonclinical toxicity studies
- To provide information which, in conjunction with the toxicity findings, contributes to the design of subsequent nonclinical toxicity studies

These data may be used in the interpretation of toxicology findings and their relevance to clinical safety issues (ICH Guidance on Toxicokinetics 1994).

The ratio of drug exposure in animals at the no observed adverse effect level (NOAEL) and in humans at the expected therapeutic dose is one of the precautionary principles to determine the risk benefit profile of pharmaceuticals. For this ratio the expressions "safety ratio" and "safety margin" were also used. It is usually based on area under plasma concentration-time curve (AUC) – which is a measure of the exposure – between animal and human (animal/human AUC ratio). However, depending on the mode of action or whichever is smaller, the ratio can also be based on the maximum concentration in plasma (animal/human C_{max} ratio).

Procedure

Number of Animals and Time Points

In the ICH Guidance on Toxicokinetics (1994), it is stated that "the number of animals to be used should be the minimum consistent with generating adequate toxicokinetic data" and that "the area under the matrix level concentration-time curve and/or the measurement of matrix concentrations at the expected peak-concentration time C_{max}, or at some other selected time C(time), are the most commonly used parameters." In large animals (like dogs), the number of animals is usually fixed by the number of animals that are necessary for safety evaluation. The withdrawal of a sufficient number of blood samples (six to nine) per animal is not a problem. However, in small animals like rodents, it is recommended not to collect more than 10% of the blood volume during the AUC sampling interval (BVA/FRAME/RSPCA/UFAW Working Group of Refinement 1993; Cayen 1995). A good practice guide to the administration of substances and removal of blood, including routes and volumes, is described in the similar paper (Diehl et al. 2001). The authors limit the total daily volumes of multiple sampling to 7.5% of the circulatory blood volume at a recovery period of 1 week, 10–15% at a recovery period of 2 weeks, and 20% at a recovery period of 3 weeks. The optimum number of time points is always a compromise between blood volume restrictions and reliable assessment of toxicokinetic parameters (AUC and C_{max}) (Diehl et al. 2001).

A more recent report of the European Partnership for Alternative Approaches to Animal Testing urges the regulatory acceptance of alternatives to animal testing: networking and communication (including cross-sector collaboration, international cooperation, and harmonization), involvement of regulatory agencies from the initial stages

of test method development, and certainty on prerequisites for test method acceptance including the establishment of specific criteria for regulatory acceptance. Data sharing and intellectual property issues affect many aspects of test method development, validation, and regulatory acceptance. In principle, all activities should address replacement, reduction, and refinement methods (albeit animal testing is generally prohibited in the cosmetics sector) (Ramirez et al. 2015).

The ICH "Focus on microsampling" Implementation Working Group also stresses out the important contribution to 3R benefits (replacement, reduction, and refinement) by reducing or eliminating the need for toxicokinetic animals. Nevertheless, the today's common analytical method sensitivity (such as that of liquid chromatography/mass spectrometry) has been improved, allowing microsampling techniques (very low volume sampling) to be widely used in toxicokinetic assessment (ICH Focus on microsampling 2017). For three different compounds, Pai et al. (1996) compared the AUCs from intensive (full) (10–15 time points with 4 to 5 rats per time point) sampling schemes with sparse sampling schemes (5 time points with 2 rats per time point). Using Monte Carlo simulation, Pai et al. (1996) could show that the deviation of AUC estimation of the sparse sampling scheme from the full sampling scheme was not larger than 10%. Thus it is seen that a sparse sampling scheme with five to seven time points with two to three animals per time point is well suited for the reliable determination of systemic exposure in small animal toxicity studies.

Main Group or Satellite Animals?

Whenever possible, toxicokinetic measurements are performed on all the animals in the toxicity study. This is the most representative approach, and it allows the individual pharmacokinetic data to be directly correlated with the toxicological findings. The second choice is toxicokinetic measurement in representative subgroups or satellite groups. Satellite groups are animals that are treated and housed under conditions identical to those of the main study animals. The use of satellite animals is indicated, for example, in small animals, where the collection of a relatively large volume of blood may influence the toxicological findings.

Analytical Methods

Integration of pharmacokinetics into toxicity testing implies early development of analytical methods for which the choice of analytes and matrices should be continually reviewed as information is gathered on metabolism and species differences. The analytical methods to be used in toxicokinetic studies should be specific for the entity to be measured and of adequate accuracy and precision. The limit of quantification should be adequate for the measurement of the range of concentrations anticipated to occur in the generation of the toxicokinetic data (ICH Guidance on Toxicokinetics 1994). The today's common analytical methods, such as ultra(high)-performance liquid chromatography (U(H)PLC) and liquid chromatography-mass spectrometry (LC-MS), have excellent performance of sensitivity and specificity, and the limit of quantification can be low to very low (Lee and Kerns, 1999).

Toxicokinetic Evaluation

The following aspects should be considered for toxicokinetic evaluation:

- Pharmacokinetic profile of the compound (exposure)
- Dose dependency of AUC and C_{max}
- Changes of exposure during the course of the toxicity study
- Sex differences

Pharmacokinetic Profile of the Compound (Exposure)

For toxicokinetic purposes it is usually sufficient to describe the systemic burden in plasma or serum of the test species with the test compound and/or its metabolites. The AUC and/or the measurement of matrix concentrations at the expected peak-concentration time, C_{max}, or at some other

selected time (e.g., $C_{(24h)}$ as trough value), $C_{(time)}$, is the most commonly used parameter. According to the supplementary notes in the ICH Guidance on Toxicokinetics (1994), for a profile (e.g., four to eight) matrix, samples during a dosing interval should be taken to make an estimate of C_{max} and/or $C_{(time)}$ and the AUC.

Dose Dependency of AUC and C_{max}

According to the ICH Guidance on Toxicokinetics (1994), it is one of the primary objectives of toxicokinetics to describe the systemic exposure achieved in animals and its relationship to dose level. At pharmacological and clinical doses, it can be generally assumed that most of the drugs show linear pharmacokinetics. Linear pharmacokinetics is given when exposure (AUC) is proportional to dose, and principal pharmacokinetic parameters like bioavailability, elimination rate, volume of distribution, and clearance are independent of dose. For toxicokinetic studies, however, nonlinear pharmacokinetics is much more frequent than linear pharmacokinetics. This in mainly due to the fact that at very high doses, most systems in the body are likely to be stressed and, possibly, saturated to some degree. Thus, in addition to the dose-proportional increase of exposure (Fig. 1), a less than proportional increase (Fig. 2) and a more than proportional increase (Fig. 3) are very frequently observed. For a less than proportional increase of exposure, a saturation of absorption processes or a concentration-dependent change of volume of distribution should be considered as potential causes. For a more than proportional increase of exposure, saturation of metabolic elimination pathways, saturation of renal or biliary excretion of parent compound, and a concentration-dependent change of volume of distribution should be considered as potential causes.

Changes of Exposure During the Course of the Toxicity Study

According to the ICH Guidance on Toxicokinetics (1994), the description of the relationship of exposure to the time course of the toxicity study belongs to the primary objectives of toxicokinetics. This objective may be achieved by deriving pharmacokinetic parameters from measurements made at appropriate time points during the course of the individual studies. In short-term studies (1 month or shorter), day 1 and a day at the end of the toxicity study may be appropriate profiling days. In long-term studies, day 1, a day after one third of the study duration, and a day at the end of the toxicity study may be appropriate sampling days. Increasing exposure may occur during the course of a study for those compounds that have a particularly long plasma half-life. Conversely, unexpectedly low exposure may occur during a study as a result

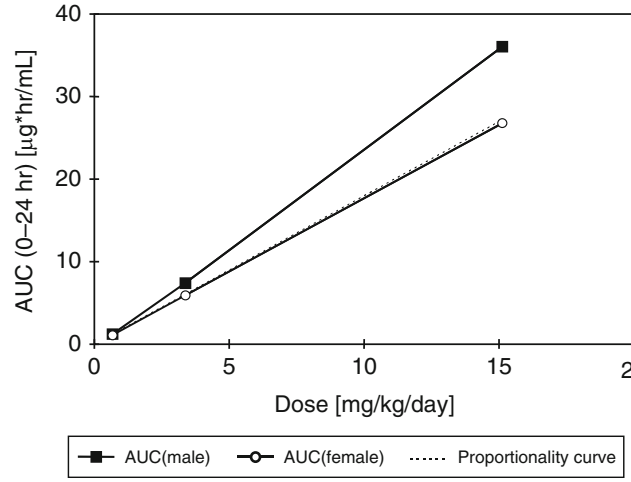


Fig. 1 Proportional dose dependency of AUC in an intravenous toxicity study in rat with the test compound A

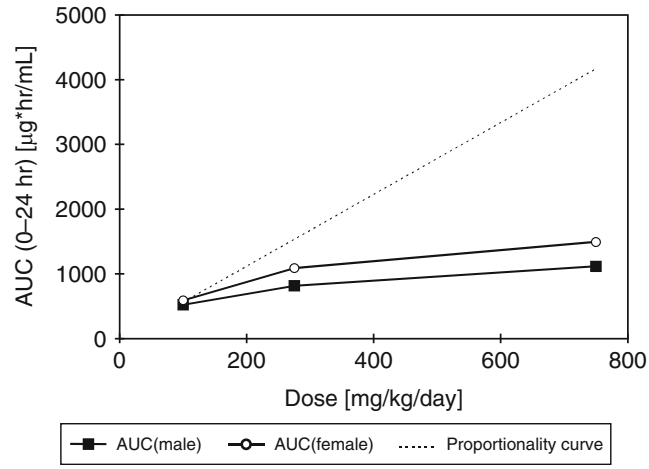


Fig. 2 Less than proportional dose dependency of AUC in an oral toxicity study in rat with the test compound B

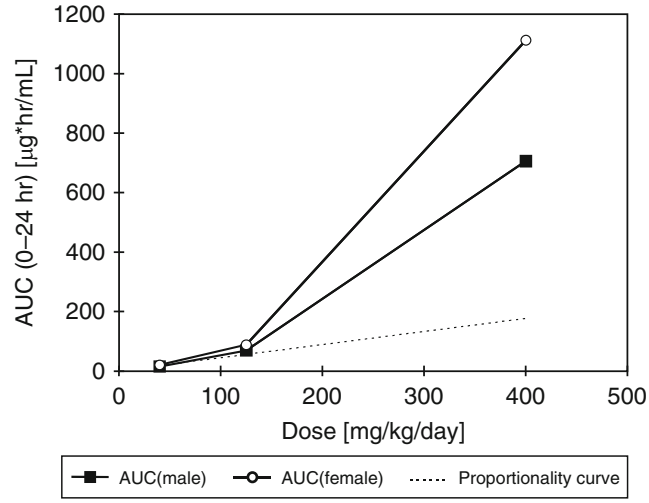


Fig. 3 More than proportional dose dependency of AUC in an oral toxicity study in dog with the test compound C

of auto-induction of metabolizing enzymes. However, other facts can also play a role in changes of exposure during the course of the study. Very often, rats and mice are used at an age at which they are not sexually mature, and during the study, sexual maturation takes place in the first 2 months, with its known impact on the rate and extent of metabolism. The harm of elimination pathways (e.g., nephro- or hepatoxicity) by the test compound can be another reason for changes in exposure. A more trivial reason such as aging or change of the administered batch with impact on bioavailability should also be considered.

Sex Differences

According to the ICH Guidance on Toxicokinetics (1994), it is normal to estimate exposure in animals of both sexes unless some justification can be made for not doing so. For evaluation both sexes should be evaluated separately. The assessment of exposure data of the two sexes can be performed by calculating the ratio of AUC, C_{max}, and elimination half-life in males and females. However, additional factors such as size of the investigated groups (with respect to random variation) and sexual maturity have to be considered. As a rule

of thumb, it can be stated that in rodents sex difference is quite common when CYP metabolism is involved as a major elimination pathway, whereas in non-rodents distinct sex differences are rather rare.

Evaluation

An important objective of toxicokinetics is to relate the no observed adverse effect dose level with the exposure (expressed as AUC or C_{max}) in the respective animal species at this dose level. From these data and from the exposure values in humans at the expected therapeutic dose, the multiple of the therapeutic exposure in human versus animal exposure at NOAEL is calculated according to the following formulas:

$$\textit{Animal/human AUC ratio} = \frac{AUC_{animal} \text{ at NOAEL}}{AUC_{human} \text{ at the maximum recommended human dose}}$$

$$\textit{Animal/human } C_{max} \textit{ ratio} = \frac{C_{max,animal} \text{ at NOAEL}}{C_{max,human} \text{ at the maximum recommended human dose}}$$

Comparison between animal and human exposure is generally based on AUC, but sometimes it may be more appropriate to use C_{max}. The synonyms "safety ratio" "safety margin" and "margin of safety" are frequently used for animal/human exposure ratio.

Most Sensitive Species

The animal/human ratios are always estimated in a conservative way, which means that the lowest exposure data (most sensitive animal species and sex) in animals and the human exposure data at the maximum recommended human dose (MRHD) are taken for calculating the ratio.

Protein Binding

The unbound drug in plasma is thought to be the most relevant indirect measure of tissue concentrations of unbound drug. The rules on how to deal with the protein-binding issue are clearly defined (Note 8 of the ICH Topic S1C(R2) 2008).While in vivo determinations of unbound drug might be the best approach, in vitro determinations of protein binding using parent and/or metabolites as appropriate (over the range of concentrations achieved in vivo in rodents and humans) might be used in the estimation of unbound AUC. When protein binding is low in both humans and rodents, or when protein binding is high and the unbound fraction of drug is greater in rodents than in humans, the comparison of total plasma concentration of drug is appropriate. When protein binding is high and the unbound fraction is greater in humans than in rodents, the ratio of the unbound concentrations should be used.

Steady-State Conditions

The AUC value used for the ratio calculation is generally AUC_{0-24} under steady-state condition for animals as well as for humans. Even if the drug is administered more than once daily to either species (e.g., the frequency of administration in laboratory animals may be increased compared to the proposed schedule for the human clinical studies in order to compensate for faster clearance rates or low solubility of the active ingredient), the exposure per day should be calculated and compared. For the rare cases in which the dosing interval is longer than 24 h, an appropriate calculation has to be performed and mentioned along with the value.

Duration of Treatment

Usually, data of several toxicity studies in the same species, but with different dosing duration, are available. The ratio estimation should be done in the most conservative way, which means that the lowest exposure data in animals under steady-state conditions should be used whenever the exposure is determined. It is not recommended that exposure be determined at the end of the life span, and exposure monitoring is not considered essential beyond 6 months (ICH Guidance on Toxicokinetics 1994). Careful attention should

be paid to the interpretation of toxicokinetic results in animals if there are already some toxicological findings. For example, very often it is not clear if the high exposure observed in these animals is the reason for their bad state or rather the consequence of it. Therefore, the direct link between observed exposure in already impaired animals and toxicological finding should be avoided.

Metabolites

Under the following circumstances, the measurement of metabolite concentrations and subsequent evaluation of animal/human exposure ratio is especially important (ICH Guidance on Toxicokinetics 1994):

- When the administered compound acts as a "prodrug" and the delivered metabolite is acknowledged to be the primary active entity
- When the compound is metabolized to one or more pharmacologically or toxicologically active metabolites, which could make a significant contribution to tissue/organ responses
- When the administered compound is very extensively metabolized and the measurement of plasma or tissue concentrations of a major metabolite is the only practical means of estimating exposure following administration of the compound in toxicity studies

In the issued FDA Guidance for Industry Safety Testing of Drug Metabolites (2008), it was additionally emphasized that it is crucial to gather toxicokinetic data from disproportionate metabolites in toxicity studies with direct dosing of the metabolite.

Biotechnology-Derived Pharmaceuticals

In the FDA Guidance for Industry S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals (1997), it is repeatedly emphasized that systemic exposure should be monitored during the toxicity studies. Thus, there is no difference compared to low-molecular-weight compounds. However, where a product has a lower affinity to, or potency in, the cells of the selected species than in human cells, testing of higher doses may be important. The multiples of the human dose that are needed to determine adequate safety margins may vary with each class of biotechnology-derived pharmaceutical and its clinical indication or indications. In addition, the effects of antibody formation on pharmacokinetic/pharmacodynamic (PK/PD) parameters, incidence and/or severity of adverse effects, complement activation, or the emergence of new toxic effects should be considered when interpreting the data.

Critical Assessment of the Method

Systemic Exposure as Surrogate for Exposure in All Other Tissues

The concept of safety margins based on exposure data is based on the assumption that plasma concentrations of a compound are the surrogate for exposure in all other tissues, including the target organ of toxicity. This approach is justified in the majority of cases. However, in some cases, the systemic exposure in plasma may go in the opposite direction to the specific exposure in the target organ. For example, strong first-pass hepatic extraction may increase the exposure in the target organ liver and concomitantly trigger the toxicity but decrease the systemic exposure in plasma. Another example for systemic exposure going in the opposite direction to target organ exposure was given by Lacy et al. (1998). Probenecid, a competitive inhibitor of organic anion transport in the proximal tubular epithelial cells, was evaluated for its effect on the chronic toxicity and pharmacokinetics of cidofovir in monkeys. Nephrotoxicity was present only in monkeys receiving cidofovir without probenecid. The coadministration of probenecid resulted in an inhibition of the active tubular secretion of cidofovir into the kidneys and concomitantly in a shift from local exposure in the kidney toward higher systemic exposure to cidofovir (as measured by AUC in plasma). The decrease of specific exposure in the kidneys is most likely the reason for the protection against nephrotoxicity.

How to Deal with Small Safety Factors

Generally, an exposure safety margin between the clinical dose and the animal NOAEL of tenfold would be acceptable (EMEA CHMP SWP Reflection Paper on PPARs 2006). However, from the information assessed, this may be unlikely. A retrospective analysis was performed on data from carcinogenicity studies of therapeutics conducted at the maximum tolerated dose (MTD) for which there was sufficient human and rodent pharmacokinetic data for comparison of AUC values (ICH Topic S1C(R2) 2008). In 35 drug carcinogenicity studies carried out at the MTD for which there were adequate pharmacokinetic data available in rats and humans, approximately one third had a relative systemic exposure ratio less than or equal to 1, and another one third had ratios between 1 and 10. In these cases, the following precautionary principles to determine the risk benefit profile should apply.

- Can reversibility of effects be demonstrated in repeated dose toxicity studies that include a drug-free period, which may provide reassurance that the findings will not be irreversible?
- If good mechanistic data for toxic effects are available, it may help in the assessment of relevance to human safety.
- A smaller safety factor might also be used when toxicities produced by the therapeutic are easily monitored by relevant and valid biomarkers, are predictable, and exhibit a moderate-to-shallow dose–response relationship with toxicities that are consistent across the tested species (both qualitatively and with respect to appropriately scaled dose and exposure).

A predicted safety margin close to 1 or even less in a clinical dose escalation study does not necessarily force a stop to the trial but requires a slower dose progression.

References and Further Reading

BVA/FRAME/RSPCA/UFAW Working Group of Refinement (1993) Removal of blood from laboratory mammals and birds. Lab Anim 27:1–22

Cayen MN (1995) Considerations in the design of toxicokinetic programs. Toxicol Pathol 23(2): 148–157

Diehl K-H, Hull R, Morton D, Pfister R, Rabemampianina Y, Smith D, Vidal J-M, van de Vorstenbosch C (2001) A good practice guide to the administration of substances and removal of blood, including routes and volumes. J Appl Toxicol 21:15–23

EMEA CHMP SWP Reflection Paper on PPARs (peroxisome proliferator activated receptors); Doc Ref: EMEA/341972/2006 (2006)

FDA Guidance for Industry S6 Preclinical Safety Evaluation of Biotechnology- Derived Pharmaceuticals (1997)

FDA Guidance for Industry Safety Testing of Drug Metabolites (2008)

ICH Focus on Microsampling (2017) Note for guidance on toxicokinetics: the assessment of systemic exposure – focus on microsampling (CHMP/ICH/320985/2016); adopted by CPMP December 2017

ICH Guidance on Toxicokinetics (1994) ICH topic S3 A; toxicokinetics: a guidance for assessing systemic exposure in toxicology studies (CPMP/ICH/384/95); approval by CPMP November 1994

ICH Topic S1C(R2) (2008) Dose selection for carcinogenicity of pharmaceuticals (EMEA/CPMP/ICH/383/1995); approval by CPMP April 2008

Lacy SA, Hitchcock MJM, Lee WAL, Tellier P, Cundy KC (1998) Effect of oral probenecid coadministration on the chronic toxicity and pharmacokinetics of intravenous cidofovir in cynomolgus monkeys. Toxicol Sci 44:97–106

Lee MS, Kerns EH (1999) LC/MS applications in drug development. Mass Spectrom Rev 18:187–279

Pai SM, Fettner SH, Hajian G, Cayen MN, Batra VK (1996) Characterization of AUCs from sparsely sampled populations in toxicology studies. Pharm Res 13 (9):1283–1290

Ramirez T, Beken S, Chlebus M, Ellis G, Griesinger C, de Jonghe S, Manou I, Mehling A, Reisinger K, Rossi LH, van Benthem J, van der Laan JW, Weissenhorn R, Sauer UG (2015) Knowledge sharing to facilitate regulatory decision-making in regard to alternatives to animal testing: report of an EPAA workshop. Regul Toxicol Pharmacol 73:210–226

Welling PG (1995) Differences between pharmacokinetics and toxicokinetics. Toxicol Pathol 23:143–147

In/In Vivo Correlation for Transporters 53

Sandra Cvijic

Contents

Abstract

Membrane transporters play one of the key roles in drug disposition and drug-drug interactions. Therefore, one of the prerequisites in the development of new drugs and dosage forms is determination of potential transporter-mediated processes a drug may undergo in the organism. The importance of this subject has been recognized by the regulatory authorities who have issued relevant guidelines on the in vitro and clinical investigations of potential drug-transporter interactions. In vitro transporter studies are designed to determine

S. Cvijic (✉)
Department of Pharmaceutical Technology and Cosmetology, University of Belgrade-Faculty of Pharmacy, Belgrade, Serbia
e-mail: sandra.cvijic@pharmacy.bg.ac.rs; gsandra@pharmacy.bg.ac.rs

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_22

whether a drug is a substrate/inhibitor/inducer of the clinically relevant transporters and to describe the interaction in terms of kinetic parameters, with the final goal to provide information on the clinical outcomes of the observed interactions with transporters. A number of in vitro methodologies have been designed to provide information on drug-transporter interplay, and the most prominent ones will be described in this chapter. Approaches to extrapolate in vitro to in vivo data will also be addressed. However, it should be stressed that knowledge on transporters and transporter-mediated processes is constantly growing, and, as new technologies are emerging, transporter science is making its headway toward the application of advanced in vitro-in vivo-in silico assessment tools. Although being routinely performed for years now, investigation of drug-transporter interactions and in vitro-to-in vivo translational modeling are still evolving, and breakthrough may happen in the following years.

Abbreviations

BCRP	Breast cancer resistance protein
BSEP	Bile salt export pump
CAR	Constitutive androstane receptor
CHO	Chinese hamster ovary cell
CYP	Cytochrome P450
HEK293	Human embryonic kidney 293 cells
LLC-PK1	Lewis-lung cancer porcine kidney 1 cells
MATE	Multidrug and toxin extrusion protein
MCT1	Monocarboxylate transporter 1
MDCK	Madin-Darby canine kidney cells
MDR1	Multidrug resistance protein 1
MRP	Multidrug resistance-associated protein
OAT1	Organic anion transporter 1
OATP	Organic-anion-transporting polypeptide
OCT	Organic cation transporter
PEPT1	Peptide transporter 1
P-gp	P-glycoprotein
PXR	Pregnane X receptor

Introduction

Drug bioperformance is a complex phenomenon ruled by a number of factors and processes, including drug interference with membrane transporters. These specific proteins mediate drug transport through physiological barriers, including plasma membranes and membranes of intracellular structures, and therefore influence all ADME properties (absorption, distribution, metabolism, and elimination) of the drug. Consequently, they may impact therapeutic and adverse drug effects. Moreover, due to the intrinsic properties of transporters such as specificity and limited capacity, they play an important role in drug-drug interactions (DDIs) and drug-nutrients interactions. Thereof, it is important to establish reliable means to estimate all these effects.

There are different ways to classify membrane transporters. Based on the direction of the transmembrane transport, they are divided into influx (facilitate entry of a molecule into the cell) and efflux (carry a molecule out of the cytosol) transporters. In addition, transporter-facilitated movement of molecules can be classified as uniport (one-way direction of one molecular entity), symport (the same one-way direction of two different molecular entities), or antiport (transport of different molecular entities in different directions), although some transporters act in bidirectional way. Regarding the mechanism of transport, transporters are differentiated as passive or active. The first group functions in energy-independent manner and carry a substrate molecule down the concentration gradient, whereas the later one consumes energy and enables movement of a substrate against the concentration gradient. This last group comprises ATP-binding cassette (ABC) transporter family, which consists of primary-active transporters (since they use energy directly, commonly from ATP hydrolysis). These are mostly efflux transporters that drive translocation of drugs and endogenous substances from the cell and are therefore predominantly localized on the basolateral membrane, e.g., intestinal epithelium, hepatocytes, renal epithelia, and blood-tissue barrier. On the other hand, secondary-active transporters use ion gradient across the membrane

to pair it with the transport of a substrate molecule. Together with passive transporters, secondary-active transporters form the second large, solute carrier (SLC) transporter family. SLC transporters predominantly modulate influx of nutrients and drugs into the cell. According to the current knowledge, there are more than 400 transporters in these two superfamilies, but not all of them are clinically relevant in terms of affecting drugs ADME properties. To read more about clinically relevant ABC and SLC transporters, refer to ▶ Chap. 54, "Relevance of Transporters in Clinical Studies" by Hagenbuch.

Transporters are ubiquitously distributed in the body, but their absolute abundance across organs and tissues is still puzzling. In recent years, advanced technologies such as mass spectrometry-based techniques and quantitative immunoblotting assays enabled localization and quantitative determination of major transporters. Also, with the progress in pharmacogenomics, understanding of the genes encoding membrane proteins has increased significantly.

An intrinsic property of transporters is their specificity for certain molecular entities. Some transporters have broad substrate and inhibitor specificities, while the others are more specific and interact only with a limited number of molecules with more or less similar structure. For this reason, some transporters lack specific probe substrates and inhibitors for in vitro and in vivo studies. Moreover, drug transport across biological membranes depends upon a variety of factors (e.g., it can be energy-dependent, concentration-dependent, pH-dependent, plasma membrane potential-dependent), requiring careful selection of in vitro setups for determination of transporter substrates and inhibitors.

In 2007, the International Transporter Consortium (ITC) was formed to coordinate and promote activities toward resolving the transporter-related issues in drug development. Alongside, regulatory agencies in the USA, EU, and Japan have issued guidelines on drug interaction studies, including transporter-mediated drug interactions. According to the regulations, transporter interaction studies are required for all new molecular entities (NME), but the choice of assays and transporters to be tested depend upon drug physicochemical and pharmacokinetic (PK) properties. The goals of these guidelines are to (i) define clinically relevant transporters that can impact drug efficacy and safety, (ii) establish standards for in vitro evaluation of transporter-mediated drug interactions, (iii) provide decision trees with "cutoff" values to determine whether it is necessary to conduct clinical transporter interaction studies, and (iv) establish standards for in vivo evaluation of transporter-mediated drug interactions.

From a number of transporters, only some of them need to be evaluated in the prospective in vitro transporter substrate and/or inhibition studies (Table 1). Regulatory opinions regarding major transporters are not fully harmonized, but an all-inclusive list for the European Medicines Agency (EMA), the US Food and Drug Administration (FDA), and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA) currently comprises nine transporters: MDR1 (P-gp), BCRP, OCT2, OAT1, OAT3, OATP1B1, OATP1B3, MATE1, and MATE2K, while OCT1, BSEP, and MRP2 are also considered as important. It should be noted that this is an evolving list susceptive to changes in line with the progress in transporter science. In vitro assaying of other transporters may also be necessary if there is evidence that a drug might interact with certain transporter(s). Examples of membrane transporters of emerging importance include OATP2B1, OAT2, OAT4, OCT3, and MRP4.

Results of transporter interaction studies can be useful in different phases of drug discovery. In vitro assessment of potential drug interactions with transporters is important to guide and assure safety during concomitant clinical studies. In vivo studies serve to prove the assumed interactions and provide information on drug treatment and DDIs for labeling. DDI issues are addressed in ▶ Chaps. 45, "Drug–Drug Interaction Studies" by Stopfer and ▶ 46, "In Vitro/In Vivo Correlation for Drug-Drug Interactions" by Wahlstrom and Wienkers. Also, if the new information on potential drug-transporter interaction arises during the life cycle of a product, additional transporter studies are recommended.

Table 1 Clinically relevant transporters, their localization, and examples of in vitro probe substrates and inhibitors (bold, regulatory recommended; italic, for consideration)

Transporter/ encoding gene	Transport direction	Localization	In vitro substrates	In vitro inhibitors
MDR1 (P-gp)/ ABCB1	Efflux	Enterocytes (luminal membrane) Hepatocytes (canalicular membrane) Renal proximal tubules (luminal membrane) Brain capillary endothelia (luminal membrane) Other tissues (placenta, testes, cornea)	Digoxin *N*-Methyl-quinidine Talinolol	Cyclosporin A GF120918 PSC833 Ketoconazole Verapamil
BCRP/ ABCG2	Efflux	Enterocytes (luminal membrane) Hepatocytes (canalicular membrane) Brain capillary endothelia (luminal membrane) Other tissues (testes, placenta, mammary glands)	Estrone-3-sulfate Prazosin Sulfasalazine	GF120918 Fumitremorgin C Ko134 Ko143
BSEP/ ABCB111	Efflux	Hepatocytes (canalicular membrane) Other tissues (kidney, testes, choroid plexus)	Taurocholate Glycocholate	Rifampicin Cyclosporin A
MRP2/ ABCC2	Efflux	Hepatocytes (canalicular membrane) Enterocytes (luminal membrane) Renal proximal tubules (luminal membrane) Other tissues (placenta, gallbladder, bronchi)	Estradiol-17b-glucuronide Carboxy-dichlorofluorescein Vinblastine	MK-571 PSC-833 Probenecid Cyclosporin A
MATE1/ SLC47A1	Efflux	Renal proximal and distal tubules (luminal membrane) Hepatocytes (canalicular membrane) Other tissues (skeletal muscle, adrenal gland, testes, heart)	Tetraethylammonium (TEA) 1-Methyl-4-phenylpyridinium Metformin	Cimetidine Quinidine Pyrimethamine
MATE2-K/ SLC47A2	Efflux	Renal proximal tubules (luminal membrane)	Tetraethylammonium (TEA) 1-Methyl-4-phenylpyridinium Metformin	Cimetidine Quinidine Pyrimethamine
OATP1B1/ SLCO1B1	Uptake	Hepatocytes (sinusoidal membrane)	Bromosulfophthalein Estrone-3-sulfate Estradiol-17β-glucuronide Pitavastatin Rosuvastatin	Rifampin Cyclosporine A Rifamycin SV Bromosulfophthalein Estropipate
OATP1B3/ SLCO1B3	Uptake	Hepatocytes (sinusoidal membrane) Other tissues (prostate, colon)	Bromosulfophthalein Estradiol-17β-glucuronide Cholecystokinin octapeptide (CCK-8)	Rifampin Cyclosporine A Rifamycin SV Bromosulfophthalein Ursolic acid
OAT1/ SLC22A6	Uptake	Renal proximal tubules (basolateral membrane) Other tissues (placenta)	Para-aminohippurate (PAH) Cidofovir	Probenecid
OAT3/ SLC22A8	Uptake	Renal proximal tubules (basolateral membrane)	Estrone-3-sulfate Cimetidine	Probenecid
OCT1/ SLC22A1	Uptake	Hepatocytes (sinusoidal membrane) Enterocytes (basolateral membrane) Renal proximal tubules (luminal membrane) Other tissues (neurons, heart, skeletal muscle, lung)	1-Methyl-4-phenylpyridinium Tetraethylammonium (TEA) Metformin	Quinidine Quinine Verapamil

(continued)

Table 1 (continued)

Transporter/ encoding gene	Transport direction	Localization	In vitro substrates	In vitro inhibitors
OCT2/ SLC22A2	Uptake (primary)	Renal proximal tubules (basolateral membrane) Other tissues (neurons, small intestine, trachea, bronchi, skin, placenta, brain, inner ear)	Tetraethylammonium (TEA) Metformin	Cimetidine Quinidine

This chapter is aimed to provide an overview of the best practices for the in vitro characterization of drug interactions with major clinically relevant transporters and the approaches to link in vitro findings to the expected clinical outcomes. The discussion is focused on the interaction of drugs with transporters involved in the intestinal drug absorption, as well as hepatic and renal elimination, since these transporters received notable attention in recent years. However, the importance of transporters governing drug disposition in other tissues (e.g., brain penetration, pulmonary absorption) should not be disregarded. Interactions with transporters in these tissues might have been neglected in the past, but it is expected that future research will shed more light on this matter.

Disposition-Based Classification Systems to Elucidate Drug-Transporter Interactions

The potential of a drug to interact with specific transporter, in a way that can limit drug absorption and disposition, can roughly be predicted based on the drug physicochemical properties, administration route, elimination route, and knowledge on the types and function of transporters. These basic considerations have been implemented in the concept of classification systems that address the key factors governing drug disposition in humans. Several disposition-based systems have been proposed so far (Fig. 1), based on somewhat different premises, targeting different objectives, and applicable in different phases of drug discovery and development. A major benefit of these systems is that they can facilitate the decision on whether to conduct in vitro and in vivo transporter studies and aid to identify target transporter(s) that should be tested for possible interaction with the investigational drug. It is important to note that these systems are not interchangeable; instead, they should be considered as supplementary.

The Biopharmaceutical Drug Disposition and Classification System (BDDCS) (Wu and Benet 2005) was constructed on the basis of the Biopharmaceutics Classification System (BCS), but in contrast to BCS which considers the key factors limiting oral drug absorption (solubility and permeability), BDDCS is focused on the parameters reflecting drug disposition (solubility and extent of metabolism). In terms of metabolism, drugs are classified as either extensively (>70%) or poorly (<30%) metabolized, since examples of intermediate (30–70%) drug metabolism are generally rare.

In view of this system, disposition of BDDCS class 1 drugs (highly soluble, extensively metabolized) is not affected by transporters, and there is no need to test them as substrates in vitro. However, their interaction (induction or inhibition) with transporters may be relevant in transporter-mediated DDI with victim drugs from other classes. BDDCS class 2 drugs (poorly soluble, extensively metabolized) are not affected by uptake transporters on the luminal intestinal membrane, but the effect of efflux transporters might be significant (due to limited solubility of these drugs, intracellular drug concentration might not be high enough to saturate efflux transporters). Consequently, this class is also prone to the effect of uptake and efflux transporters in the liver and brain. Finally, disposition of BDDCS classes 3 and 4 drugs (poorly metabolized) may be affected by interaction with uptake and efflux transporters

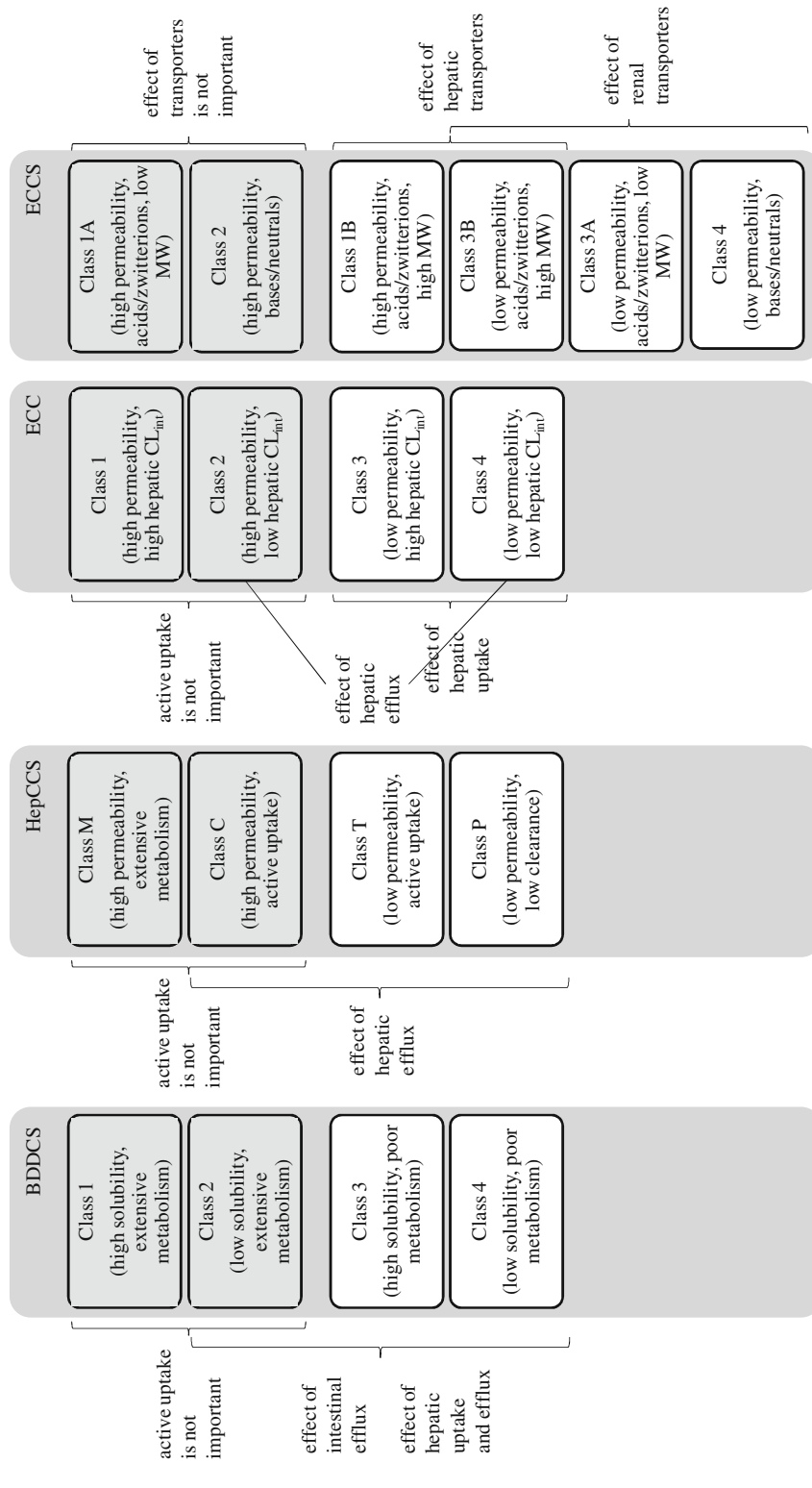


Fig. 1 The role of transporters in the proposed drug disposition and clearance classification systems

in different organs, and these drugs are candidates for transporter interaction studies.

A limitation to classify a drug within BDDCS system in the early stage of drug development is that experimental data on the extent of metabolism are not available before phase I clinical trials. An alternative approach is to use in vitro permeability data (e.g., using Caco-2 or MDCK cells) as an indicator of the extent of drug metabolism. It should also be noted that certain unexplained outliers within the proposed BDDCS drug classes have been identified, with class I statins being the most prominent ones.

The hepatic clearance classification system (HepCCS) (Fan et al. 2014) resulted from the efforts to estimate the contribution of hepatic clearance in an early drug discovery phase. The basic idea of this system is that hepatic drug clearance is influenced by a complex interplay of drug passive permeability, active uptake into the hepatocytes, and rate of metabolism and that gradual assessment of the influence of these factors might indicate the rate-determining step in hepatic drug clearance. The suggested classification criteria included passive permeability determined in MDCK cells (with a cutoff value of 5×10^{-6} cm/s) and liability to active uptake transporters (assessed using cryopreserved rat hepatocytes).

According to this system assumption, class M (highly permeable, cleared by hepatic metabolism) and class P (poorly permeable, low clearance) compounds are not prone to interfere with transporters. But, poorly permeable class T compounds interact with hepatic uptake transporters, and suitable in vitro or animal studies should be incorporated in the development pipeline of these drugs. Class T drugs can also be substrates for efflux transporters; however, efflux transport has not received much attention within HepCCS. In addition, it is postulated that, if the drug efflux is negligible, class T parent drug or metabolite might accumulate in hepatocytes, and routine measurements of drug plasma concentration might underestimate actual drug concentration relevant for the assessment of DDI. Class C is a distinct class comprising of highly permeable drugs, but substrates for uptake transporters. Hepatic clearance of class C drugs is not expected to be influenced by the activity of transporters because these drugs perfuse freely across the membranes.

HepCCS classification data can be quite useful because they imply suitability of an in vitro system to predict human in vivo clearance and potential drug-transporter interaction. But the focus of this system is on hepatic drug clearance, while other major clearance mechanisms, such as renal secretion, are not considered.

The extended clearance concept (ECC) (Camenisch et al. 2015) is similar to HepCCS in a way that it takes into account the interplay between drug permeability, (hepatic) metabolism, and biliary excretion and classifies drugs into four groups depending on the rate-determining step for hepatic elimination (passive diffusion, active uptake, or biliary efflux). However, there are differences between HepCCS and ECC, e.g., concerning classification criteria and experimental methods to determine clearance parameters. In particular, ECC classifies drugs based on in vitro determined permeability and metabolism data. In terms of permeability, a drug is classified as highly permeable when passive permeability across sinusoidal hepatocyte membrane ($PS_{inf,passive}$) is much higher than hepatic blood flow ($PS_{inf,passive} \geq 3$-fold Q_h). The extent of metabolism is expressed in terms of hepatic intrinsic clearance, which represents the sum of metabolic clearance and biliary secretion clearance. All of the key classification parameters can be determined experimentally in early phases of drug development.

Similar to HepCCS presumptions, the overall hepatic clearance of highly permeable ECC classes 1 and 2 is not considered to be affected by uptake transporters. But in contrast to HepCCS, highly permeable drugs prone to active uptake have not been classified as a separate group. ECC class 2 drugs have low hepatic clearance, and they can be secreted in bile as unchanged drugs. Therefore, the effect of efflux transporters should be considered. Low permeability ECC classes 3 and 4 are susceptive to the effect of uptake transporters, and these are generally candidate drugs for in vitro and in vivo uptake transporter interaction studies. In addition, ECC class 4 drugs are potential substrates for efflux

transporters, and these interactions should also be investigated.

An inherent drawback of ECC is that it is focused on the parameters describing hepatic drug disposition, and thus the system is informative only of the role of hepatic transporters. The contribution of other (in particular, renal) routes of elimination has also been implied, e.g., for low permeability ECC classes 3 and 4 drugs. However, the role of renal transporters has not been addressed here.

Extended clearance classification system (ECCS) is another framework aimed to identify the predominant drug clearance mechanism (El-Kattan and Varma 2018). It considers several pharmacokinetic processes: active hepatic uptake (predominantly by OATP transporters), hepatobiliary transport (mediated by MRP2, BCRP, and P-gp), hepatic drug metabolism, and renal clearance (mediated by OCT2, OAT1, OAT2, and OAT3). Hepatobiliary efflux is not considered as a rate-limiting step in drug elimination since drug exposure to efflux transporters depends upon the rate of active uptake. Comprehensive analyses of large database of drugs with diverse physicochemical and pharmacokinetic properties implied that predominant clearance pathway of a drug depends upon drug molecular weight (MW), ionization state, and passive permeability. The following cutoff values were proposed for classification: MW of 400 Da, membrane permeability of 5×10^{-6} cm/s determined using MDCK cells, and drug ionization estimated based on pKa value.

According to ECCS framework, acids and zwitterions with high MW (classes 1B and 3B), regardless of permeability, are recognized as potential substrates for hepatic uptake transporters (OATPs), and these drugs should be subjected to appropriate in vitro, and if necessary, clinical transporter interaction studies. According to the ECCS premises, substrates for hepatic uptake transporters also interact with hepatobiliary transport pathway (mediated by MRPs, BCRP, and P-gp), indicating that substrate affinity for these efflux transporters should be assessed in vitro. The major ECCS-envisaged difference between class 1B (highly permeable) and class 3B (poorly permeable) is that the first class refers to drugs that are principally eliminated as metabolites via hepatic route (due to the extensive liver metabolism), while the latter one comprises drugs eliminated unchanged via hepatic (OATP-mediated) and/or renal (OAT1, OAT2, OAT3-mediated) route. Whether a class 3B drug will be cleared predominantly through the bile or urine depends upon the substrate selectivity for OATPs vs. OATs. Therefore, class 3B drugs should also be tested for potential interactions with renal transporters. Two other ECCS classes of low permeability drugs (classes 3A and 4) are typically eliminated via the renal route, and their clearance is affected by renal transporters including uptake transporters on the basolateral membrane (OATs for class 3A and class 4 drugs and OCT2 for class 4 drugs) and efflux transporters on the apical membrane (MATE1/2K and P-gp for class 4 and potentially BCRP, MRP2, and P-gp for class 3A). Therefore, appropriate studies should be conducted to describe ECCS class 3A and 4 drug interactions with uptake and efflux renal transporters. In addition, oral pharmacokinetics of ECCS low permeable drugs (classes 3A/3B/4) is likely influenced by intestinal uptake (OATP2B1, PEPT1, MCT1, etc.) and efflux (P-gp, BCRP, MRP2) transporters, and relevance of these interactions for drug PK and possible DDIs should also be assessed. Highly permeable ECCS classes 1A and 2 are not prone to the effect of transporters, and they are mostly cleared by metabolism.

In Vitro Transporter Studies

In order to assess whether an investigational drug interacts with transporter (as substrate, inhibitor or inducer), the first step is to conduct relevant in vitro studies. Decision trees on whether and when to conduct in vitro and in vivo transporter studies are provided in the relevant guidelines issued by FDA, EMA, and PMDA (EMA 2012; PMDA 2014; FDA 2017).

Several aspects need to be considered when designing in vitro drug-transporter interaction studies. Firstly, the design of the study depends

on the investigational objective, e.g., whether the objective is to identify the transporter(s) interacting with the drug, to estimate kinetic parameters that describe drug-transporter interaction, or to provide information to guide concomitant in vivo interaction studies. Also, study design depends upon the type of interaction to be investigated (e.g., transporter-mediated drug uptake or efflux, drug-mediated inhibition or induction of transporter, DDI study, complex transporter-enzyme interaction studies). For each type of study, appropriate in vitro system and experimental procedure need to be defined.

This section provides an overview of the standard tests, experimental conditions, protocols, and probe substrates/inhibitors used for the in vitro assessment of transporter-mediated drug interactions, along with the basic descriptions regarding data analysis and interpretation.

In Vitro Systems

A drug-transporter interaction can be assessed in vitro using various systems, e.g., different cell-based systems, membrane vesicles or artificial systems, depending on the type and objective of the study, and characteristics of the investigational drug (e.g., high or low permeability). When possible, the expression and activity of the transporter in the chosen system should closely resemble the conditions in vivo. Expression of transporters can be confirmed by quantitative polymerase chain reaction (qPCR), immunochemistry, or functional assay. The selected test system and experimental procedure should be validated and well characterized (e.g., source of the cell lines, cell culture conditions, incubation time, buffer pH, concentration of probe substrate/inhibitor, sampling intervals, method for data analysis, cutoff values). The activity of transporters needs to be verified using positive and negative controls with known probe substrates, and the obtained results should comply with the values from literature.

Commonly used cell-based systems include Caco-2 cells and transfected cell lines such as MDCK, LLC-PK1, CHO, and HEK293. One of the comprehensive reviews on the cell-based in vitro models to predict drug permeability, including transporter-mediated translocation, is provided by Sarmento et al. (2012). Also, some of the most prominent in vitro test systems are addressed in the review of Brouwer et al. (2013).

Caco-2 cells, derived from human colon adenocarcinoma cells, morphologically and functionally resemble intestinal epithelial cells. They express most of the key human intestinal transporters and are generally used to assess overall (passive and active) intestinal permeability of a drug. However, the expression of transporters highly depends upon the origin of cells, number of passage, culture conditions, and seeding density. For transporter studies, the cells are seeded on a semi-porous filter as a monolayer with a density of 2.5×10^5 cells/cm^2 and cultivated in normal serum-containing cell culture medium, e.g., Dulbecco's modified eagle's medium. The cultivation is relatively long (usually 20–30 days), which is one of the limitations of this in vitro system. Viability of the cells needs to be confirmed using, e.g., Trypan blue. Also, integrity of the monolayer needs to be verified by measuring transepithelial electrical resistance (TEER) or paracellular flux of a low permeability compound (e.g., phenol red, ^{14}C-mannitol). Prior to the transport experiments, the cells are washed with transport medium and incubated at 37 °C for 30 min. Afterward, a drug solution is added to one side of the cells – basolateral (B) or apical (A), and the samples are taken from the other side in predetermined time intervals. The amount of drug can be quantified using various methods such as high-performance liquid chromatography (HPLC) or liquid scintillation spectrometry. A major disadvantage of using Caco-2 cells for transporter experiments is high interlaboratory variability in the obtained results. This is mainly caused by variations in the permeability and expression of transporters. Therefore, each laboratory is responsible to check permeability of the cells using proper controls (e.g., propranolol, inulin) and determine the expression of transporters.

MDCK cells are often used to study drug permeability, but the activity of membrane transporters in parental cells is relatively low. The cells are usually seeded in well plates at the concentration of 2.5×10^5 cells/ml, in culture with a suitable

medium, e.g., Earle's balanced salt solution supplemented with 10% fetal bovine serum, at 37 °C, in the atmosphere with 95% humidity and 5% CO_2. In general, testing procedure is similar to the one applied for Caco-2 cells, except that MDCK cells require only 3 days of growing before they are ready for permeability experiments.

Transfected (recombinant, overexpressing) cells are suitable to test interaction with a single transporter, although a suitable cell line can also be transfected to host more than one transporter. The choice of the cell line depends upon the type of the assay, e.g., HEK293 and CHO for the expression of uptake transporters or MDCK and LLC-PK1 for efflux transporters. These cells usually require shorter cultivation time (e.g., 3 days for MDCK transfectant cells), while the experimental procedure depends upon the type of study (substrate or inhibition study). If the compound of interest is a potential substrate, incubation usually lasts less than 10 min, and the experiments need to be conducted in vector-free (parental, wild-type) cells as control. The difference in the efflux ratio between the transfected and parental cells indicates transporter-mediated uptake or efflux process. In the inhibition study (when the accumulation of the probe substrate is measured), incubation time is adjusted depending on the compound properties (i.e., permeability) and mechanism of inhibition. In certain cases, time-dependent inhibition may occur, and therefore it is important to use initial uptake rates of the probe substrate when calculating relevant transporter kinetic parameters.

Other cell-based systems for transporter studies include cells used in fluorescent dye-based assays, e.g., calcein assay for P-gp and MRPs transporters or Hoechst 33342 assay for BCRP transporter (Bircsak et al. 2013).

Hepatocyte cultures have been historically used to determine hepatic drug clearance (mostly CYP-mediated), but nowadays they are also used to in vitro assess drug interaction with hepatic transporters. Different hepatocyte cultures of human or animal origin can be used for these purposes, e.g., suspension, plate, and sandwich-cultured hepatocytes (SCH), depending on the type of the study.

Hepatocytes in suspension contain freshly isolated or cryopreserved hepatocytes in the appropriate medium, e.g., William's E medium or Hanks' balanced salt solution. Prior to the experiment, the suspension is preincubated at 37 °C for about 10 min, followed by incubation during predetermined time intervals with different concentrations of the test drug, both at 37 °C and on ice (or at 4 °C). In the subsequent step, hepatocytes are separated from excess substrate using either oil spin or filterplate method. Quantification of the drug in hepatocytes is performed using scintillation counting (for radioactively labeled drug) or liquid chromatography-mass spectrometry (LC-MS). Determination of the uptake transport rate is determined by subtracting the value obtained at 4 °C (corresponds to passive diffusion) from the value obtained at 37 °C (overall drug uptake including active and passive transport). Although widely used, this approach to determine active uptake rate has been criticized due to the fact that membrane fluidity, and consequently, drug passive permeability also changes with temperature (the membrane becomes more rigid), and this may lead to erroneous results. An alternative approach to determine passive transport, in order to subtract it from the total cellular uptake, is incubation with specific transporter inhibitors. But the problem with this approach is the lack of specific inhibitors (Zamek-Gliszczynski et al. 2013).

SCH can be obtained from humans and animals, including genetically modified animals with knockdown of a specific transporter. Hepatocytes are cultured between the two extracellular matrices, and as opposed to hepatocytes in suspension, functionality and polarity of the cells is preserved in this system. From all hepatocyte cultures, only SCH express fully functional canalicular efflux transporters, and therefore they are able to mimic biliary excretory function. Regarding that they also maintain enzyme activity, SCH can be used to study enzyme-transporter interplay and DDI. Experimental protocol with SCH system depends upon the goal of the study, e.g., different protocols are adopted for studying interaction with uptake transporters or for determination of biliary clearance. Example protocols are described in literature (De Bruyn et al. 2013; Pfeifer et al. 2013;

Cantrill and Houston 2017). In general, the isolated hepatocytes are seeded at a density of 1.5×10^5 (rat SCH) or 1.8×10^5 cells/cm^2 (human SCH) on multi-well plates. Following the initial culturing period when the cells attach to the plate, the unattached excess is removed, and the culture is covered with another plate. Depending on the culturing time and conditions, expression and functionality of transporters (and metabolizing enzymes) may vary significantly (De Bruyn et al. 2013).

Membrane vesicles came to use more recently, but they are now available from different sources, and they are mainly used to investigate drug interaction with efflux transporters from the ABC family. Depending on the source, membrane can express various transporters. Inverted (inside-out) membrane vesicles can be derived from insect cells (e.g., baculovirus-infected insect cell lines Sf9 or Sf21, High Five, and Sf+ insect cells), from cDNA-transfected mammalian cell lines (e.g., CHO, HeLa, V79 hamster, HEK293), from tissues (e.g., kidney or liver), or from artificial membranes that express the transporter of interest. Testing with vesicles from transfected cell lines requires the use of vesicles from non-transfected (wild-type) cells because these cells may also possess some background transporter activity. Vesicles derived from tissues are useful for simultaneous investigation of multiple transporters on apical or basolateral membrane, but the procedure requires special care because the preparation may contain vesicles from the opposite membrane expressing transporters that share the investigational drug as a substrate.

Prepared membrane vesicles can be resuspended in hypotonic or isotonic buffer, or freezed, and stored at −80 °C or in the liquid nitrogen. Commercial preparations are also available. A raw vesicle mixture is composed of inside-out vesicles, right-side-out vesicles, and open lamellar membrane fragments. In efflux assays only inside-out vesicles can interact with ATP and transport a substrate into the vesicle, so the raw mixture can be purified to increase the concentration of these vesicles. The experiments are performed by incubating vesicles in a buffer containing different concentrations of the tested substrate (radiolabeled or fluorescence labeled) at 37 °C in the presence of ATP. Parallel testing in the presence of AMP serves as a negative control for passive permeability, while probes with the reference standard serve as a positive control. After predetermined incubation time, active uptake process is terminated by adding the ice-cold solution, and the vesicles are separated by filtration or centrifugation (centrifugation is preferable for highly hydrophobic substances that tend to bind to the filter). Separation of the vesicles should be performed quickly, because the stop solution does not withhold efflux of substrate out of the vesicle. The collected vesicles are then lysed, and the accumulated drug is quantified based on radioactivity, fluorescence, or using LC-MS. One of the benefits of the membrane vesicle systems is that the measured concentration of a substrate or inhibitor represents the unbound concentration which is used to estimate kinetic parameters of drug transport.

Transport assays in membrane vesicles are usually performed in indirect manner (as an inhibition test), whereas the amount of accumulated probe substrate inside the vesicles is measured after incubation with the investigational drug in the presence of ATP. Indirect assay can only show whether the tested drug interacts with transporter (manifested as inhibition of the substrate transport), but it does not provide information on the nature of interaction (whether the drug is inhibitor or competitive substrate). Membrane vesicles can also be used in the direct assay, if the tested drug is available in, e.g., radiolabeled form. In this case, the outcome indicates the nature of the interaction with transporter.

Advances in microfluidics and lab-on-a-chip technologies provided new potential means to study drug disposition, e.g., 3D cell culture platforms that mimic the conditions of renal proximal tubule epithelial cells (Gozalpour and Fenner 2018). These systems are gaining more attention, and they may be more routinely used to study transporter-mediated drug interactions in the future.

Probe Substances

A number of probe substrates and inhibitors can be used in transporter interaction studies, and some of

them are listed in Table 1. Examples of well-known probe substrates, inhibitors, and inducers for the key clinically relevant transporters are provided in the FDA online database (https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm093664.htm), together with the relevant notes concerning their selectivity. Namely, some of these probe substances are selective substrates or inhibitors of a certain transporter, but some interact with more than one transporter. It should be noted that the FDA list is not a definitive one (nor all-inclusive), and it is occasionally updated. On the other hand, EMA guideline (EMA 2012) does not provide a comprehensive list of probe substances, and instead the manufacturers are instructed to refer to the relevant literature. An example of a good review on probe substrates and inhibitors for major transporters (except for P-gp) has been provided by Momper et al. (2016).

Approaches to Estimate Transporter Kinetic Parameters

In order to understand and quantitatively assess the role of transporters in drug PK and DDIs, drug-transporter interactions need to be characterized in terms of relevant kinetic parameters. Depending on the type of interaction (uptake, efflux, inhibition, induction), different parameters are used to describe transport kinetics, e.g., net transport rate for P-gp and BCRP efflux substrate, inhibition constant (K_i) and/or half-maximal inhibitory concentration (IC_{50}) for transporter inhibitor, and intrinsic clearance for uptake transporter substrate. These parameters can be obtained by different methods, including in vitro assays, preclinical testing in animals (transporter gene-knockout models), and human in vivo studies.

One of the commonly used approaches to analyze kinetics of drug-transporter interactions refers to the employment of conventional static methods. These methods are based on relatively simple relationships and equations, but they imply certain approximations and simplifications of the in vivo conditions, e.g., that only one process is governing drug disposition (active uptake or efflux) while other concomitant processes (passive diffusion, metabolism, intracellular and extracellular binding) are neglected. Such methods can be useful for collecting prior information to guide future interaction studies, but they are not applicable when complex interactions are involved in drug disposition.

On the other hand, transporter kinetic parameters can be assessed using more complex mechanistic compartmental models. These models account for cellular and media compartments, and they take into account parallel processes (active transport, passive diffusion, drug binding, and metabolism) that influence drug concentration in the predefined compartments. Additional advantage of mechanistic models is that transporter kinetic parameters are estimated based on multiple drug concentration data at different time points to account for dynamic changes during, e.g., drug uptake process. Some of the mechanistic compartmental models have been detailed in the review published by ITC group (Zamek-Gliszczynski et al. 2013).

Substrate Studies

According to the regulatory guidelines (EMA 2012; PMDA 2014; FDA 2017), in vitro assessment of a drug potential to act as a substrate for transporter is recommended for several transporters (efflux or influx), localized mostly in the major clearance organs (the gut, liver, and kidneys). Regarding the gut efflux transporters, substrate potential should be investigated for P-gp and BCRP. Substratability studies for hepatic uptake transporters are suggested for OATP1B1 and OATP1B3. In case of renal transporters, substratability for uptake transporters OCT2, OAT1, and OAT3 and efflux transporters MATE1 and MATE2-K should be evaluated. The decision tree to evaluate whether a drug is a substrate of the major transporters is shown in Fig. 2.

In vitro substrate studies are performed in the presence/absence of the known inhibitors. Inhibitors should be selected based on likelihood of co-administration with the tested drug, and a general suggestion is to start with selective inhibitors (if available).

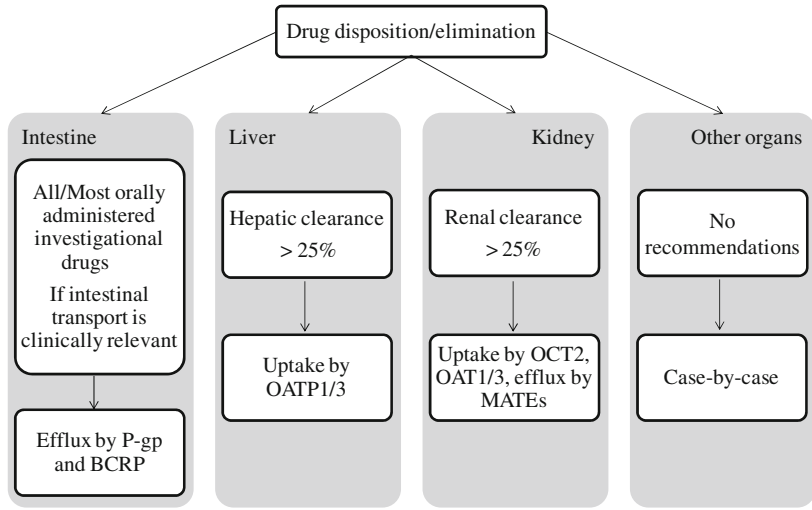


Fig. 2 Recommendations on the investigation of a drug substrate potential for major transporters

Substrates for P-gp and BCRP Efflux Transporters

In cases when intestinal wall secretion exceeds 25% of the total clearance for orally administered drug, in vitro testing should be performed to identify the transporter involved and to qualitatively or semiquantitatively describe the interaction. The involvement of gut secretion is assessed when intestinal drug transport is clinically relevant (e.g., for oral, nasal, or inhalation delivery route), and it can be estimated based on drug PK and mass balance studies. In practice, substrate potential for P-gp and BCRP efflux transporters is assessed for all (or most) orally administered investigational drugs (Fig. 2).

Standard systems for determination of whether a drug is a substrate for P-gp or BCRP efflux transporters include Caco-2 cells, but other cell cultures (e.g., MDCK, LLC-PK1) or membrane vesicles overexpressing the transporter of interest can also be used. To confirm adequate expression of transporters, experiments with known P-gp or BCRP probe substrates are performed as positive control.

Routinely used cell-based test to identify substrates for efflux transporters is bidirectional transport assay. The test system comprises two compartments (donor and acceptor) divided by a cell monolayer. The conditions are usually kept the same in both compartments (pH 7.4). Transport of a drug (expressed as apparent permeability coefficient, P_{app}) is determined in both directions (B–A and A–B) using the following equation:

$$P_{app} = \frac{(dQ/dt) \times V}{A \times C_0} \quad (1)$$

Here, dQ/dt is the permeability rate, C_0 the initial concentration in donor compartment, V the volume of receiver compartment, and A surface area of the monolayer.

Based on these values, the efflux ratio (R_E) is calculated as:

$$R_E = \frac{P_{app}\ (B - A)}{P_{app}(A - B)} \quad (2)$$

In cases when transfected cell lines are used, efflux ratio is calculated for both transfected (R_T) and wild-type cells (R_W), and the resulting ratio is expressed as:

$$R_E = \frac{R_T}{R_W} \quad (3)$$

R_E value higher than 2 indicates that a drug is a substrate for the tested efflux transporter. In contrast, R_E around 1 implies that there is minimal or no drug efflux. Low R_E values, i.e., $R_E < 0.5$, suggest the involvement of uptake transporters.

Bidirectional transport assay should be performed with various concentrations of the test drug (at least four different concentrations) covering the range of clinically relevant concentrations (concentration at the luminal intestinal membrane), except when tested concentration of a drug is limited by its solubility or cytotoxicity. Also, transport rate should be linear over the employed concentration range and under the selected experimental conditions.

To confirm that investigational drug is a substrate for the efflux transporter of interest, additional study with P-gp or BCRP inhibitor should be performed. If possible, a selective inhibitor should be used, at concentration of at least ten times its K_i value. At least 50% reduction in R_E value in the presence of the inhibitor in comparison with R_E in the absence of inhibitor or R_E equal to unity in the presence of inhibitor confirms that a drug is a substrate for the tested transporter. In such a case, additional in vivo interaction studies should be considered. Other cutoff values may also be used if they are adequately justified and validated with known drug substrates. The problem with the identification of P-gp substrates with the addition of inhibitor lies with the fact that no specific P-gp inhibitor have been identified so far. Therefore, when using Caco-2 cells, which express various transporters, the study should be performed with two or more P-gp inhibitors to insulate specific contribution of this transporter.

Inverted membrane vesicles can also be used to identify drug substrates for efflux transporters. Here the drug-binding site of the transporter is located at the outer membrane surface, so the transported substrate is accumulated within the vesicles. Vesicles are subsequently filtered, and the amount of drug is quantified using a suitable method (as noted before). Transport activity is determined by subtracting the amount of drug accumulated in control vesicles incubated with AMP from the transporter-overexpressing vesicles incubated with ATP.

Substrate uptake rates determined in membrane vesicles are calculated based on mg protein, so they represent relative values in relation to the specific vesicle system. In order to calculate absolute values, i.e., Michaelis-Menten constant (K_m), a range of substrate concentrations should be tested. The ITC group (Brouwer et al. 2013) suggests to test at least seven substrate concentrations, whereas the highest one should correspond to at least 90% of the maximum transport velocity, along with a control (zero concentration) probe. Subsequently, kinetic parameters can be obtained from the Michaelis-Menten equation (Eq. 4) by applying regression analysis:

$$V = \frac{V_{max} \times [S]}{K_m + [S]} + CL_{passive} \times [S] \qquad (4)$$

In Eq. 4, V is the uptake rate, V_{max} the maximum uptake rate, $CL_{passive}$ passive diffusion clearance, and [S] substrate concentration.

Membrane vesicle-based test may not be appropriate for highly permeable lipophilic compounds because extensive passive diffusion in both directions may overrun the efflux process (i.e., highly permeable drugs may escape the vesicle via passive diffusion through the lipid bilayer). An alternative way to evaluate whether a drug is a substrate for efflux transporters is ATPase assay, which is especially useful for highly permeable drugs. The test is performed on transporter-rich membranes obtained from mammalian cells or baculovirus-insect cells. The activity of transporter is reflected in changes in ATPase activity, which can be estimated by quantifying the amount of the generated inorganic phosphate. This test is useful in identifying highly permeable drug substrates for ABC transporters, but since this is an indirect assay, it is recommended to confirm the results by an additional assay.

Substrates for SLC Transporters

In vitro uptake studies are routinely used to investigate whether a drug is a substrate for the key transporters from the SLC family, namely, OATP transporters in hepatocytes and OCT, OAT, and MATE transporters in renal tubules. According to the regulatory recommendations, OATP-mediated

hepatic uptake studies should be performed when hepatic transformation is major or one of the major routes of drug transformation, i.e., when hepatic metabolism or biliary secretion contributes to more than 25% of the total drug clearance (Fig. 2) or when uptake of a drug into the liver is clinically important (e.g., to exhibit pharmacological effect). In a similar manner, uptake by renal transporters (OCT, OAT, and MATE) should be assessed when renal secretion exceeds 25% of the total drug clearance and when in vivo renal clearance (in preclinical species) is higher than glomerular filtration rate (GFR). Assuming there is no reabsorption of a drug, active renal secretion can be estimated based on renal clearance (CLr), GFR, and unbound fraction of drug in plasma ($f_{u,p}$) using Eq. 5. In addition, renal clearance mechanisms should also be elucidated for drugs that exhibit pharmacological or toxic effect in kidneys.

$$\text{Active renal secretion} = \text{CLr} - \left(f_{u,p} \times \text{GFR}\right) \tag{5}$$

Commonly used in vitro systems for the assessment of OATP-mediated transport include hepatocyte cultures and transfected cell lines (CHO, HEK293, MDCK). Substrates for renal transporters (OCT, OAT, and MATE) are usually assayed on transfected cell lines (CHO, HEK293, MDCK), although MATE-mediated uptake can also be investigated in membrane vesicles. A note herein is that, when estimating MATE substrates, pH in the test system should be adjusted because activity of this transporter is affected by proton gradient.

When conducting an uptake study, a substrate can rapidly accumulate inside the cell or vesicle forcing the uptake process to deviate from linearity. For this reason, it is important to select early time points (in the linear phase) to estimate the uptake rate and calculate relevant kinetic parameters. The initial uptake rate can be estimated using linear or dynamic regression analysis. Active uptake clearance (CL_{uptake}) is then determined using Eq. 6, based on the data from the initial linear uptake phase:

$$CL_{uptake} = \frac{\text{Amount}_{t2} - \text{Amount}_{t1}}{(t2 - t1) \times C_{medium}} \tag{6}$$

In Eq. 6, Amount_{t1} and Amount_{t2} are cumulative amounts of drug over time t1 and t2, respectively, and C_{medium} concentration of drug in medium.

In addition, percentage of the active uptake (in relation to passive diffusion) can be estimated from the slope of the initial uptake phase determined in the presence and absence of a known inhibitor:

$$\%\text{Active uptake} = \left(1 - \frac{\text{Slope with inhibitor}}{\text{Slope without inhibitor}}\right) \times 100 \tag{7}$$

When a drug is identified as a substrate for the uptake transporter, kinetic parameters (K_m, V_{max}, CL_{int}) are calculated to characterize the transport process. Uptake kinetic parameters can be estimated using Eq. 4 and regression analysis of the experimental data. Then active uptake clearance (or unbound active uptake clearance, $CL_{active,u}$) can be expressed as:

$$CL_{active,u} = \frac{V_{max}}{K_m} \tag{8}$$

whereas the total unbound uptake clearance of a drug ($CL_{uptake,u}$) is the sum of passive ($CL_{passive,u}$) and active clearances ($CL_{active,u}$):

$$CL_{uptake,u} = CL_{passive,u} + CL_{active,u} \tag{9}$$

A drug is considered as OATB substrate if uptake ratio, in terms of drug uptake in transfected cells in comparison to drug uptake in empty vector-transfectant cells, is equal to or higher than 2. In addition, the test should be performed with a selective inhibitor of the transporter of interest (e.g., rifampin for OATP1B1/OATP1B3), at concentration of at least ten times its K_i or IC_{50} value. This test confirms that a drug is a substrate for the tested transporter if drug uptake in the presence of selective inhibitor is reduced for at least 50% in comparison to the uptake in the absence of

inhibitor. Other cutoff values may also be used if properly justified and validated with known probe substrates.

Hepatobiliary Transport

Drug transport from the blood, through the hepatocytes, into the bile requires the activity of both uptake transporters located at the basolateral (sinusoidal) membrane of hepatocytes and efflux transporters on the canalicular hepatocyte membrane. Therefore, this kind of transport cannot be assessed in cell systems expressing only uptake or efflux transporters. Up to now, the most suitable in vitro assay to assess the interplay between uptake and efflux hepatocyte transporters and estimate biliary clearance of a drug has been SCH assay (Nakakariya et al. 2012; De Bruyn et al. 2013; Yang et al. 2016). Based on SCH-generated data, it is possible to calculate biliary efflux clearance and biliary excretion index (BEI) of a drug using the following equations (Liu et al. 1999a):

$$CL_{bile,app} = \frac{Amount_{cell+bile} - Amount_{cell}}{t \times C_{medium}} \quad (10)$$

$$CL_{bile,int} = \frac{Amount_{cell+bile} - Amount_{cell}}{t \times C_{cell}} \quad (11)$$

$$BEI = \frac{Amount_{cell+bile} - Amount_{cell}}{Amount_{cell+bile}} \times 100\% \quad (12)$$

$$C_{cell} = \frac{Amount_{cell}}{V_{intracell}} \quad (13)$$

Depending on whether the drug concentration in medium (C_{medium}) or drug intracellular concentration (C_{cell}) is used, in vitro apparent biliary clearance from medium to bile ($CL_{bile,app}$) and in vitro intrinsic biliary clearance from hepatocytes to bile ($CL_{bile,int}$) can be estimated, respectively. Here, C_{medium} is the initial drug concentration in the incubation medium, C_{cell} is concentration of a drug in hepatocytes which can be calculated using Eq. 13, and t is the incubation time. $V_{intracell}$ in Eq. 13 is the volume of intracellular space per mg protein. $Amount_{cell}$ is the amount of drug accumulated in the cells, and $Amount_{cell+bile}$ is the sum of drug amount in the cells and bile canaliculi. These values can be estimated experimentally by modulating opening of tight junctions to bile canaliculi using incubation media with and without Ca^{2+}/Mg^{2+}. The principle is described within the patented B-Clear® technology (Liu et al. 1999b). BEI is a measure of drug accumulation in the bile and serves as a qualitative indicator of biliary excretion of drug.

Another method to calculate biliary clearance of a drug (CL_{bile}) was proposed by Cantrill and Houston (2017) (Eq. 14), whereas hepatocyte-to-media unbound concentration ratio (Kp_u) is used to reflect drug partition between hepatocytes and external medium and hence the contribution of active uptake process (Eq. 15):

$$CL_{bile} = \frac{Amount_{cell+bile} - Amount_{cell}}{t \times (Kp_u \times C_{medium})} \quad (14)$$

$$Kp_u = \frac{CL_{uptake}}{CL_{passive}} \quad (15)$$

In vitro obtained hepatic drug clearance values are usually expressed in μl/min/mg protein, and for further scaling to in vivo estimates, these values should be converted to ml/h/kg using the available data on liver weight and protein content in liver tissue of the species of interest. The converted values can also be used as inputs in physiologically based pharmacokinetic (PBPK) models to predict PK of a drug in humans or in preclinical species.

Inhibition Studies

Inhibitors affect transporter activity by noncompetitive binding to the receptor site, thus hindering other drugs to interact with the transporter, or they can obstruct the process that generates the required energy for active transport (e.g., P-gp inhibitors that block ATP hydrolysis).

According to the regulatory guidelines (EMA 2012; PMDA 2014; FDA 2017), the potential of a drug to inhibit a transporter should be accessed for

transporters that are known to be involved in clinically relevant drug interactions. There are certain differences between the regulatory documents, but the inclusive list of recommended target transporters comprises P-gp, BCRP, OATP1B1, OATP1B3, OCT2, OAT1, OAT3, MATEs, OCT1, and BSEP.

An inhibition study can be performed as unidirectional or bidirectional transport assay with an annotation that monolayer cell cultures are not suitable for efflux inhibition tests because there is no accepted method to calculate K_i value. Preferable method to estimate drug inhibition potential for an efflux transporter includes testing with membrane vesicles. Also, due to the pronounced variability regarding P-gp inhibition parameters between laboratories, EMA (2012) suggests to use at least two systems to test a drug inhibition potential for this transporter. If justified, this approach can be used for other transporters as well. An alternative to detect a drug inhibition effect on P-gp and MRP1 transporters is to use whole cell-type calcein assay. This is an indirect type of test that measures fluorescence of free calcein trapped inside the cell due to the inhibition of P-gp and MRP1 activity. The test is based on the fact that hydrophobic calcein ester (calcein AM), added to the cell culture, can be actively transported out of the cell. But if efflux transporters are blocked, hydrolyzed calcein derivative (fluorescent) will accumulate within the cell. So, if fluorescence in the presence of different concentrations of the investigational drug increases, this indicates that a drug is an inhibitor of the tested transporter (Glavinas et al. 2011).

The inhibitory potential of a drug is assessed in the presence of a known probe substrate for the transporter of interest. Since inhibitory potential of a drug can be substrate-specific, the best option for the in vitro study is to use the same probe substrate as in the prospective clinical study. If this is not an option, a preferred in vitro probe substrate is the one that generates lower IC_{50} value for the known inhibitors (to minimize the chance for false-negative results). Also, it should be noted that investigation of a drug inhibition potential for OATP1B1/OATP1B3 transporters requires an additional preincubation step for at least 30 min to detect possible time-dependent inhibition which could reduce a drug IC_{50} value, e.g., as observed for cyclosporine (Gertz et al. 2013).

The test should be performed with a range of substrate concentrations (at or preferably below its K_m, taking care that the transport is linear over the employed concentration range) and the range of concentration for the investigational drug, starting from the highest concentration (higher than clinically expected at the site of interaction) to potentiate inhibitory effect, but not to exceed drug solubility and cytotoxic concentrations.

Clinically relevant concentration of the tested inhibitor (I) depends upon the localization of the interacting transporter:

I. Intestinal luminal concentration (I_{gut}) for luminal intestinal transporters (P-gp, BCRP):

$$I_{gut} = \frac{\text{Dose}}{250\ \text{ml}} \tag{16}$$

II. Maximal unbound hepatic inlet concentration ($I_{u,in,max}$) for hepatic uptake transporters (OATP1B):

$$I_{u,in,max} = f_{u,p} \times \left(\frac{C_{max} + F_a \times F_g \times k_a \times \text{Dose}}{Q_h \times R_{bp}} \right) \tag{17}$$

In Eq. 17, F_a is fraction of drug absorbed, F_g fraction of drug that escapes the intestine unchanged, k_a absorption rate constant, Q_h hepatic blood flow, and R_{bp} blood-to-plasma concentration ratio. Certain approximations can be made when F_a, F_g, and k_a are unknown, and then a worst-case scenario is assumed with $F_a \times F_g = 1$ and $k_a = 0.1$. Also, $f_{u,p} = 1\%$ can be used when experimentally determined value is lower than 1% (due to uncertainties in the measurements).

III. Maximal unbound plasma concentration ($I_{max,u}$) for renal transporters (OAT, OCT, MATE):

$$I_u = C_{max,u} \quad (18)$$

If a drug shows inhibition effect, the test is repeated with additional drug concentrations to calculate IC_{50} and K_i values. According to EMA (2012), it is preferable to use K_i to assess a drug potential to inhibit a transporter. IC_{50} is suggested only as an alternative when K_i cannot be determined, if linear conditions are maintained, and there is no time-dependent inhibition.

IC_{50} can be estimated using several methods (Balimane et al. 2008; Volpe et al. 2014). In general, the calculated P_{app} values for a probe substrate in the presence of different inhibitor concentrations are plotted against inhibitor concentrations, and the obtained curve is fitted by regression analysis to determine the inhibitor concentration that corresponds to 50% reduction in the probe substrate P_{app}. The final value is usually expressed as a mean for at least three separate measurements. There are also different methods to calculate K_i. One of the approaches is to convert IC_{50} to K_i using the following equation:

$$K_i = \frac{IC_{50}}{1 + \frac{S}{K_m}} \quad (S - \text{substrate concentration}) \quad (19)$$

Several aspects need to be considered when determining IC_{50} and K_i values, as highlighted by the ITC group (Brouwer et al. 2013). One thing is that the obtained values may vary depending on the method employed, and this may complicate comparison of the results obtained in different laboratories. Moreover, IC_{50} does not provide information on the type of inhibition (competitive, noncompetitive, uncompetitive). Also, in case of competitive inhibition, IC_{50} depends upon the substrate concentration (in oppose to K_i). This annotation should be taken into account if IC_{50} is used for in vitro-in vivo extrapolation (IVIVE). IC_{50} is generally determined with a single probe substrate concentration, and when the substrate concentration is far below K_m, IC_{50} will be equivalent to K_i value. On the other hand, substrate concentration above its K_m (>50% K_m) leads to erroneous K_i estimates based on Eq. 19.

Another approach to analyze inhibition data refers to the use of Dixon plots (distinguishes between competitive and noncompetitive or uncompetitive inhibition) or, e.g., "quotient velocity plot" (distinguishes between all types of inhibition, including competitive, noncompetitive, and mixed-type inhibition) (Yoshino and Murakami 2009). These graphical methods enable determination of the inhibition type and K_i which can be considered as a more robust parameter than IC_{50}.

In vitro determined IC_{50} and K_i values are subsequently used to estimate a drug potential to in vivo inhibit transporters of interest, which will be discussed in section "In Vitro/In Vivo Extrapolation."

Induction Studies

Although it is known that certain drugs may induce membrane transporters, up to date there are no official recommendations on how to conduct in vitro transporter induction studies. The only suggestion given so far concerns investigation of potential P-gp inducers in cases when an investigational drug is identified as CYP inducer (due to similar mechanism of inducing CYP enzymes and P-gp transporter by interacting with PXR and CAR nuclear receptors). In case a drug is a CYP inducer, an induction study should be conducted using the same assumptions and test conditions as suggested for the investigation of CYP induction potential.

Interaction with Metabolites

Metabolites are hydrophilic in nature and usually not able to passively cross lipid membrane barrier, so they require active transport. Inhibition of these processes may lead to accumulation of potentially toxic or active metabolites in tissues, which can lead to potentially dangerous clinical consequences. For this reason, investigation of metabolite-transporter interaction should be included in

drug development pipeline. Regulatory recommendation is that metabolite should be tested as a substrate for transporter(s) when a metabolite contributes to ≥50% of the total drug activity (EMA 2012; FDA 2017). Inhibition potential of metabolites is usually assessed in vivo, along with the parent drug. But in the cases when parent drug is not identified as inhibitor, in vitro inhibition studies with metabolites should still be performed when a metabolite is less polar than the parent drug and exposure to metabolite exceeds 25% of parent drug exposure or when a metabolite is more polar than the parent drug and exposure to metabolite exceeds 100% of parent drug exposure (FDA 2017).

Still, these studies can only be planned in later stages of drug development, after the metabolites have been identified in the relevant in vivo studies. Additionally, in silico quantitative structure-activity relationship (QSAR) may contribute to early investigation of the presumed metabolites of NDE and their elimination routes.

Critical Assessment of In Vitro Methods

In vitro studies are, with no doubt, important part of the overall assessment of transporters' effect on drug absorption, disposition, and DDIs. However, almost all of these studies have certain limitations, depending on the in vitro system and type of the study.

Most of the suggested probe substrates and inhibitors are interacting with multiple transporters, meaning that they usually lack sensitivity (e.g., rosuvastatin is suggested as OATP1B1 substrate, but it is also a substrate for, e.g., BCRP, MRP2, P-gp, OAT3; metformin is substrate for OCT2, but it also interacts with MATE1/MATE2 and OCT1). The criteria for the selection of appropriate probe substrates and inhibitors are also lacking, except for P-gp, and need to be established in the near future.

Additional matter to consider in the in vitro studies is whether qualitative information on the involvement of a single transporter in drug absorption and disposition is enough to decide upon the need to conduct clinical interaction studies. Namely, neglecting the role of multiple transporters on drug interactions may lead to false-positive (implying unnecessary clinical studies) or false-negative results (when necessary clinical testing is left out). Also, for drugs whose PK is influenced by the combined effect of transporter(s) and metabolic enzymes, a standard transporter in vitro assay might not be able to indicate clinical relevance of the tested transporter for a particular drug. The use of multiple tests with different systems might be the most accurate option to predict the transporter-mediated drug interactions.

As noted above, for certain types of transporter studies, various in vitro test systems and methods can be used, so different laboratories can obtain different results. And comparison of such results can be challenging. Even with a single method, and the same type of material, the results may vary (Cantrill and Houston 2017).

In case of cell-based studies, certain drugs cannot be tested in clinically relevant concentration ranges because of the limited solubility or high cytotoxicity. Organic solvents might be added to increase drug solubility, but in limited concentration (less than 1% vol/vol) due to potential effect on cell integrity and transporter activity. As for cytotoxic drugs, an alternative option is testing in membrane vesicle-based systems.

Nonspecific binding to the system components (cells, apparatus) is also a potential problem because it may reduce effective drug concentration. Therefore, the concomitant mass balance tests (percent recovery) are routinely done to evaluate nonspecific binding. Percent recovery is expressed as the sum amount of substrate remaining in donor and acceptor compartments at the end of test in relation to the initial substrate amount in the donor compartment. The test is considered valid if percent recovery is more than 70%. A drug can also bind to membrane vesicles, and therefore its binding affinity should be checked prior to experiments. Moreover, a drug can bind to the filter used to separate membrane, and this phenomenon may be obviated by using nonreactive filter materials or by preincubation of filter with excess unlabeled substrate (Brouwer et al. 2013).

Membrane vesicles assays have many advantages (suitable system for different types of transporter studies, can express various proteins, the estimated unbound substrate concentration is useful to calculate relevant kinetic parameters), but they have some drawbacks. One of them concerns lipophilic compounds that may penetrate the membrane by passive diffusion or bind to the membrane, indicating false-negative results in drug-transporter interaction studies. In these cases, alternative tests (e.g., cell monolayers) are preferable to study transporter interactions.

As for suspended hepatocyte cultures, their functionality may be altered due to cryopreservation and membrane leakage (Yang et al. 2016). In addition, possible bias in cells polarity may happen during isolation of the cells.

When it comes to the use of SCH, one of the limitations observed in rat SCH is downregulation of uptake transporters. However, this has not been an issue with human hepatocytes when cultured under appropriate conditions (Kotani et al. 2011). Reduced expression of uptake transporters in hepatocytes can further limit concentration of a drug substrate available for efflux transporters and lead to underprediction of in vitro obtained biliary efflux values. In order to predict relevant in vivo values based on these data, scaling factors (SFs) have to be used. But the problem is that these factors are drug- and transporter-specific, and evaluation of the predictive value of such in vitro data is difficult (Cantrill and Houston 2017).

Looking at the type of study, certain approximations are often made in the estimation of kinetic parameters describing transporter-mediated drug permeability in a bidirectional in vitro assay. This particularly concerns K_m value which is assumed to be similar in both directions if the experimental conditions (buffer pH) on both sides are the same. However, some observations indicate that K_m estimates can be direction-dependent (Harwood et al. 2013).

General problem with the inhibition studies is that decreased transport of probe substrate in the presence of the investigational drug does not provide information on the nature of the interaction (investigational drug can be either inhibitor or competitive substrate). However, subsequent substrate studies can indicate whether a drug is a substrate. For this reason, substrate and inhibition tests should be performed in combination.

In Vitro/In Vivo Extrapolation

Quantitative prediction of the influence of transporters on drug disposition, tissue exposure, and DDIs is challenging. Based on the current knowledge and proposed theories on drug-transporter interaction mechanisms, several approaches to link in vitro data to the expected in vivo outcomes (in vitro-in vivo correlation (IVIVC) or IVIVE) have been suggested, and they will be described in this section.

ITC group and regulatory agencies have provided recommendation on whether to conduct follow-up clinical transporter studies, whereas the decision is made based on in vitro results and IVIVE. To make a decision, data from multiple in vitro assays are needed, along with supportive data from preclinical species and first-in-human studies (if available). Other data such as patient population, drug therapeutic index, safety profile, and therapeutic indications indicating likely co-administration of other drugs that are substrates or inhibitors for the same transporter pathways should also be considered when planning in vivo interaction studies. To exemplify, if in vitro experiments suggest that a drug is a substrate for renal tubule transporters, this finding should be confirmed in the in vivo study. Moreover, if the investigational drug is likely to be co-administered with a drug that is a known substrate or inhibitor of OCT, OAT, or MATE transporters, potential interaction should be assessed before phase III (preferably before phase II) clinical trials.

In vitro obtained K_m and V_{max} values are preferable transporter kinetic parameters to be used for IVIVE. Clearance value can also be considered, but it should be kept in mind that this parameter does not take into account saturable kinetics of drug transport. A major problem when using in vitro kinetic data for IVIVE is that these values depend upon the applied method, so the experimentally obtained values may vary considerably.

As for the in vivo parameters, drug plasma levels have been traditionally used for IVIVE to assess the interaction of transporters in vivo. However, there are cases when plasma levels do not follow changes in drug tissue exposure, and in these cases, IVIVE may lead to erroneous estimates regarding the interactions with transporters. Technological achievements in imaging techniques may help to overcome these issues. Namely, imaging techniques such as positron-emission tomography, single-photon emission computed tomography, magnetic resonance imaging, etc. can provide valuable information on drug-transporter interactions in vivo, including tracking of dynamic changes in tissue drug concentration and estimation of critical transporter-related kinetic parameters. These data can eventually be correlated to the in vitro findings.

Due to the differences in the expression and activity of transporters in the in vitro systems and in vivo environment, quantitative prediction of transporter-mediated drug PK, tissue exposure, and DDI based on in vitro data requires the use of scaling factors. A scaling factor is defined by dividing fitted in vivo clearance value with the in vivo clearance predicted based on in vitro measurements. One of the suggested approaches is to estimate the transporter-specific scaling factor as the average (geometric mean) factor for different drug substrates, assuming that in vitro measurements for different drugs used to calibrate the SF are done in the same system (Jones et al. 2012). Another approach concerns estimation of the compound-specific SF (Poirier et al. 2009). This empirical value can be obtained by fitting with the animal PK data, and then the same factor can be used to predict human PK based on preclinical data (e.g., in human hepatocytes).

An ideal approach to correlate in vitro to in vivo data is to take into account the effect of all the important transporters enrolled in the disposition of a drug. Such a complex in vitro system that accounts for the effect of all the transporters in different tissues does not exist, nor is likely to be established. SCH model can be used to gain knowledge on the involvement of multiple transporters in liver. Alternatively, multiple assays can be combined (e.g., cell-based systems, membrane vesicles, substrate and inhibition studies) to assess the contribution of each pathway separately.

In general practice, IVIVE approaches for transporter-mediated drug processes can be regarded as static or dynamic. Static approaches are based on drug clearance data (e.g., "Qgut model" for intestinal bioavailability, "well-stirred model," and "extended clearance model" for hepatic drug clearance, "well-stirred renal model") and generally do not take into account changes in drug (substrate or inhibitor) concentration during dosing interval. Also, plasma or luminal concentrations are often used as a substitute for intracellular drug concentration. Moreover, this approach may be suitable to evaluate the effect of a single transporter, but assessing the interaction of multiple transporters or enzyme-transporter interplay is rather difficult. Although this approach has limitations, it is widely used in the initial phase of transporter studies to provide preliminary information on possible drug-transporter interaction before conducting additional in silico and/or in vivo studies. On the other hand, dynamic approaches usually refer to the application of more complex PBPK simulations and modeling tools that can track changes in drug concentration in plasma and tissues as a result of parallel processes a drug undergoes in the organism, including interaction with multiple transporters and/or enzymes. Basic principles of both approaches will be described in the following text.

IVIVE for Oral Absorption

Drugs that are substrates for intestinal absorption may display dose-dependent nonlinear pharmacokinetics, whereas increased drug concentrations (exceeding saturation capacity of transporter) lead to nonproportional decrease in drug absorption (for apical influx transporters) and, in opposite, increase in drug absorption (for apical efflux transporters). For drugs that are substrates for both apical influx and efflux transporters (e.g., quinidine), the situation can be more complicated. Although it is not a rule that substrates for intestinal transporters will exhibit erratic absorption, in vitro screening of possible interactions with

intestinal transporters in early drug development phase, and concomitant IVIVE, may help to detect problems with poor oral drug bioavailability.

The influence of transporter interference on drug absorption depends upon its expression in the intestinal tissue and activity for a certain drug (expressed as K_m). In order to establish correlation between in vitro and in vivo data for drugs that are substrates for intestinal transporters, several aspects need to be considered. First, the in vitro component of IVIVC comprises both drug-related and system-related parameters. Drug-related parameters are addressed in the expression for in vitro intestinal drug clearance:

$$CL_{int,active} = \frac{V_{max}}{K_m + C} \quad (20)$$

where $CL_{int,active}$ is drug intrinsic clearance due to active transport and C unbound drug concentration at the transporter binding site.

Furthermore, by applying the concept of the Q_{gut} model, whereas "drug flow" through the enterocytes (Q_{gut}) encompasses both perfusion and permeability drug transport (Eq. 21), intestinal drug bioavailability (F_g) can be predicted based on drug permeability and intrinsic clearance ($CL_{int,gut}$) in the enterocytes (Eq. 22).

$$Q_{gut} = \frac{CL_{perm} \times Q_{ent}}{CL_{perm} + Q_{ent}} \quad (21)$$

$$F_g = \frac{Q_{gut}}{Q_{gut} + \left(f_{u,gut} + CL_{int,gut}\right)} \quad (22)$$

Here, Q_{ent} is blood flow through the enterocytes, CL_{perm} drug clearance defining permeability through the enterocytes, and $f_{u,gut}$ the fraction of unbound drug in the enterocytes.

This model seems to be well predictive for drugs with relatively high F_g ($F_g > 0.5$), but not for drugs with high intestinal clearance because saturation phenomena and nonlinear processes are not taken into account (Zamek-Gliszczynski et al. 2013).

As for the system-related parameters, it has been shown that drug-transporter kinetic determinants also depend on the experimental system (e.g., direction-dependent K_m value in certain systems, V_{max} value dependent on the expression of transporter in the system) and employed conditions. To overcome these issues, more complex multi-compartment mathematical models have been proposed to estimate intrinsic kinetic parameters of active drug transport (Harwood et al. 2013).

IVIVE for Hepatobiliary Transport

Several successful attempts to establish a relationship between in vitro and in vivo biliary clearance data have been reported in literature and among them (Nakakariya et al. 2012). Still, the predictive power of the proposed extrapolation methods is hard to determine, since available data on the rate and extent of biliary elimination of drugs from clinical studies are rather limited.

I. One of the approaches to estimate in vivo biliary clearance of a drug is scaling of in vitro obtained values (Eq. 23). But the problem with this approach is that scaling factors seem to be both drug- and transporter-specific (Cantrill and Houston 2017). Also, the overall intrinsic clearance in vivo for OATP substrates can be underestimated if hepatic uptake intrinsic clearance from SCH or plated hepatocytes is used for prediction and very large scaling factor have to be used to bridge the discrepancy (Li et al. 2014):

$$CL_{bile,int,pred} = CL_{bile,int,in\ vitro} \times SF \quad (23)$$

In Eq. 23, $CL_{bile,int,pred}$ is the predicted in vivo intrinsic biliary clearance and $CL_{bile,int,in\ vitro}$ intrinsic biliary clearance determined in vitro.

II. On the other hand, biliary clearance of a drug can be estimated based on a well-stirred hepatic model. To estimate drug intrinsic biliary clearance in vivo ($CL_{bile,int,in\ vivo}$) from in vivo clearance data ($CL_{bile,in\ vivo}$), the following equation applies:

$$CL_{bile,int,in\ vivo} = \frac{Q_p \times CL_{bile,in\ vivo}}{Q_p - CL_{bile,in\ vivo}} \quad (24)$$

where Q_p represents liver blood flow.

This equation can be modified to account solely for the fraction of drug unbound in plasma (Fukuda et al. 2008):

$$CL_{bile,int,in\ vivo} = \frac{Q_p \times CL_{bile,in\ vivo}}{Q_p - \frac{CL_{bile,in\ vivo}}{R_{bp}}} \times \frac{1}{f_{u,p}} \quad (25)$$

where R_{bp} is blood-to-plasma concentration ratio.

In a similar manner, biliary clearance of a drug in vivo can be estimated using in vitro data:

$$CL_{bile,app,pred} = \frac{Q_p \times CL_{bile,app,in\ vitro}}{Q_p + CL_{bile,app,in\ vitro}} \quad (26)$$

Or if the fraction of inbound drug is taken into account:

$$CL_{bile,app,pred} = \frac{Q_p \times f_{u,p} \times CL_{bile,app,in\ vitro}}{Q_p + f_{u,p} \times CL_{bile,app,in\ vitro}} \quad (27)$$

In Eqs. 26 and 27, $CL_{bile,app,pred}$ is the predicted in vivo apparent biliary clearance and $CL_{bile,app,in\ vitro}$ apparent biliary clearance determined in vitro.

There are speculations about whether it is better to use $CL_{bile,app,in\ vitro}$ or $CL_{bile,int,in\ vitro}$ to predict biliary clearance of a drug in vivo, but generally the data go in favor of intrinsic clearance values (Nakakariya et al. 2012). Also, the correlation is improved if unbound plasma concentration is used in the estimation of in vivo values (Fukuda et al. 2008).

Further modification of the method to estimate drug biliary clearance in vivo was proposed by Li et al. (2010), who introduced the so-called g factor that incorporates quantitative data on the amount of hepatobiliary transporters (MRP2, BCRP, and BSEP) in rat SCH:

$$CL_{bile,int,pred,g} = \frac{Q_p \times g \times CL_{bile,int,pred}}{Q_p + g \times CL_{bile,int,pred}} \quad (28)$$

III. The third approach to estimate hepatic drug clearance in vivo is based on the extended clearance concept which takes into account multiple transports and/or metabolic processes when predicting overall drug hepatic intrinsic clearance. Transporter-enzyme-mediated overall hepatic intrinsic clearance ($CL_{int,H}$) is mathematically described as:

$$\begin{aligned} CL_{int,H} &= \left(CL_{int,met} + CL_{int,bile}\right) \\ &\quad \times \left(\frac{PS_{uptake}}{PS_{efflux} + CL_{int,met} + CL_{int,bile}}\right) \\ &= \left(CL_{int,met} + CL_{int,bile}\right) \times K_{p,uu} \end{aligned} \quad (29)$$

where:

$$PS_{uptake} = PS_{active} + PS_{passive} \quad (30)$$

$$PS_{efflux} = PS_{basal\text{-}active} + PS_{passive} \quad (31)$$

In Eqs. 29, 30, and 31, $CL_{int,met}$ is metabolic intrinsic clearance, $CL_{int,bile}$ biliary intrinsic clearance, PS_{active} active uptake clearance, $PS_{basal\text{-}active}$ active efflux, $PS_{passive}$ passive diffusion clearance on either direction across the sinusoidal membrane, and $K_{p,uu}$ the unbound drug concentration in the liver relative to plasma at steady state. $K_{p,uu}$ is difficult to measure in vivo, and an alternative in vitro approach is to calculate ratio of drug uptake (in human hepatocyte suspension) at 37 °C and on ice (when $PS_{passive} \gg (CL_{int,met} + CL_{int,bile})$), as described in Eqs. 32 and 33. Different methods to estimate $K_{p,uu}$ are described in the review of Shitara et al. (2013).

$$K_{p,uu} = \frac{PS_{active} + PS_{passive}}{PS_{passive}} = \frac{PS_{uptake}}{PS_{passive}} \quad (32)$$

$$K_{p,uu} = \frac{K_p(37°C)}{K_p(\text{on ice})} \quad (33)$$

In cases when hepatic uptake is the rate-determining step for drug elimination ($CL_{int,met} + CL_{int,bile} \gg PS_{passive}$ and $PS_{efflux} = 0$), the overall hepatic drug clearance can be expressed as:

$$CL_{int,H} = \left(PS_{active} + PS_{passive}\right) \quad (34)$$

Izumi et al. (2017) evaluated different IVIVE approaches for the prediction of human hepatic clearance of OATP substrates and suggested that the best approach, that yielded reasonable correlation with the overall hepatic intrinsic clearance in vivo, is based on the in vitro parameter that describes the rate-determining step for drug hepatic elimination, i.e., uptake intrinsic clearance (PS_{uptake}) while neglecting hepatic metabolism. In addition, it was shown that introducing $K_{p,uu}$ value to account for the difference between drug concentration in hepatocytes and plasma improves correlation between the predicted and in vivo observed clearance values. PS_{uptake} can be estimated from the initial slope of linear equation that illustrates changes in the ratio of hepatocytes to buffer drug concentration during time in relation to the ratio of drug exposure and concentration in the buffer, as shown in Eq. 35:

$$\frac{X_H}{C_{buffer}} = PS_{uptake} \times \frac{AUC_{(0-t),buffer}}{C_{buffer}} + V_0 \quad (35)$$

In Eq. 35, X_H is the amount of drug in hepatocytes, C_{buffer} concentration of drug in the incubation buffer, $AUC_{(0-t),buffer}$ area under the drug concentration-time curve in the incubation buffer, and V_0 initial distribution volume that represents instantaneous drug adsorption to the surface of hepatocytes.

PS_{uptake} determined in vitro at 37 °C can be transposed to the corresponding in vivo value (using hepatocellularity for the human liver and relevant SFs). However, PS_{uptake} values obtained in human hepatocyte suspensions are subjected to inter-batch variability, so using batch-specific SF can improve the predictions (Izumi et al. 2017).

IVIVE for Renal Clearance

In vitro-in vivo prediction methods for transporter-mediated renal drug clearance have not been well established, and the confidence in the prediction is rather low. Namely, data on transporter-mediated renal drug clearance is difficult to obtain in vitro, due to current inability to design a system that will reproduce physiology and functionality of nephron. Also, there are no available methods to assess renal drug clearance in humans directly.

Renal clearance of drugs (CL_{renal}) is comprised of three processes (glomerular filtration, active tubular secretion, and tubular reabsorption), which is represented in the following equation:

$$CL_{renal} = \left(f_{u,b} \times GFR + CL_{sec}\right) \times \left(1 - F_{reabs}\right) \quad (36)$$

where CL_{sec} is renal secretory clearance and F_{reabs} fraction of reabsorbed drug.

If the well-stirred renal model is assumed, secretory clearance can be expressed as:

$$CL_{sec} = Q_r \times \frac{f_{u,b} \times CL_{int,sec}}{Q_r + f_{u,b} \times CL_{int,sec}} \quad (37)$$

Here Q_r is the renal blood flow and $CL_{int,sec}$ intrinsic secretion clearance which depends upon the activity of uptake and efflux transporters as illustrated in the following equation:

$$CL_{int,sec} = \frac{PS_{influx,b} \times PS_{efflux,a}}{PS_{influx,b} + PS_{efflux,a}} \quad (38)$$

In Eq. 38, $PS_{influx,b}$, and $PS_{efflux,a}$ represent relevant influx and efflux intrinsic clearances through basolateral (b) and apical (a) membranes of proximal tubules.

A common approach to estimate renal active clearance of drugs includes empirical assessment based on GFR and scaling of animal data. However, allometric scaling based on animal data may not be the best option because of the variations in transporters expression and activity between species. Also, there is a risk of upregulation or

downregulation of other transporters in transporter-knockout animals which may tangle interpretation of the results. An additional option is to use SFs to translate in vitro (e.g., from kidney slices) to in vivo data. Several studies demonstrated that clearance data obtained in kidney slices can correlate to the in vivo clearance values for some drugs, but SFs needed to be applied (Watanabe et al. 2011). Alternative approach to predict renal drug clearance in humans is to use PBPK modeling, but this approach also has some drawbacks, including uncertainties regarding the expression levels of renal transporters.

IVIVE for Inhibition Studies

In case of transporter inhibition studies, in vitro-generated K_i or IC_{50} values for the investigational drug are correlated to the in vivo drug concentration (intestinal, liver, or plasma concentration, depending on the localization of the transporter of interest), and the outcome is used to estimate potential clinically significant interactions (Fig. 3).

The criteria to evaluate a drug inhibition potential are not harmonized between the leading regulatory authorities, and in addition recently proposed FDA draft guideline (FDA 2017) introduced certain changes in comparison to the FDA guideline from 2012 (FDA 2012). In summary, the following criteria and cutoff values are recommended, depending on the regulatory document:

I. Orally administered investigational drug has a potential to inhibit P-gp or BCRP transporters in vivo if:
   $I_{gut}/IC_{50} \geq 10$ (FDA 2017)
   $I_1/IC_{50} \geq 0.1$ or $I_{gut}/IC_{50} \geq 10$ (FDA 2012; PMDA 2014)
   $I_{1u}/(K_i$ or $IC_{50}) \geq 0.02$ or $I_{gut}/(K_i$ or $IC_{50}) \geq 10$ (EMA 2012)
   Here I_1 and I_{1u} are total (unbound and bound) and unbound systemic concentrations of the inhibitor, respectively.
   Comparison of the prediction performance of different criteria indicated superiority of the criterion suggested in the novel FDA draft guideline (Zhou et al. 2016). However, one clear disadvantage of using IC_{50} value for the assessment of P-gp inhibition potential of a drug is that it can vary considerably between the laboratories, which may lead to inconsistency in the interpretation of the results. To minimize false results, it is advisable to run inhibition studies with a number of well-known inhibitors (Zamek-Gliszczynski et al. 2013).
II. Investigational drug has a potential to inhibit OATP1B1/OATP1B3 transporters in vivo if:
   $R = 1 + I_{u,in,max}/IC_{50} \geq 1.1$ (FDA 2017)
   $C_{max}/(K_i$ or $IC_{50}) \geq 0.1$ and $R \geq 1.25$ (FDA 2012)
   $R \geq 1.04$ (EMA 2012)
   $R \geq 1.25$ (PMDA 2014)
   where R represents the ratio of victim drug AUC in the presence and absence of inhibitor (investigational drug). Again, criterion indicated in the novel FDA draft guidance seems to be the most appropriate (Vaidyanathan et al. 2016). As in the case of P-gp inhibition, IC_{50} values also show large variability depending on the substrate drug and experimental conditions (e.g., with or without preincubation step), and this may influence the validity of the suggested cutoff criteria. Additional problem with the proposed R criteria is that it treats hepatic drug transport solely as OATP-mediated process and neglects the contribution of passive diffusion and/or efflux and hepatic drug metabolism. This often leads to overestimation of the contribution of active hepatic uptake. A way to overpass this limitation is to use relative activity factors (RAFs) when estimating contribution of a specific transporter in the overall hepatic uptake based on the in vitro data.
III. Investigational drug has a potential to inhibit OAT transporters in vivo if:
   $C_{max,u}/IC_{50} \geq 0.1$ (FDA 2012, 2017)
   $C_{max,u}/(K_i$ or $IC_{50}) \geq 0.02$ (EMA 2012)
   $C_{max,u}/IC_{50} \geq 0.25$ (PMDA 2014)

According to the results of Dong et al. (2016), the most suitable criterion for OAT

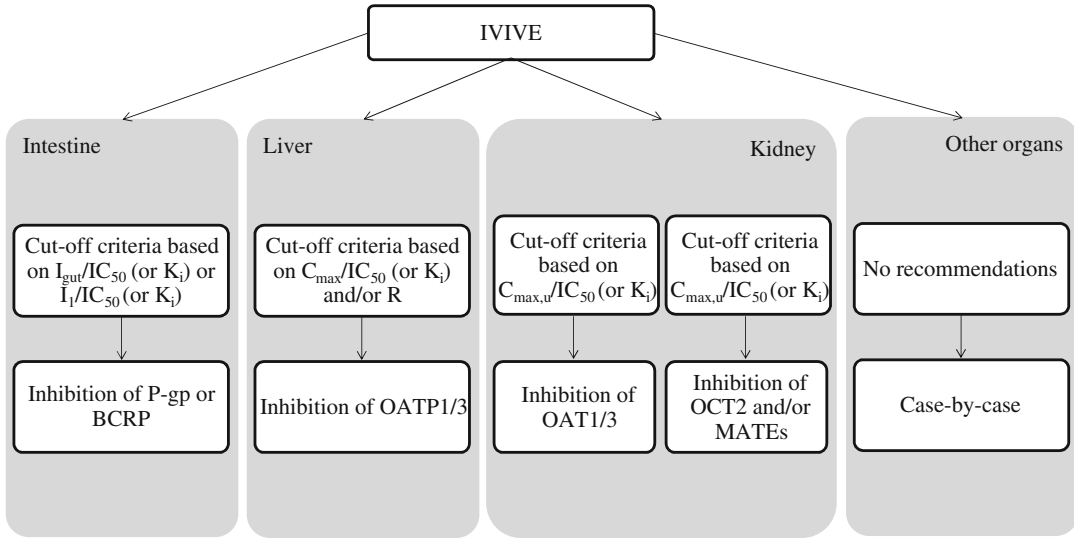


Fig. 3 Recommendations on the in vivo investigation of a drug inhibition potential for major transporters based on in vitro data

transporters appear to be the one proposed by FDA and PMDA.

IV. Investigational drug has a potential to inhibit OCT2 and/or MATE transporters in vivo if:
   $C_{max,u}/IC_{50} \geq 0.1$ (OCT2) or $C_{max,u}/IC_{50} \geq 0.02$ (MATE) (FDA 2017)
   $C_{max,u}/IC_{50} \geq 0.1$ (OCT2) (FDA 2012)
   $C_{max,u}/(K_i \text{ or } IC_{50}) \geq 0.02$ (OCT2) or $C_{max,u}/(K_i \text{ or } IC_{50}) \geq 0.02$ (MATEs) (EMA 2012)
   $C_{max,u}/IC_{50} \geq 0.25$ (OCT2) or $C_{max,u}/IC_{50} \geq 0.25$ (MATEs) (PMDA 2014)

Based on the findings of Dong et al. (2016), it seems the best to apply different cutoffs for OCT2 and MATEs, as proposed in the new FDA guideline. In each of the abovementioned cases, other cutoff values may also be considered if properly justified using an in vitro system calibrated with known inhibitors and non-inhibitors. But to highlight again, for each investigational drug and interacting transporter, the final decision on whether to conduct clinical DDI studies will depend on the therapeutic indication of the investigational drug and likeliness of co-medication with drugs that are known substrates for a particular transporter.

PBPK Modeling

In vivo systems are quite complex, and their correct representation in the vitro environment is not feasible. Therefore, in silico tools that represent some of the complexity of the in vivo conditions provide a valuable mean to assess bioperformance of drugs and elucidate contribution of different mechanisms on drug pharmacokinetics.

PBPK models are physiologically based mathematical models that use a series of differential equations to simulate drug transit and bioperformance in the body. These models usually comprise a set of compartments (organs, tissues, and their substructures) linked by the vascular system. They integrate various physiological data (e.g., characteristics of different cells and spatial structures, blood flow rates) and are able to simulate numerous physiological and biochemical processes including transporter-mediated drug disposition and enzyme-mediated metabolism. PBPK models go even beyond feasible experiments since they are able to test mechanical hypotheses and assess parameters that are difficult or impossible to measure in vivo. Also, simulation and modeling may indicate

the involvement of processes and mechanisms affecting drug bioperformance that have not been identified in vitro or in vivo.

Some of the commercially available software packages for PBPK modeling have integrated data on the localization and expression levels of certain human transporters. Data on some animal transporters are also provided. If information on the suspected transporter is not included in the software database, these data have to be input manually, and they should preferably be experimentally obtained. Relative expression of a transporter in organs and tissues can be determined by traditional methods such as immunoblotting and RNA-based methods. More recently, advances in quantitative proteomics enabled determination of absolute transporter abundances in human tissues and commonly used in vitro systems. But these data should be used with caution because obtained expression levels can vary depending on, e.g., age, gender, disease state, and especially drug history, along with specificities related to the applied methodology.

PBPK modeling tools have been accepted by the regulatory authorities in the USA and EU, and nowadays labeling information for several marketed drugs include information on potential DDIs based solely on in silico modeling results. But, these examples are mostly related to simulations regarding the influence of metabolic enzymes on DDIs. As yet, PBPK modeling has not been exploited much to support regulatory submission that address issues related to transporter-mediated interactions, and examples of such submissions are rather scarce. Some of the examples have been listed by Pan et al. (2016) and include the employment of PBPK modeling to understand the role of (i) P-gp on intestinal absorption and possible DDIs of naloxegol and ceritinib as P-gp substrates and ibrutinib as P-gp inhibitor and (ii) hepatic OATP on possible DDIs with simeprevir as OATP substrate.

Tissue Models

In a PBPK model, each tissue is displayed as a series of separate compartments. A common tissue structure is composed of vascular, intracellular, and extracellular compartment. The intestine, liver, and kidney (and other organs such as the lungs, brain) models are usually more complex. Drug partitioning between extra- and intracellular space depends on passive diffusion through the cellular membrane and active transit mediated by membrane transporters (influx and/or efflux). In line with this, tissues can be treated as perfusion-limited (drug transfer is limited by blood flow rate through the tissue) or permeability-limited (drug transport is a saturable process partly governed by the expression and activity of membrane transporters, besides passive diffusion) (Fig. 4). Passive diffusion is characterized by the PSt factor (permeability-surface area for the given tissue) (Eq. 39), while carrier-mediated transport is described by Michaelis-Menten equation and transporter-specific kinetic parameters (K_m and V_{max}). The PSt for each tissue can be scaled from the PSt for the liver, using the cell volume (V_{cell}) in each tissue (Eq. 40). Liver PSt can be estimated from drug passive permeability determined in hepatocyte culture. K_m and V_{max} values are usually determined in the in vitro experiments, and, based on the assay type, they can be scaled to

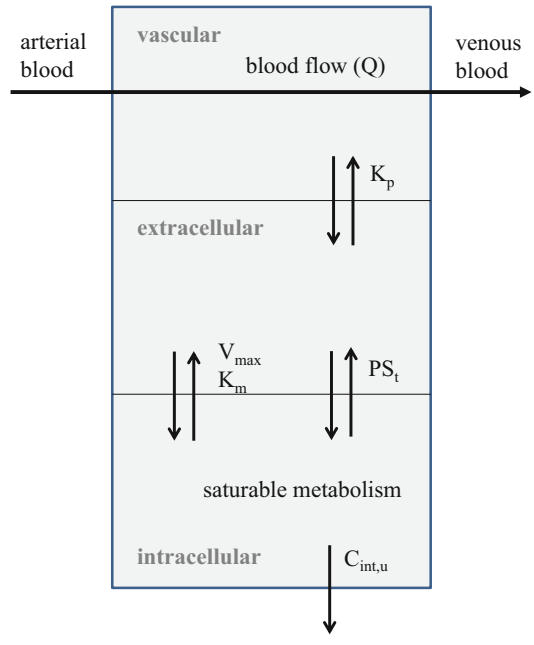


Fig. 4 Schematic representation of permeability-limited tissue in the PBPK model

the relevant in vivo values. In vitro intrinsic clearance can also be scaled to the relevant in vivo value. Alternatively, K_m and V_{max} might be obtained by fitting PBPK results to the in vivo PK data, when in vivo data are available (after clinical studies).

$$PS_t = P_{app} \times S \qquad (S - \text{exchange surface area}) \tag{39}$$

$$PS_{t(\text{tissue1})} = PS_{t(\text{tissue2})} \times \frac{V_{cell(\text{tissue1})}}{V_{cell(\text{tissue2})}} \tag{40}$$

Knowledge of the expression levels of transporters in different tissues is necessary to enable adequate scaling from in vitro and preclinical studies to human data. Namely, if relative expression levels of a transporter (relative amount in each compartment/tissue compared to its V_{max} measurement environment) are known, then in vitro V_{max} can be scaled to the in vivo value. In order to do so, knowledge on the expression of transporter in the in vitro system used for V_{max} determination is also required (e.g., if hepatocytes are used, it can be assumed that the expression level of transporter in vitro is equal to its expression level in the liver). If the expression level in vitro is not known, the obtained V_{max} needs fitting to the in vivo data.

To extrapolate in vitro K_m to the in vivo value, it is necessary to know relevant drug concentration in vitro (e.g., unbound intracellular concentration for the efflux transporters). Herein, unbound concentration of drug in a tissue depends upon unbound drug concentration in plasma, plasma/tissue partitioning, tissue binding, cellular membrane permeability, and degradation via cellular metabolic pathways. There are several methods to estimate drug partitioning into tissues (expressed as K_p value), but the choice of equations depends upon the processes a drug is assumed to undergo in the organism. For instance, lipophilic drugs with high passive permeability will most likely pass cellular membranes by simple passive diffusion, and distribution into body tissues will be perfusion-limited. In these cases, K_p can be calculated from drug physicochemical properties (e.g., molecular weight, lipophilicity, dissociation constant(s), blood-to-plasma concentration ratio, fraction unbound in plasma and tissue) (Kuepfer et al. 2016).

A combined equation that defines drug entry into extracellular tissue compartment, considering three-compartment tissue model (vascular-extracellular-intracellular), and combination of passive diffusion, active transport, and tissue metabolism are given in a following form:

$$\begin{aligned}\left(V_{ec} + \frac{V_v \times R_{bp}}{K_p}\right) \times \frac{dC_{ec}}{dt} &= Q_T \times \left(C_{art} - \frac{C_{ec} \times R_{bp}}{K_p}\right) - PS_t \\ &\times \left(C_{ec,u} - C_{ic,u}\right) - \sum_{i=1}^{nInfTr} \\ &\times \frac{V^i_{max} \times C_{ec,u}}{K^i_m + C_{ec,u}} + \sum_{j=1}^{nEffTr} \\ &\times \frac{V^j_{max} \times C_{ic,u}}{K^j_m + C_{ic,u}}\end{aligned} \tag{41}$$

In a similar manner, change in intracellular drug concentration over time can be expressed as:

$$\begin{aligned}V_{ic} \times \frac{dC_{ic}}{dt} &= PS_t \times \left(C_{ec,u} - C_{ic,u}\right) \\ &+ \sum_{i=1}^{nInfTr} \frac{V^i_{max} \times C_{ec,u}}{K^i_m + C_{ec,u}} \\ &- \sum_{j=1}^{nEffTr} \frac{V^j_{max} \times C_{ic,u}}{K^j_m + C_{ic,u}} \\ &- \sum_{k=1}^{nEnz} \frac{V^k_{max} \times C_{ic,u}}{K^k_m + C_{ic,u}} \\ &- \left(CL_{ic,u} \times C_{ic,u}\right)\end{aligned} \tag{42}$$

In Eqs. 41 and 42, C_{ec} and C_{ic} are drug concentrations in extracellular and intracellular compartment, respectively ($C_{ec,u}$ and $C_{ic,u}$ relevant unbound drug concentrations); V_{ec} and V_{ic} volumes of extracellular and intracellular compartment, respectively; V_v volume of vascular compartment; Q_T tissue blood flow; C_{art} arterial blood concentration; $CL_{ic,u}$ unbound intrinsic clearance in tissue; and nInfTr, nEffTr, and nEnz

numbers of influx transporters, efflux transporters, and enzymes, respectively.

Decision on whether to use perfusion- or permeability-limited model for a certain organ might be difficult to make, since both passive and active diffusion contribute to the overall drug transport. With PBPK modeling, it might be the best to test different hypotheses and analyze the outcomes, to eventually choose the most suitable drug-specific model.

As noted before, PBPK models for certain organs/tissues can be more specific to capture the complexity of physiological structure and processes.

In novel PBPK models, gastrointestinal tract (GIT) is represented as a series of separated compartments defined by a number of physiological parameters, including the expression of some major transporters and metabolic enzymes. Based on the relevant drug-related and physiological parameters, a series of differential equations is used to describe drug transport, dissolution, and absorption in various segments of the GIT.

Kidney model comprises an additional kidney tubule compartment to account for drug active and passive secretion and reabsorption, in addition to filtration and metabolic clearance. Kidney filtration can be estimated based on GFR, while the other transport processes (perfusion- or permeability-limited) can be simulated based on the input data from in vitro or animal studies (as for other tissues).

Hepatic PBPK models include the additional gall bladder compartment to account for biliary excretion of a drug. The simulated processes include emptying of the gallbladder over a defined period of time and reabsorption of a drug (enterohepatic circulation model). Thereof, the necessary input kinetic parameters regarding hepatic drug transport include intrinsic passive diffusion clearance ($CL_{int,pass}$), intrinsic uptake clearance ($CL_{int,uptake}$), and intrinsic biliary clearance ($CL_{int,bile}$), while basolateral efflux is usually neglected. Input hepatic clearance values for PBPK are usually obtained using SCH, and in this case, linear kinetics is assumed. An alternative is to use K_m and V_{max} values to simulate nonlinear hepatic drug clearance. These values can be input directly in a PBPK model or transformed into in vivo hepatic clearance based on a well-stirred hepatic model (described in "In Vitro/In Vivo Extrapolation" section).

There are different approaches to scale transporter-mediated intrinsic clearance obtained in vitro to the relevant in vivo parameters, some of them reviewed by Yang et al. (2016).

I. Simple scaling based on physiological parameters (e.g., liver weight and number of hepatocytes per g liver) often lead to overprediction or underprediction of the relevant in vivo values, and therefore, transporter (or compound)-specific SFs should be used. These empirical SFs can be estimated by comparing in vitro data to the estimated in vivo clearance from intravenous data (human or animal).
II. Another way to estimate in vivo clearance values from in vitro data is based on the quantitative data on transporter(s) expression/abundance.

In addition, relative expression factor (REF) and RAF were proposed as correction factors to bridge the gap between data parameters obtained in different systems (e.g., human hepatocytes vs. recombinant systems or in vitro transport clearance vs. in vivo secretory clearance) and facilitate scaling of in vitro to in vivo data. These factors were initially introduced for metabolic enzymes, but the same concept has been applied to transporters. RAF refers to the difference in activity between the in vitro and in vivo system (based on intrinsic drug clearance in the tissue or V_{max}), while REF describes the difference in transporter expression between in vitro and in vivo system (based on relative mRNA or protein quantification). Absolute scaling factor that describes the difference in "functional transporter expression" (i.e., transporter expression based on absolute protein quantitation and activity) between in vitro and in vivo system can also be used, but kinetic data that link transporter expression with its function are scarce.

There are numerous examples of successful application of PBPK modeling approach to estimate the impact of transporters on drug

disposition and DDIs. PubMed database search on terms "PBPK" and "transporters," performed on August 28, 2018, retrieved 133 results, whereas more than 60% of these publications emerged in the last 5 years.

Critical Assessment of IVIVC Methods

Despite the widespread use of IVIVE and in silico predictions to project ADME behavior of drugs, there are still many challenges associated with translation of in vitro to in vivo data. One of them is lack of precise data on transporter expression and activity in the in vitro systems, preclinical species, and humans. Furthermore, ethnic variability and polymorphism in the transporter-encoding genes may cause variations in transporter expression in different tissues and, consequently, distinct effects on a drug PK and DDIs. Statins are a typical example for the influence of transporter (OATP1B1) polymorphism on drug exposure to central and peripheral tissues, i.e., increased exposure to peripheral tissues may lead to myopathy. Prediction of drug-transporter interactions for patients with organ impairment or for specific population groups (pediatrics, geriatrics, pregnant women) is also challenging. It is well known that the expression and activity of transporters changes during aging and in certain physiological and pathological states. But due to the ethical reason, in it is not always feasible to assess the impact of transporters on drug disposition in vivo. Here, PBPK modeling might be a helpful alternative.

Another issue regarding IVIVE is related to the use of blood or plasma drug concentrations as a substitute for the unbound drug concentration in tissues to assess transporter-mediated drug interaction in vivo, which may be a reason for false results, e.g., if DDI is reflected on drug tissue concentration, but not on plasma concentration. In these cases, PBPK modeling may be a useful alternative to in vivo studies to estimate drug tissue concentration. Another available option is to determine drug distribution in different organs and tissues using imaging techniques such as positron-emission tomography.

PBPK modeling undoubtedly offers numerous advantages in addressing a wide range of PK issues. But, no matter how advanced, PBPK models are not self-sufficient, and they require input and/or validation data from in vitro and in vivo studies. Namely, due to the gaps in our current knowledge about human physiology and transporter expression and function, modeling strategy is often driven by assumptions. Also, there are issues regarding the validity of the in vitro data used as inputs or the conformance of the applied modeling approach. This has been elaborated in the review of Yang et al. (2016). In brief, scaling of transporter-mediated in vitro clearance values (or relevant K_m and V_{max} values), based on physiological parameters, may lead to underestimation or overestimation of the in vivo clearance values. In fact, most of the published examples regarding PBPK modeling of transporters' effect relay on the use of scaling factors to bridge the differences in the expression and activity of transporters between in vitro systems and preclinical species and humans. But as mentioned before, these values are compound-specific, system-specific, and species-specific, and their introduction in a PBPK model brings some distrust in the prediction results.

But despite these limitations, it is rational to believe that new knowledge, and consequently, updates in PBPK models (e.g., regarding model components and organ substructure), will enhance predictability of these tools and lead to increased confidence in in silico results and more wider use of modeling tools.

Concluding Remarks

In recent years, considerable scientific progress, advanced technologies, and new knowledge have enabled deeper understanding of the role of transporter in drug disposition and DDIs. Huge efforts of a number of scientists, research groups, working parties, and regulatory organizations resulted in established methods for in vitro and in vivo assessment of transporter-mediated drug interactions. In addition, development of predictive mathematical models enabled translation of in

vitro to in vivo data through the means of IVIVC/IVIVE. As a step upward, in silico PBPK models provided an integrative in vitro-in vivo-in silico platform to mechanistically explain drug interference with transporters. However, transporter science is evolving, and information presented in this chapter, including in vitro setups, decision criteria, IVIVE methods, etc., reflects current thinking on specific transporter issues, which may change when more transporter-related data become available.

References and Further Reading

Balimane PV, Marino A, Chong S (2008) P-gp inhibition potential in cell-based models: which "calculation" method is the most accurate? AAPS J 10:577–586

Bircsak KM, Gibson CJ, Robey RW et al (2013) Assessment of drug transporter function using fluorescent cell imaging. Curr Protoc Toxicol 57:23–26

Brouwer KL, Keppler D, Hoffmaster KA et al (2013) In vitro methods to support transporter evaluation in drug discovery and development. Clin Pharmacol Ther 94:95–112

Camenisch G, Riede J, Kunze A et al (2015) The extended clearance model and its use for the interpretation of hepatobiliary elimination data. ADMET DMPK 3:1–14

Cantrill C, Houston JB (2017) Understanding the interplay between uptake and efflux transporters within in vitro systems in defining hepatocellular drug concentrations. J Pharm Sci 106:2815–2825

De Bruyn T, Chatterjee S, Fattah S et al (2013) Sandwich-cultured hepatocytes: utility for in vitro exploration of hepatobiliary drug disposition and drug-induced hepatotoxicity. Expert Opin Drug Metab Toxicol 9:589–616

Dong Z, Yang X, Arya V et al (2016) Comparing various in vitro prediction criteria to assess the potential of a new molecular entity (NME) to inhibit organic anion transporter 1 AND 3 (OAT1 and OAT3) in vivo. Clin Pharmacol Ther 99(Suppl 1):S94–S95

El-Kattan AF, Varma MV (2018) Navigating transporter sciences in pharmacokinetics characterization using the extended clearance classification system. Drug Metab Dispos 46:729–739

European Medicine Agency (EMA), Committee for Human Medical Products (CHMP) (2012) Guideline on the investigation of drug interactions. https://www.ema.europa.eu/documents/scientific-guideline/guideline-investigation-drug-interactions_en.pdf. Accessed 29 Sept 2018

Fan PW, Song Y, Berezhkovskiy LM et al (2014) Practical permeability-based hepatic clearance classification system (HepCCS) in drug discovery. Future Med Chem 6:1995–2012

FDA Guidance for Industry (2012) Drug interaction studies – study design, data analysis, implications for dosing, and labeling recommendations (draft guidance). https://www.xenotech.com/regulatory-documents/2012/2012_guidance.aspx. Accessed 29 Sept 2018

FDA Guidance for Industry (2017) In vitro metabolism and transporter-mediated drug-drug interaction studies (draft guidance). https://www.fda.gov/downloads/Drugs/Guidances/UCM581965.pdf. Accessed 29 Sept 2018

Fukuda H, Ohashi R, Tsuda-Tsukimoto M et al (2008) Effect of plasma protein binding on in vitro-in vivo correlation of biliary excretion of drugs evaluated by sandwich-cultured rat hepatocytes. Drug Metab Dispos 36:1275–1282

Gertz M, Cartwright CM, Hobbs MJ et al (2013) Application of PBPK modeling in the assessment of the interaction potential of cyclosporine against hepatic and intestinal uptake and efflux transporters and CYP3A4. Pharm Res 30:761–780

Glavinas H, von Richter O, Vojnits K et al (2011) Calcein assay: a high-throughput method to assess P-gp inhibition. Xenobiotica 41:712–719

Gozalpour E, Fenner KS (2018) Current state of in vitro cell-based renal models. Curr Drug Metab 19:310–326

Harwood MD, Neuhoff S, Carlson GL et al (2013) Absolute abundance and function of intestinal drug transporters: a prerequisite for fully mechanistic in vitro-in vivo extrapolation of oral drug absorption. Biopharm Drug Dispos 34:2–28

Izumi S, Nozaki Y, Komori T et al (2017) Comparison of the predictability of human hepatic clearance for organic anion transporting polypeptide substrate drugs between different in vitro-in vivo extrapolation approaches. J Pharm Sci 106:2678–2687

Jones HM, Barton HA, Lai Y et al (2012) Mechanistic pharmacokinetic modeling for the prediction of transporter-mediated disposition in humans from sandwich culture human hepatocyte data. Drug Metab Dispos 40:1007–1017

Kotani N, Maeda K, Watanabe T et al (2011) Culture period-dependent changes in the uptake of transporter substrates in sandwich-cultured rat and human hepatocytes. Drug Metab Dispos 39:1503–1510

Kuepfer L, Niederalt C, Wendl T et al (2016) Applied concepts in PBPK modeling: how to build a PBPK/PD model. CPT Pharmacometrics Syst Pharmacol 5:516–531

Li N, Singh P, Mandrell KM et al (2010) Improved extrapolation of hepatobiliary clearance from in vitro sandwich cultured rat hepatocytes through absolute quantification of hepatobiliary transporters. Mol Pharm 7:630–641

Li R, Barton HA, Varma MV (2014) Prediction of pharmacokinetics and drug-drug interactions when hepatic transporters are involved. Clin Pharmacokinet 53:659–678

Liu X, Chism JP, LeCluyse EL et al (1999a) Correlation of biliary excretion in sandwich-cultured rat hepatocytes and in vivo in rats. Drug Metab Dispos 27:637–644

Liu X, LeCluyse EL, Brouwer KR et al (1999b) Biliary excretion in primary rat hepatocytes cultured in a collagen-sandwich configuration. Am J Phys 277: G12–G21

Momper JD, Tsunoda SM, Ma JD (2016) Evaluation of proposed in vivo probe substrates and inhibitors for phenotyping transporter activity in humans. J Clin Pharmacol 56(Suppl 7):S82–S98

Nakakariya M, Ono M, Amano N et al (2012) In vivo biliary clearance should be predicted by intrinsic biliary clearance in sandwich-cultured hepatocytes. Drug Metab Dispos 40:602–609

Pan Y, Hsu V, Grimstein M et al (2016) The application of physiologically based pharmacokinetic modeling to predict the role of drug transporters: scientific and regulatory perspectives. J Clin Pharmacol 56(Suppl 7):S122–S131

Pfeifer ND, Yang K, Brouwer KL (2013) Hepatic basolateral efflux contributes significantly to rosuvastatin disposition I: characterization of basolateral versus biliary clearance using a novel protocol in sandwich-cultured hepatocytes. J Pharmacol Exp Ther 347:727–736

PMDA Japan (2014) Guideline on the investigation of drug interactions. http://www.nihs.go.jp/mss/T140710-jimu.pdf. Accessed 29 Sept 2018

Poirier A, Cascais AC, Funk C et al (2009) Prediction of pharmacokinetic profile of valsartan in human based on in vitro uptake transport data. J Pharmacokinet Pharmacodyn 36(6):585–611

Sarmento B, Andrade F, da Silva SB et al (2012) Cell-based in vitro models for predicting drug permeability. Expert Opin Drug Metab Toxicol 8:607–621

Shitara Y, Maeda K, Ikejiri K et al (2013) Clinical significance of organic anion transporting polypeptides (OATPs) in drug disposition: their roles in hepatic clearance and intestinal absorption. Biopharm Drug Dispos 34:45–78

Vaidyanathan J, Yoshida K, Arya V et al (2016) Comparing various in vitro prediction criteria to assess the potential of a new molecular entity to inhibit organic anion transporting polypeptide 1B1. J Clin Pharmacol 56 (Suppl 7):S59–S72

Volpe DA, Hamed SS, Zhang LK (2014) Use of different parameters and equations for calculation of IC_{50} values in efflux assays: potential sources of variability in IC_{50} determination. AAPS J 16:172–180

Watanabe T, Kusuhara H, Watanabe T et al (2011) Prediction of the overall renal tubular secretion and hepatic clearance of anionic drugs and a renal drug-drug interaction involving organic anion transporter 3 in humans by in vitro uptake experiments. Drug Metab Dispos 39:1031–1038

Wu CY, Benet LZ (2005) Predicting drug disposition via application of BCS: transport/absorption/elimination interplay and development of a biopharmaceutics drug disposition classification system. Pharm Res 22:11–23

Yang K, Guo C, Woodhead JL et al (2016) Sandwich-cultured hepatocytes as a tool to study drug disposition and drug-induced liver injury. J Pharm Sci 105:443–459

Yoshino M, Murakami K (2009) A graphical method for determining inhibition constants. J Enzyme Inhib Med Chem 24:1288–1290

Zamek-Gliszczynski MJ, Lee CA, Poirier A et al (2013) International Transporter Consortium (ITC) recommendations for transporter kinetic parameter estimation and translational modeling of transport-mediated PK and DDIs in humans. Clin Pharmacol Ther 94:64–79

Zhou T, Arya V, Zhang L (2016) Comparing various in vitro prediction criteria to assess the potential of a new molecular entity (NME) to inhibit P-glycoprotein (P-GP) in vivo. Clin Pharmacol Ther 99(Suppl 1): S89–S90

Relevance of Transporters in Clinical Studies

54

Bruno Hagenbuch

Contents

Abstract

It has become clear that drug disposition is not just a result of passive diffusion and metabolizing enzymes. Numerous transporters were identified in recent years to be involved in the absorption, distribution, and excretion of essentially all drugs. While transporters of the solute carrier (SLC) family are mainly involved in the uptake of drugs into cells, ATP-binding cassette (ABC) transporters are responsible for their efflux. Among the more than 420 SLC and 47 ABC transporters, only about 25 seem to be important for the disposition of over-the-counter and prescription drugs. Among these the Food and Drug Administration (FDA), the European Medicines Agency (EMA), and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA) have identified seven transporters which need to be tested for investigational drugs and an additional five transporters that are considered to be important. Two of the seven transporters, the multidrug resistance protein 1 (MDR1) and the breast cancer

B. Hagenbuch (✉)
Department of Pharmacology, Toxicology and Therapeutics, The University of Kansas Medical Center, Kansas City, KS, USA
e-mail: bhagenbuch@kumc.edu

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_23

resistance protein (BCRP), are ABC transporters. The other five, the organic cation transporter 2 (OCT2), the organic anion transporter 1 (OAT1) and 3 (OAT3), and the organic anion transporting polypeptide 1B1 (OATP1B1) and 1B3 (OATP1B3), are SLC transporters. If additional transporters become clinically relevant, they may be added by the regulatory agencies to the list or required transporters.

Introduction

Initially, drug disposition, which includes absorption, distribution, metabolism, and excretion, was thought to be mainly the result of passive diffusion and metabolizing enzymes. More recently, however, it became clear that drug transporters play an important role in absorption, distribution, and excretion of drugs and their metabolites (Giacomini et al. 2010; Hillgren et al. 2013). Most drugs are administered orally and have to cross initially a barrier built of polarized epithelial cells in the intestine, the enterocytes (Fig. 1a) (Drozdzik et al. 2014). Once in the portal blood, drugs first reach the liver where they can be removed by what is called first-pass metabolism, a combination of drug uptake into hepatocytes (Stieger and Hagenbuch 2016), metabolism by drug-metabolizing enzymes, and excretion into bile (Fig. 1b) (Pfeifer et al. 2014). Drugs that are not (completely) cleared by this first-pass effect will reach the systemic circulation and will be carried to their target organ where they bind to a receptor or are taken up into the cells to affect their drug target. To act in the brain, drugs also have to cross the blood-brain barrier, another polarized epithelial layer that contains drug uptake and efflux transporters in the plasma membranes facing the blood or the brain (Fig. 1c) (Abdullahi et al. 2017). Eventually, most drugs not cleared by the liver will be excreted via the kidneys by either filtration or secretion. Renal secretion also involves the transport of the drug across the basolateral membrane into the cell and then across the brush-border membrane into the tubule (Fig. 1d) (Liu et al. 2016).

In 2010 the International Transporter Consortium published the first recommendations on which transporters it considered to be important for drug absorption and disposition and extended this list in 2013 (Giacomini et al. 2010; Hillgren et al. 2013). In the meantime, the Food and Drug Administration (FDA), the European Medicines Agency (EMA), and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA) have published guidelines for the pharmaceutical industry listing which drug transporters need to be tested when evaluating an investigational drug (Table 1).

This chapter will review the clinical importance of drug uptake and efflux transporters with an emphasis on those transporters that are highlighted by the regulatory agencies for investigational drugs.

Major Drug Transporters with Clinical Relevance

About 10% of all human genes are transporter related. Among these are ATP-binding cassette (ABC) transporters and solute carrier (SLC) transporters. ABC transporters are primary active efflux transporters that utilize the energy derived from the hydrolysis of ATP to transport their substrates against electrochemical concentration gradients. There are 7 ABC families that contain 47 members and 3 pseudogenes (https://www.genenames.org/cgi-bin/genefamilies/set/417). All the important drug efflux transporters are classified within three ABC families: in family ABCB there are the multidrug resistance protein 1 (MDR1; gene symbol *ABCB1*) and the bile salt export pump (BSEP; *ABCB11*); family ABCC contains the multidrug resistance-associated protein 2 (MRP2; *ABCC1*); and in family ABCG there is the breast cancer resistance protein (BCRP; *ABCG2*) (Table 1 and Fig. 1). The SLC transporters are in general uptake transporters, but some may also mediate the efflux of substrates out of the cell. SCL transporters can be secondary active, transporting their substrates against an electrochemical gradient, or facilitated transporters, transporting their substrates along an

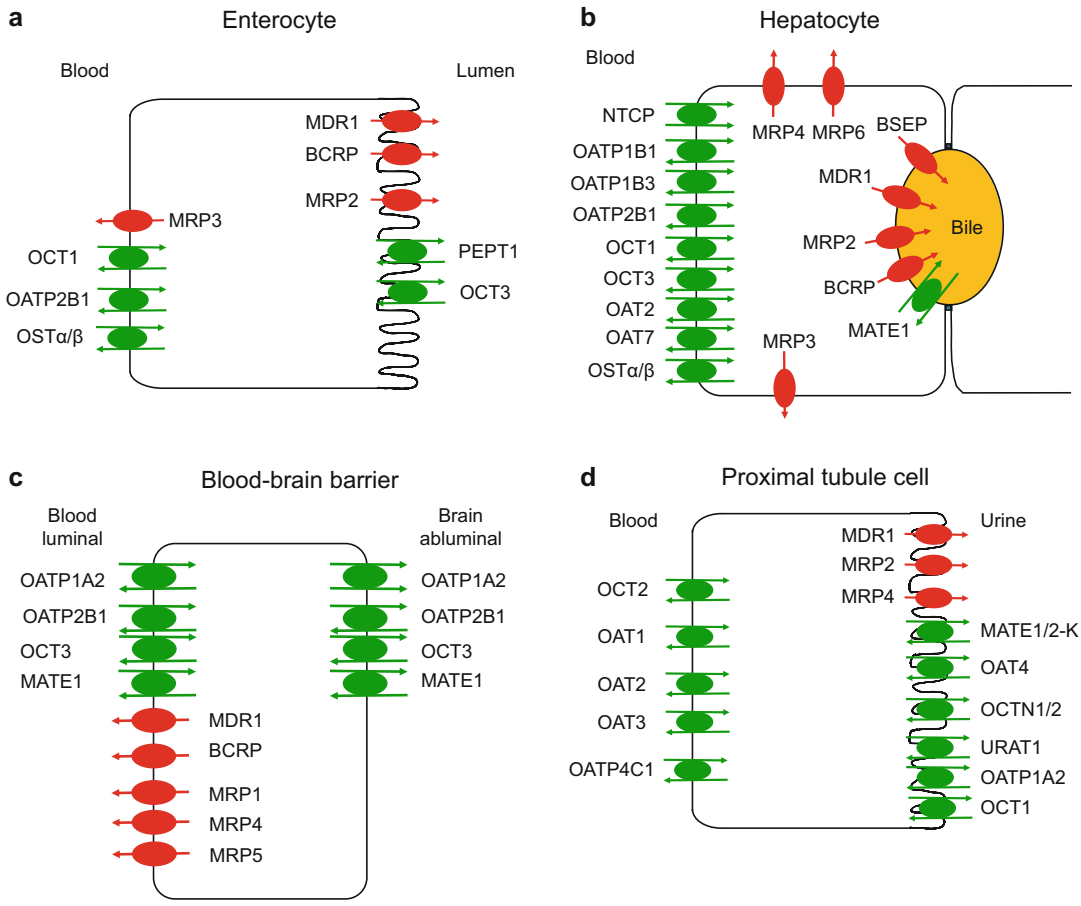


Fig. 1 Schematic localization of transporters involved in the disposition of drugs in the small intestine (**a**), hepatocytes (**b**), the blood-brain barrier (**c**), and the proximal tubule of the kidney (**d**)

electrochemical gradient. There are currently 65 SLC families that contain over 420 members and more than 20 pseudogenes (http://slc.bioparadigms.org/). All of the important drug uptake transporters that are recommended or considered by the regulatory agencies are classified within three SLC families: in the SLC22 family, there are the organic cation transporter 1 (OCT1; *SLC22A1*) and OCT2 (*SLC22A2*), as well as the organic anion transporter 1 (OAT1; *SLC22A6*) and OAT3 (*SLC22A8*); the SLC47 family consists of the multidrug and toxin extrusion protein 1 (MATE1; *SLC47A1*) and MATE2 (*SLC47A2*); and in the SLCO family, there are the organic anion transporting polypeptide 1B1 (OATP1B1; *SLCO1B1*) and OATP1B3 (*SLCO1B3*) (Table 1 and Fig. 1).

ATP-Binding Cassette (ABC) Transporters

The Multidrug Resistance Protein 1 (MDR1)

The gene for the human multidrug resistance protein 1 (MDR1; gene symbol *ABCB1*) was originally identified from multidrug-resistant carcinoma cells in 1986 (Roninson et al. 1986) as the gene encoding P-glycoprotein (Ueda et al. 1986), a glycoprotein of 170 kDa that was originally discovered in 1976 (Juliano and Ling 1976). Since then it has become clear that MDR1 is also expressed in normal cells of the body, in particular in epithelial cells (Schinkel and Jonker 2003).

Table 1 Human transporters involved in the uptake and efflux of endo- and xenobiotics

Protein name	Gene symbol	Recommended or considered by regulatory agencies
MDR1	*ABCB1*	FDA, EMA, PMDA
BSEP	*ABCB11*	Considered by FDA, EMA, PMDA
MRP1	*ABCC1*	
MRP2	*ABCC2*	Considered by PMDA
MRP3	*ABCC3*	
MRP4	*ABCC4*	
BCRP	*ABCG2*	FDA, EMA, PMDA
NTCP	*SLC10A1*	
ASBT	*SLC10A2*	
PEPT1	*SLC15A1*	
PEPT2	*SLC15A2*	
OCT1	*SLC22A1*	Considered by EMA and PMDA
OCT2	*SLC22A2*	FDA, EMA, PMDA
OAT1	*SLC22A6*	FDA, EMA, PMDA
OAT3	*SLC22A8*	FDA, EMA, PMDA
ENT1	*SLC29A1*	
ENT2	*SLC29A2*	
MATE1	*SLC47A1*	PMDA, considered by FDA and EMA
MATE2	*SLC47A2*	PMDA, considered by FDA and EMA
OSTα	*SLC51A*	
OSTβ	*SLC51B*	
OATP1A2	*SLCO1A2*	
OATP1B1	*SLCO1B1*	FDA, EMA, PMDA
OATP1B3	*SLCO1B3*	FDA, EMA, PMDA
OATP2B1	*SLCO2B1*	

Transporters are classified either as ATP-binding cassette (*ABC*) transporters or as solute carrier (*SLC*) transporters
EMA European Medicines Agency, *FDA* Food and Drug Administration, *PMDA* Pharmaceuticals and Medical Devices Agency

In enterocytes, MDR1 is expressed in the apical membrane (Fig. 1a) where it protects the body from toxic xenobiotics. Pumping its substrates out of the cells, MDR1 can restrict substrate uptake and thus affect the bioavailability of numerous drugs. In general, MDR1 substrates are hydrophobic and amphipathic. They can be between 200 and over 4000 Da and are mostly uncharged or positively charged. They include numerous anticancer agents, antimicrobials, several HIV protease inhibitors, immunosuppressant drugs like cyclosporine A, and various cardiovascular drugs including calcium channel blockers and digoxin (Schinkel and Jonker 2003; Terada and Hira 2015; Saidijam et al. 2018). See Lund et al. (2017) for a compilation of drugs that modulate MDR1 activity and can lead to adverse drug-drug interactions. Given the broad substrate spectrum and the fact that several of these drugs often are prescribed together, there is a real possibility for adverse drug-drug interactions either because of inhibition or induction of MDR1. The quinidine-digoxin drug-drug interaction, e.g., could be explained by inhibition of MDR1-mediated efflux of digoxin from enterocytes by quinidine, resulting in increased absorption and thus increased plasma concentration of digoxin. Similar effects due to MDR1 inhibition were reported for amiodarone, dronedarone, and propafenone, three antiarrhythmics that are likely prescribed to patients that also take oral digoxin (Wessler et al. 2013). In contrast, rifampicin administration induced the expression of MDR1 in the duodenum and decreased absorption and thus bioavailability of digoxin (Wessler et al. 2013).

MDR1 is also a key player in the protection of the brain from toxic xenobiotics. It is expressed in the luminal membrane (Fig. 1c) where it prevents the uptake of its substrates from blood into the endothelial cells of the blood-brain barrier (Schinkel and Jonker 2003). As a consequence, the brain penetration of numerous drugs is rather low but can be increased in the presence of MDR1 inhibitors or in the absence of a functional MDR1. This was nicely demonstrated using *mdr1a (−/−)* mice (Schinkel et al. 1996). These mice are lacking the mouse homolog of the human MDR1 protein, and therefore they accumulated drugs in the brain 10- to 100-fold above the levels of their wild-type controls (Schinkel and Jonker 2003).

In the liver (Fig. 1b), MDR1 is expressed at the canalicular membrane of hepatocytes and can impact the excretion of mainly cationic xenobiotics that have been taken up into hepatocytes via OCT1 and OCT3. In proximal tubule cells (Fig. 1d), MDR1 is expressed at the brush-border membrane and can affect the secretion of substrates that have been taken up into the tubular cells via OCT2.

Given that mutations or polymorphisms can lead to an inactive (or overactive) MDR1, it is also important to search for and characterize such gene variants. At least 390 sequence variants have been identified in the coding region of the *ABCB1* gene, but the majority of these variants only occur at low frequencies (<1%). Wolking et al. (2015) reviewed the literature and concluded that multiple studies with the three most frequent polymorphisms c.1236C > T (G412G), c.2677G > T/A (A893S/T), and c.3435C > T (I1145I) with different drug substrates have been performed. Some of them have been associated with effects on drug disposition, response, and toxicity, but overall the findings were conflicting and had limited clinical implications. These findings indicate that in the case of MDR1, inhibition due to drug-drug interactions and induction of expression are predominant and clinically more relevant than polymorphisms.

Because MDR1 can affect the oral bioavailability, the distribution, and the excretion of drugs, the regulatory agencies expect that investigational drugs are tested in vitro whether they are substrates of MDR1. Class I drugs (i.e., highly soluble and highly permeable) according to the biopharmaceutical classification system (Amidon et al. 1995) only need to be tested as MDR1 substrates if there are potential safety concerns regarding distribution into the brain or the kidneys. The in vitro test should either be performed by measuring the transepithelial flux, e.g., using Caco-2 cell layers or by inhibiting the transepithelial flux with at least one known MDR1 inhibitor at a concentration of more than ten times its K_i value. An efflux ratio of at least two or inhibition of the efflux ratio by more than 50% by the known inhibitor indicates that the investigational drug is an MDR1 substrate. If the drug is a substrate, an in vivo study might be necessary based on the safety margin of the drug, on its therapeutic index, and on the fact that a likely co-medication is an MDR1 inhibitor (Lund et al. 2017).

The Bile Salt Export Pump (BSEP)

The human bile salt export pump (BSEP; gene symbol *ABCB11*), a glycoprotein of 170 kDa, is expressed almost exclusively in the canalicular membrane of hepatocytes (Fig. 1b) where it is responsible for the export of mainly conjugated bile acids (Stieger 2011). Normal BSEP function is required because mutations in the *ABCB11* gene that result in either reduced or completely absent BSEP function result in cholestatic liver disease that can be mild (benign recurrent intrahepatic cholestasis type 2) or life threatening (progressive familial intrahepatic cholestasis type 2) (Stieger 2011). Numerous drugs and xenobiotics inhibit BSEP function in vitro, and a correlation of their IC_{50} values with reported hepatotoxicity revealed that if the IC_{50} value is below 25 μM, chances for drug-induced liver injury increased (Morgan et al. 2010; Stieger 2011). However, no correlation between maximal free plasma concentration and BSEP inhibition or liver injury could be established. This is likely due to the fact that the uptake step for these drugs, transport across the basolateral membrane

into hepatocytes, is rate limiting (Hillgren et al. 2013).

Sequencing of the *ABCB11* gene revealed numerous mutations and polymorphisms. For example, the c.1331T>C variant leading to an alanine at position 444 instead of a valine (V444A) is associated with low BSEP expression (Stieger 2011) and was overrepresented in a population with cholestatic liver disease, suggesting that it can contribute to or predispose for liver disease (Droge et al. 2017).

Although not required by the regulatory agencies, the International Transporter Consortium recommends that BSEP function should be tested under certain conditions. The European Medicines Agency states that BSEP inhibitory potential should be considered in particular if plasma bile acid levels are increased in animal studies. Furthermore, if a BSEP inhibitor is given to humans, their serum bile acid levels should be monitored along with liver serum markers because of the potential of drug-induced liver injury (Hillgren et al. 2013).

The Multidrug Resistance-Associated Protein 2 (MRP2)

The multidrug resistance-associated protein 2 (MRP2; gene symbol *ABCC2*) is a 190 kDa glycoprotein expressed in the canalicular membrane of hepatocytes, as well as in the apical membrane of enterocytes and proximal tubular cells (Fig. 1). It was originally identified as a canalicular multispecific organic anion transporter mediating the efflux of conjugated anionic substrates including bilirubin glucuronides and numerous drug conjugates into bile. Its functional absence leads in humans to the Dubin-Johnson syndrome, a rare benign disorder characterized by conjugated hyperbilirubinemia (Schinkel and Jonker 2003). Numerous studies have characterized the role MRP2 plays in the hepatobiliary excretion of drugs, and with its strategic localization in the apical membrane of enterocytes, MRP2 and its inhibition could affect the bioavailability of its drug substrates. Furthermore, it also can contribute directly to the renal excretion of drugs. It seems that both MDR1 and MRP2 play an important role in protecting the human body from potentially toxic xenobiotics.

Given its broad substrate specificity, drug-drug interactions seem possible. However, only a few of these interactions have been described so far. In a recent screening of 124 natural compounds, only 3.2% were inhibitors of MRP2, while the breast cancer resistance protein (BCRP) was inhibited by 36% of the compounds (Sjostedt et al. 2017). Besides inhibitors also stimulators of MRP2 have been characterized in vitro, but so far only a few studies have investigated the stimulatory effect on MRP2 in vivo, and it could be verified in a rat model (Heredi-Szabo et al. 2009).

Numerous genetic polymorphisms have been identified in the *ABCC2* gene, but only a few lead to decreased MRP2 function, and conjugated hyperbilirubinemia is a possible consequence. Because hyperbilirubinemia could be a sign of hepatotoxicity, the International Transporter Consortium recommends that in cases of drug-induced hyperbilirubinemia, inhibition of MRP2 should be tested (Hillgren et al. 2013). Furthermore, the Japanese Pharmaceuticals and Medical Devices Agency suggests that inhibition of MRP2 could lead to increased drug concentrations in hepatocytes or drug-induced increases in plasma concentrations of endogenous compounds.

The Breast Cancer Resistance Protein (BCRP)

The breast cancer resistance protein (BCRP; gene symbol *ABCG2*) is an ABC half-transporter of 75 kDa that probably functions as a dimer. It was originally identified in the multidrug-resistant human breast cancer MCG-7 cell line. Initial characterization demonstrated that expression of BCRP conferred resistance to several anticancer agents (Saidijam et al. 2018). Later studies discovered that similar to MDR1, BCRP is expressed in the apical membrane of many epithelia including the enterocytes, hepatocytes, and endothelial cells of the blood-brain barrier (Fig. 1) and protects the organism from numerous xenobiotics (Terada and Hira 2015).

Functional characterization revealed that besides anticancer agents, BCRP transports numerous drugs from many different classes including antivirals, antibiotics, tyrosine kinase inhibitors, nonsteroidal anti-inflammatory drugs (NSAIDs), and statins (Lee et al. 2015). Although BCRP transports such a wide variety of drugs, drug-drug interactions exclusively due to BCRP are rare, except for limiting oral absorption in the intestine, because of an overlap in substrate specificity with other transporters and drug-metabolizing enzymes. For example, curcumin, a natural polyphenol and the main curcuminoid of turmeric, increased AUC of the NSAID sulfasalazine in healthy volunteers between 2- and 3.2-fold (Lee et al. 2015). Similarly, the AUC of rosuvastatin increased in healthy volunteers in the presence of the immunosuppressant cyclosporine A or the protease inhibitors tipranavir and ritonavir. A suggested mechanism includes inhibition of uptake into hepatocytes via OATP1B1 (see below) and increased absorption due to BCRP inhibition in the intestine (Lee et al. 2015).

BCRP expressed at the blood-brain barrier, together with MDR1, protects the brain from potentially toxic xenobiotics. While this is good for the normal function of the brain, it also limits the brain penetration of drugs that have their target in the brain. One example is imatinib mesylate, for which there is some in vitro evidence that it could be used to treat malignant gliomas. However, a clinical study showed that imatinib mesylate had minimal activity in malignant gliomas, probably because it is a substrate of BCRP (Urquhart and Kim 2009). Therefore, inhibitors of BCRP, and potentially dual BCRP/MDR1 inhibitors, could be useful tools to increase drug delivery to the brain.

More than 80 polymorphisms have been documented in the *ABCG2* gene. However, most of these are rare and found in less than 1% of the population. The most frequent polymorphism, c.421C>A, results in a lysine at position 141 instead of a glutamine (Q141K). This mutation leads to a protein that is less stable than the wild-type and results in a reduced BCRP function, due to reduced plasma membrane expression levels. The allele frequency of this polymorphism is about 5–10% in Caucasians and African-Americans but between 30% and 60% in East Asians (Hira and Terada 2018). In addition, this polymorphism has been identified in a genome-wide association study as being associated with increased serum urate concentrations and an increased risk of gout. This demonstrates another important role BCRP plays in the elimination of uric acid in the intestine (Cleophas et al. 2017).

Because BCRP has the potential to affect the oral bioavailability, the tissue distribution, and the hepatic excretion of drugs, the regulatory agencies treat BCRP similar as MDR1 and expect that investigational drugs are tested in vitro whether they are substrates of BCRP following the same principles as outlined for MDR1 above. Like for MDR1, class I drugs according to the biopharmaceutical classification system are excluded.

Solute Carrier (SLC) Transporters

The Organic Cation Transporters (OCTs) of the SLC22 Family

Organic Cation Transporter 1 (OCT1)

The organic cation transporter 1 (OCT1; gene symbol *SLC22A1*) is a glycoprotein of approximately 70 kDa expressed mainly at the sinusoidal or basolateral membrane of human hepatocytes (Fig. 1b). In addition, OCT1 is expressed at the basolateral membrane of human enterocytes (Fig. 1a), at the apical membrane of proximal tubule cells (Fig. 1d), and in several additional tissues including the lung, heart, skeletal muscle, brain, mammary and adrenal gland, eye, adipose tissue, and immune cells (Koepsell 2013).

OCT1 transports a wide variety of substrates that in general have a molecular weight of less than 500 Da, are mainly hydrophobic, and carry a positive charge. Common model substrates for in vitro studies are 1-methyl-4-phenylpyridinium (MPP^+), *N*-methylquinine, and tetraethylammonium (TEA) (Koepsell 2013). Besides these model substrates, numerous endogenous cationic substrates including the

neurotransmitters acetylcholine, dopamine, norepinephrine, and serotonin are transported by OCT1. Among the drug substrates are antiarrhythmics, antibiotics, anticholinergics, the antidiabetic metformin, antihypertensives, anticancer agents, antivirals, β2-agonists, diuretics, and H2-antagonists (Nies et al. 2011). The importance of OCT1 for metformin uptake into its target cells, human hepatocytes, was demonstrated when in healthy volunteers metformin was co-administered with verapamil, an OCT1 inhibitor. After co-administration with verapamil, metformin did not reduce maximal blood glucose concentrations to the same degree as when the volunteers only received metformin (Patel et al. 2016). In contrast, metformin exhibited a larger reduction in blood glucose levels after healthy volunteers were treated with rifampicin, an agonist of the pregnane X receptor (PXR). As a result OCT1 mRNA levels were increased about fourfold in peripheral blood cells. This study suggests that increased expression of OCT1 resulted in increased hepatic uptake and activity of metformin (Patel et al. 2016).

Numerous single nucleotide polymorphisms were identified in the *SLC22A1* gene, several of which affect metformin efficacy. Reduced function polymorphisms c.181C>T (R81C), c.1201G>A (G401S), c.1260GAT>del (420del), and c.1393G>A (G465R) resulted in higher metformin AUC, higher maximal plasma concentrations, and reduced glucose-lowering effects during oral glucose tolerance testing (Wagner et al. 2016). These findings are consistent with reduced uptake of metformin into hepatocytes. OCT1 is not only important for metformin uptake but mediates the uptake of many additional clinical important drugs. For example, lamivudine, a drug used to treat HIV infection and a substrate of OCT1, was transported less efficiently by the above polymorphic variants, and carriers of these polymorphisms would have lower drug efficiency.

The International Transporter Consortium has listed OCT1 as a clinical relevant transporter because OCT1 activity positively correlates with how patients with chronic myeloid leukemia respond to imatinib and because OCT1 seems to be involved in the interindividual response to metformin (Giacomini et al. 2010). However, so far the regulatory agencies do not require OCT1 to be tested for drug interactions. The European Medicines Agency and the Japanese Pharmaceuticals and Medical Devices Agency consider to add OCT1 to the required drug transporters to be tested in the future.

Organic Cation Transporter 2 (OCT2)

The organic cation transporter 2 (OCT2; gene symbol *SLC22A2*) is a glycoprotein of 555 amino acids and is mainly expressed at the basolateral membrane of proximal tubule cells (Fig. 1d) (Koepsell 2013). There it plays a crucial role in the secretion of organic cations by mediating the first step, uptake across the basolateral membrane into the tubular cells, before these cations are excreted across the brush-border membranes by MATE1 and MATE2K (see below). Additional tissues with minor OCT2 expression are the small intestine, lung, placenta, thymus, brain, and inner ear. Similar to OCT1, also OCT2 has a broad substrate specificity and transports numerous OCT1 substrates but also some distinct compounds (Nies et al. 2011). Model OCT2 substrates include 1-methyl-4-phenylpyridinium (MPP^+); tetraethylammonium (TEA); endogenous monoamines like norepinephrine, dopamine, and serotonin; the antineoplastic drug oxaliplatin; the antiviral lamivudine; the antidiabetic drug metformin; and the antihypertensive atenolol (Yin and Wang 2016).

The combined uptake by OCT2 across the basolateral membrane and the secretion by MATE1/2-K across the brush-border membrane are crucial functions for the renal elimination of metformin. Several drug-drug interactions at OCT2 affect renal metformin secretion. Cimetidine, a histamine H_2-receptor antagonist, is well known to reduce renal clearance of metformin. Early studies in healthy volunteers suggested that the OCT2 substrate and inhibitor cimetidine would inhibit OCT2-mediated uptake of metformin into tubular cells and thus reduce renal clearance (Yin and Wang 2016). Similarly, the antiviral dolutegravir, when co-administered with metformin, increased metformin AUC by 2.5-fold.

Based on in vitro studies where dolutegravir was identified as a weak OCT2 inhibitor, its effect is probably only partially due to OCT2 inhibition (Yin and Wang 2016). The anticancer drug cisplatin leads to dose-limiting nephrotoxicity because the kidney accumulates this drug to a higher degree than other organs. In vitro studies demonstrated that cisplatin is a good substrate of OCT2. In addition, cancer patients with the polymorphism c.808G>T (S270A) which results in lower OCT2-mediated uptake had reduced cisplatin-induced nephrotoxicity (Yin and Wang 2016). Thus, selective OCT2 inhibitors might be useful in protecting cancer patients from cisplatin toxicity.

Several polymorphisms have been identified in the *SLC22A2* gene with four of the non-synonymous variants showing ethnic-specific allele frequencies of more than 1%. These polymorphisms are c.495G>A (M165I), c.808G>T (A270S), c.1198C>T (R400C), and c.1294A>C (K432Q) (Fujita et al. 2006). When these four variants where tested in vitro using the *Xenopus laevis* oocyte expression system, all four variants were able to transport MPP^+, but differences in kinetics and in inhibition studies were observed, suggesting substrate-dependent effects. Such substrate-dependent inhibition was further characterized by comparing IC_{50} values obtained for the inhibition of the fluorescent substrate, *N*,*N*,*N*-trimethyl-2-[methyl(7-nitrobenzo[c][l,2,5] oxadiazol-4-yl)amino]ethanaminium iodide (NBD-MTMA), of MPP^+ and of metformin (Belzer et al. 2013). The results demonstrated that inhibition of OCT2-mediated metformin transport was about ten times more effective than for OCT2-mediated MPP^+ uptake. Thus, when drug-drug interactions at OCT2 are investigated, it might be more predictive if more than one transport substrate is used and ideally the most likely co-medication is included.

The regulatory agencies recommend that investigational drugs with significant renal secretion (active secretion of parent drug by the kidney is ≥25%) should be tested as substrates for OCT2. If the in vitro results show that the investigational drug is an OCT2 substrate (transport of at least twofold above the negative control) and an OCT2 inhibitor, i.e., inhibits OCT2-mediated transport of the investigational drug by ≥50% at a concentration of at least ten times the K_i value, in vivo studies might be necessary (Giacomini et al. 2010; Hillgren et al. 2013).

The Organic Anion Transporters (OATs) of the SLC22 Family

Organic Anion Transporter 1 (OAT1) and 3 (OAT3)

The organic anion transporter 1 (OAT1, gene symbol *SLC22A6*) is a glycoprotein of 550 amino acids and is mainly expressed at the basolateral membrane of the proximal tubule cells (Fig. 1d) (Koepsell 2013). It plays an important role in the renal secretion of numerous organic anions. OAT1 seems to work as an organic anion/α-ketoglutarate exchanger, exchanging the intracellular α-ketoglutarate for mainly hydrophilic and small (less than 300 kDa) organic anion. Besides various endogenous compounds including different dicarboxylates and prostaglandins, OAT1 mediates the basolateral uptake of numerous drugs such as angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, diuretics, antibiotics, antivirals, histamine H2-receptor antagonists, and NSAIDs (Burckhardt and Burckhardt 2011). In in vitro assays, the model organic anion p-aminohippurate (PAH) is frequently used to characterize OAT1. The prototypical inhibitor of OAT1 is probenecid, although it is not an OAT1-selective inhibitor. Besides probenecid several NSAIDs inhibit OAT1-mediated transport (Burckhardt and Burckhardt 2011).

The organic anion transporter 3 (OAT3, gene symbol *SLC22A8*) is a 542-amino acid, and similar to OAT1, it is mainly expressed at the basolateral membrane of the proximal tubule cells (Fig. 1d) (Koepsell 2013). Together with OAT1 it is involved in the renal secretion of various endogenous and exogenous organic anions. Like OAT1, OAT3 seems to work as an organic anion/α-ketoglutarate exchanger. Thus, both OATs are expressed in the same basolateral membrane and use the same mode of transport.

Similarly, the substrate specificities of the two transporters overlap but are not identical. While OAT1 transports mainly small and hydrophilic organic anions, OAT3 substrates are in general larger and more hydrophobic. Endogenous OAT3 substrates include bile acids like cholate and taurocholate, sulfated hormones like dehydroepiandrosterone sulfate and estrone-3-sulfate, prostaglandins, and urate. In addition, OAT3 transports numerous drugs including angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, diuretics, antibiotics, histamine H2-receptor antagonists, and several statins (Burckhardt and Burckhardt 2011). In in vitro assays, the model organic anion used to characterize OAT3 function is estrone-3-sulfate. Probenecid is also an inhibitor of OAT3. In addition, NSAIDs, cimetidine, bumetanide, and some dicarboxylates were reported to inhibit OAT3.

Because both, OAT1 and OAT3, handle a broad range of substrates, several drug-drug interactions have been described. The organic anion transport inhibitor probenecid inhibits both OAT1 and OAT3. Probenecid is a uricosuric drug which inhibits the reabsorption of uric acid in the tubules and is used to treat hyperuricemia associated with gout. Probenecid can also be used to prolong penicillin serum levels because both OAT1 and OAT3 can transport penicillin and are involved in the secretion of this β-lactam antibiotic. However, based on in vitro data, it seems that the major effect is due to inhibition of OAT3, while the effect on cephalosporins is probably due to inhibition of both OATs (Burckhardt and Burckhardt 2011). This drug-drug interaction actually is used as a beneficial side effect of the drug and not an adverse effect. Similarly, probenecid is used to prevent nephrotoxicity of cidofovir, an antiviral drug used to treat cytomegalovirus-induced eye infections in people with AIDS. The underlying mechanism is inhibition of OAT1-mediated accumulation of cidofovir in the proximal tubule cells (Yin and Wang 2016).

In the *SLC22A6* gene encoding OAT1 and the *SLC22A8* gene encoding OAT3, only very few amino acid changing polymorphisms have been identified. For OAT1, the polymorphism c.149G>A (R50H) resulted in increased affinities of the mutated protein for adefovir, cidofovir, and tenofovir. This might lead to increased nephrotoxicity for patients carrying this polymorphism (Burckhardt and Burckhardt 2011). For OAT3, the polymorphism c.913A>T (I305F) which is found at a frequency of 3.5% in Asian-Americans showed reduced estrone-3-sulfate transport compared to the wild-type OAT3 (Burckhardt and Burckhardt 2011). In a recent report, it was shown that Asians with this variant had reduced renal secretion and clearance of the cephalosporin antibiotic cefotaxime (Yee et al. 2013). All of the other identified reduced-function variants are found at less than 3% allele frequency. Thus, the R50H mutation in OAT1 and the I305F mutation in OAT3 could potentially impact renal drug elimination and increase drug concentrations in subjects carrying these mutations.

Regarding regulatory agencies, OAT1 and OAT3 are treated similar as OCT2 because they also play a role in renal secretion of numerous drugs. Thus, the recommendations described above for OCT2 are the same for OAT1 and OAT3 (Giacomini et al. 2010; Hillgren et al. 2013).

The Multidrug and Toxin Extrusion (MATE) Proteins of the SLC47 Family

Multidrug and Toxin Extrusion 1 (MATE1) and 2 (MATE2)

The SLC47 family contains two genes, *SLC47A1* and *SLC47A2* encoding multidrug and toxin extrusion 1 (MATE1) and MATE2 as well as the splice variant MATE2-K. MATE1 is composed of 570 amino acids, MATE2 has 602 amino acids, and the splice variant MATE2-K which is missing part of exon 7 has 566 amino acids. MATE1 is expressed at the apical membrane of the proximal and distal tubule and at the canalicular membrane of hepatocytes. MATE2-K is also expressed in the apical membrane of the proximal tubule, while MATE2 mRNA was not detected in the kidney. Both MATE1 and MATE2-K are involved in the tubular secretion of organic cations that were transported into tubular cells by OCT2. A potentially similar mechanism is

proposed for hepatocytes where the basolateral OCT1 takes up cationic substrates into the cells and the canalicular MATE1 secretes them into bile. The transport mechanism for MATE-mediated secretion is H^+-coupled electroneutral organic cation exchange (Motohashi and Inui 2013). Substrates transported by MATE1/2-K include well-known OCT substrates including TEA, MPP^+, cimetidine, metformin, guanidine, procainamide, quinine, topotecan, cisplatin, oxaliplatin, but also some anionic compounds like estrone-3-sulfate, acyclovir, and ganciclovir. While some of these compounds including TEA, MPP^+, and quinidine had similar affinities for both, MATE1 and MATE2-K, affinities for choline and cimetidine were very different for the two transporters, suggesting that they are multispecific transporters with overlapping but also distinct substrate specificities.

Initially drug interactions with metformin, such as the above-described cimetidine-metformin interaction, were explained as cimetidine inhibition of OCT2-mediated metformin uptake across the basolateral membrane. However, in vitro studies revealed that cimetidine is a much stronger inhibitor of MATE1/2-K than of OCT2 with a K_i value more than 20-fold lower, suggesting that the main inhibitory effect is at the secretory step mediated by MATE1/2-K (Motohashi and Inui 2013). Pyrimethamine, an antiparasitic compound, is a selective inhibitor of MATE1/2-K. When co-administered with metformin, increased metformin AUC and decreased renal clearance of metformin were reported, demonstrating that inhibition of MAT1/2-K could lead to clinically relevant drug-drug interactions (Yin and Wang 2016).

The genetic variant c.-66T>C in the 5′UTR and the intronic c.922-158G>A in the *SLC47A1* gene are associated with a higher glucose-lowering effect of metformin, suggesting that they result in lower expression levels MATE1. The variant c.-130G>A in the 5′UTR of the *SLC47A2* gene has been associated with a decreased glucose-lowering effect of metformin (Staud et al. 2013). So far no non-synonymous polymorphisms in MATE1/2-K have been reported.

Because of observed drug-drug interactions and the potential for adverse effects of new drugs, the regulatory agencies recommend the same procedures for MATE1/2-K as for OCT2 and OAT1/OAT3 described above (Giacomini et al. 2010; Hillgren et al. 2013).

The Organic Anion Transporting Polypeptides (OATPs) of the SLCO Family

Organic Anion Transporting Polypeptide 1B1 (OATP1B1) and 1B3 (OATP1B3)

The organic anion transporting polypeptide 1B1 (OATP1B1, gene symbol *SLCO1B1*) is a glycoprotein of 691 amino acids with a molecular weight of about 85 kDa. It is exclusively and evenly expressed at the basolateral (or sinusoidal) membrane of human hepatocytes throughout the liver lobule. OATP1B3 (*SLCO1B3*) is also a glycoprotein with 702 amino acids and a molecular weight of about 120 kDa. In the liver, OATP1B3 is expressed at the basolateral membrane mainly around the central vein with less expression toward the portal vein (Hagenbuch and Stieger 2013). Besides the liver, OATP1B3 has also been documented in several cancers, but it seems that outside the liver the cancer-type OATP1B3 is mainly expressed, which is missing the N-terminal 28 amino acids and is hardly expressed at the plasma membrane. As a consequence its transport activity is strongly reduced (Chun et al. 2017).

OATP1B1 and OATP1B3 have a broad and partially overlapping substrate specificity. They transport various endogenous compounds including bile acids, bilirubin and its conjugates, thyroid hormones, and several steroid conjugates (Hagenbuch and Stieger 2013). Besides these endogenous substrates, both OATPs also transport numerous drugs including statins, antihypertensives, antibiotics, and anticancer agents (Roth et al. 2012). Given that the two proteins share 80% amino acid identity, it is not astonishing that they share most of the substrates. However, there are some substrates that are specifically transported only or mainly by one of the

two liver OATPs. Estrone-3-sulfate at low concentrations (nanomolar range) is preferentially transported by OATP1B1, but at higher concentrations the low affinity high capacity OATP1B3 can take over. For OATP1B3 at least two selective substrates have been documented: cholecystokinin-8 (CCK-8) and telmisartan (Roth et al. 2012). In addition, fluorescein-containing substrates like fluorescein-methotrexate or 8-fluorescein-cAMP are in general better substrates of OATP1B3 (Bednarczyk 2010; Gui et al. 2010).

There are a number of clinically relevant drug-drug interactions that involve the liver-specific OATPs. Because polymorphisms that result in an inactive or less active transporter mainly in the *SLCO1B1* gene have been linked to altered drug disposition, it is in general assumed that OATP1B1 plays the more important role than OATP1B3 for the disposition of most drug substrates. In 2008, the SEARCH collaborative group was able to link the *SLCO1B1*variant c.521T>C (V174A, also known as OATP1B1*5) to an increased risk of statin-induced myopathy (Link et al. 2008). Several studies demonstrated that the immunosuppressant cyclosporine A, a known inhibitor of several OATPs, affected drug bioavailability in transplant patients that were treated with statins, repaglinide or bosentan (Patel et al. 2016). In several studies, statin AUC was increased in the presence of cyclosporine A between 3.5 and almost tenfold. Similar increases in AUCs for statins were also observed in patients treated with gemfibrozil, rifampicin, and HIV protease inhibitors. One of the common properties of all these drugs is that they are known inhibitors of OATP1B1 and OATP1B3. Because increased plasma levels of statins have been associated with rhabdomyolysis, any drug-drug interactions that can increase statin plasma concentrations have to be carefully monitored (Patel et al. 2016).

As indicated above, polymorphisms that lead to an inactive or less active transporter have been identified in the *SLCO1B1* gene. Reports regarding the variant c.388A>G (N130D, OATP1B1*1b) are conflicting with increased, decreased, or unchanged effects, also depending on the drug substrate (Gong and Kim 2013). When assessed in vitro, OATP1B1*5 expression at the plasma membrane is reduced to about 35% of wild-type OATP1B1. Consistent with this reduced expression level, the majority of the studies reported an increase in AUC for patients with this polymorphism. Similarly, OATP1B1*15 which consists of c.388A>G plus c.521T>C is associated with increased AUCs for several drugs including pravastatin, pitavastatin, and rosuvastatin as well as some antihypertensives, anticancer drugs, and the cholesterol-lowering ezetimibe (Gong and Kim 2013). For OATP1B3, only a few polymorphisms have been identified. The most frequent variants are c.334T>G (S112A) and c.699G>A (M233I), both showing similar expression levels in in vitro experiments. OATP1B3-M233I was associated with reduce uptake of CCK-8 and rosuvastatin (Gong and Kim 2013). The AUC of mycophenolic acid glucuronide in renal transplant patients was increased in patients with the c.334T>G and c.699G>A haplotype although in addition conflicting results were reported. Given that only a few studies are available that investigated OATP1B3 polymorphisms, additional studies are required to see whether, e.g., OATP1B1 function could compensate for decreased OATP1B3 activity. Patients with Rotor syndrome, a disorder with conjugated hyperbilirubinemia and coproporphyrinuria, were compared to their normal family members. In addition, studies in OATP1A/OATP1B knockout mice were performed. These studies revealed that only a combination of a defective OATP1B1 with a defective OATP1B3 lead to the disease, demonstrating that either a functional OATP1B1 or functional OATP1B3 was sufficient for normal bilirubin disposition (van de Steeg et al. 2012).

With regard to OATP1B1/OATP1B3 inhibitors, several compounds have been identified that inhibit uptake mediated by both transporters. The most frequently used inhibitors are rifampicin, cyclosporine A, rifamycin SV, bromosulfophthalein, and MK-571. However, many of these inhibitors also interact with other drug transporters. For example, MK-571 was originally established as an MRP2 inhibitor, but later experiments revealed that it also inhibited uptake mediated by both, OATP1B1 and OATP1B3 (Brouwer et al. 2013). Estropipate (also known as piperazine

estrone sulfate) is a somewhat selective inhibitor for OATP1B1 with an IC_{50} of 0.06 μM as compared to 19.3 μM for OATP1B3. In contrast, ursolic acid was identified as a somewhat selective OATP1B3 inhibitor with an IC_{50} of 2.3 μM as compared to 12.5 μM for OATP1B1 (Gui et al. 2010). However, given that both OATPs have shown substrate-dependent inhibition patterns, more detailed research is needed to elucidate the underlying mechanisms and to hopefully identify really selective OATP1B1 and OATP1B3 inhibitors. Such selective inhibitors might be useful as co-medication with new chemical entities that are good drugs but have a low bioavailability due to extensive liver first-pass metabolism.

OATP1B1 and OATP1B3 are key uptake transporters for numerous drugs and are expressed at the basolateral membrane of human hepatocytes. The regulatory agencies expect that investigational drugs are tested in vitro to examine whether they are substrates of OATP1B1 and/or OATP1B3 in cases where ADME studies indicate that the hepatic uptake or elimination of the investigational drug is significant (≥25% of total drug clearance) or if the hepatic uptake is clinically important (for biotransformation or if the drug target is in the liver). If the in vitro results show that the investigational drug is an OATP1B1 or OATP1B3 substrate (transport of at least twofold above the negative control), and a known OATP inhibitor such as rifampicin can decrease OATP1B1 and/or OATP1B3-mediated uptake by more than 50% at a concentration of at least ten times the K_i value, the investigational drug is considered an OATP1B1 or OATP1B3 substrate. If the investigational drug is a substrate, in vivo studies might be necessary (Giacomini et al. 2010; Hillgren et al. 2013).

Summary and Outlook

So far seven transporters, which are known to be involved in adverse drug-drug interactions due to inhibition of the transporters or due to polymorphisms, are required by the regulatory agencies to be tested for investigational drugs. Five additional transporters either are already required by certain agencies or are considered by some or all of the three major agencies. However, because some of the drug-drug interactions involve multiple transporters, and because there are at least another 13 transporters that are known to be involved in drug disposition, the list of required transporters can increase, and it is even possible that certain transporters that are currently not considered will be required to be tested in the future. The most likely candidates are listed in Table 1, but several other transporters that are drug targets or mediate the transport of only a few drugs with so far no reported significant adverse effects could become relevant in the future.

Acknowledgments The author would like to acknowledge the National Institutes of Health grant GM077336.

References and Further Reading

Abdullahi W, Davis TP, Ronaldson PT (2017) Functional expression of P-glycoprotein and organic anion transporting polypeptides at the blood-brain barrier: understanding transport mechanisms for improved CNS drug delivery? AAPS J 19:931–939

Amidon GL, Lennernas H, Shah VP et al (1995) A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm Res 12:413–420

Bednarczyk D (2010) Fluorescence-based assays for the assessment of drug interaction with the human transporters OATP1B1 and OATP1B3. Anal Biochem 405:50–58

Belzer M, Morales M, Jagadish B et al (2013) Substrate-dependent ligand inhibition of the human organic cation transporter OCT2. J Pharmacol Exp Ther 346:300–310

Brouwer KL, Keppler D, Hoffmaster KA et al (2013) In vitro methods to support transporter evaluation in drug discovery and development. Clin Pharmacol Ther 94:95–112

Burckhardt G, Burckhardt BC (2011) In vitro and in vivo evidence of the importance of organic anion transporters (OATs) in drug therapy. Handb Exp Pharmacol 201:29–104

Chun SE, Thakkar N, Oh Y et al (2017) The N-terminal region of organic anion transporting polypeptide 1B3 (OATP1B3) plays an essential role in regulating its plasma membrane trafficking. Biochem Pharmacol 131:98–105

Cleophas MC, Joosten LA, Stamp LK et al (2017) ABCG2 polymorphisms in gout: insights into disease susceptibility and treatment approaches. Pharmgenomics Pers Med 10:129–142

Droge C, Bonus M, Baumann U et al (2017) Sequencing of FIC1, BSEP and MDR3 in a large cohort of patients with cholestasis revealed a high number of different genetic variants. J Hepatol 67:1253–1264

Drozdzik M, Groer C, Penski J et al (2014) Protein abundance of clinically relevant multidrug transporters along the entire length of the human intestine. Mol Pharm 11:3547–3555

Fujita T, Urban TJ, Leabman MK et al (2006) Transport of drugs in the kidney by the human organic cation transporter, OCT2 and its genetic variants. J Pharm Sci 95:25–36

Giacomini KM, Huang SM, Tweedie DJ et al (2010) Membrane transporters in drug development. Nat Rev Drug Discov 9:215–236

Gong IY, Kim RB (2013) Impact of genetic variation in OATP transporters to drug disposition and response. Drug Metab Pharmacokinet 28:4–18

Gui C, Obaidat A, Chaguturu R et al (2010) Development of a cell-based high-throughput assay to screen for inhibitors of organic anion transporting polypeptides 1B1 and 1B3. Curr Chem Genomics 4:1–8

Hagenbuch B, Stieger B (2013) The SLCO (former SLC21) superfamily of transporters. Mol Asp Med 34:396–412

Heredi-Szabo K, Glavinas H, Kis E et al (2009) Multidrug resistance protein 2-mediated estradiol-17beta-D-glucuronide transport potentiation: in vitro-in vivo correlation and species specificity. Drug Metab Dispos 37:794–801

Hillgren KM, Keppler D, Zur AA et al (2013) Emerging transporters of clinical importance: an update from the International Transporter Consortium. Clin Pharmacol Ther 94:52–63

Hira D, Terada T (2018) BCRP/ABCG2 and high-alert medications: biochemical, pharmacokinetic, pharmacogenetic, and clinical implications. Biochem Pharmacol 147:201–210

Juliano RL, Ling V (1976) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. Biochim Biophys Acta 455:152–162

Koepsell H (2013) The SLC22 family with transporters of organic cations, anions and zwitterions. Mol Asp Med 34:413–435

Lee CA, O'Connor MA, Ritchie TK et al (2015) Breast cancer resistance protein (ABCG2) in clinical pharmacokinetics and drug interactions: practical recommendations for clinical victim and perpetrator drug-drug interaction study design. Drug Metab Dispos 43:490–509

Link E, Parish S, Armitage J et al (2008) SLCO1B1 variants and statin-induced myopathy – a genomewide study. N Engl J Med 359:789–799

Liu Y, Zheng X, Yu Q et al (2016) Epigenetic activation of the drug transporter OCT2 sensitizes renal cell carcinoma to oxaliplatin. Sci Transl Med 8:348ra397

Lund M, Petersen TS, Dalhoff KP (2017) Clinical implications of P-glycoprotein modulation in drug-drug interactions. Drugs 77:859–883

Morgan RE, Trauner M, van Staden CJ et al (2010) Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. Toxicol Sci 118:485–500

Motohashi H, Inui K (2013) Multidrug and toxin extrusion family SLC47: physiological, pharmacokinetic and toxicokinetic importance of MATE1 and MATE2-K. Mol Asp Med 34:661–668

Nies AT, Koepsell H, Damme K et al (2011) Organic cation transporters (OCTs, MATEs), in vitro and in vivo evidence for the importance in drug therapy. Handb Exp Pharmacol 201:105–167

Patel M, Taskar KS, Zamek-Gliszczynski MJ (2016) Importance of hepatic transporters in clinical disposition of drugs and their metabolites. J Clin Pharmacol 56(Suppl 7):S23–S39

Pfeifer ND, Hardwick RN, Brouwer KL (2014) Role of hepatic efflux transporters in regulating systemic and hepatocyte exposure to xenobiotics. Annu Rev Pharmacol Toxicol 54:509–535

Roninson IB, Chin JE, Choi KG et al (1986) Isolation of human mdr DNA sequences amplified in multidrug-resistant KB carcinoma cells. Proc Natl Acad Sci U S A 83:4538–4542

Roth M, Obaidat A, Hagenbuch B (2012) OATPs, OATs and OCTs: the organic anion and cation transporters of the SLCO and SLC22A gene superfamilies. Br J Pharmacol 165:1260–1287

Saidijam M, Karimi Dermani F, Sohrabi S et al (2018) Efflux proteins at the blood-brain barrier: review and bioinformatics analysis. Xenobiotica 48: 506–532

Schinkel AH, Jonker JW (2003) Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview. Adv Drug Deliv Rev 55:3–29

Schinkel AH, Wagenaar E, Mol CA et al (1996) P-glycoprotein in the blood-brain barrier of mice influences the brain penetration and pharmacological activity of many drugs. J Clin Invest 97:2517–2524

Sjostedt N, Holvikari K, Tammela P et al (2017) Inhibition of breast cancer resistance protein and multidrug resistance associated protein 2 by natural compounds and their derivatives. Mol Pharm 14:135–146

Staud F, Cerveny L, Ahmadimoghaddam D et al (2013) Multidrug and toxin extrusion proteins (MATE/SLC47); role in pharmacokinetics. Int J Biochem Cell Biol 45:2007–2011

Stieger B (2011) The role of the sodium-taurocholate cotransporting polypeptide (NTCP) and of the bile salt export pump (BSEP) in physiology and pathophysiology of bile formation. Handb Exp Pharmacol 201:205–259

Stieger B, Hagenbuch B (2016) Recent advances in understanding hepatic drug transport. F1000Res 5:2465

Terada T, Hira D (2015) Intestinal and hepatic drug transporters: pharmacokinetic, pathophysiological, and pharmacogenetic roles. J Gastroenterol 50:508–519

Ueda K, Cornwell MM, Gottesman MM et al (1986) The mdr1 gene, responsible for multidrug-resistance, codes

for P-glycoprotein. Biochem Biophys Res Commun 141:956–962

Urquhart BL, Kim RB (2009) Blood-brain barrier transporters and response to CNS-active drugs. Eur J Clin Pharmacol 65:1063–1070

van de Steeg E, Stranecky V, Hartmannova H et al (2012) Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver. J Clin Invest 122:519–528

Wagner DJ, Hu T, Wang J (2016) Polyspecific organic cation transporters and their impact on drug intracellular levels and pharmacodynamics. Pharmacol Res 111:237–246

Wessler JD, Grip LT, Mendell J et al (2013) The P-glycoprotein transport system and cardiovascular drugs. J Am Coll Cardiol 61:2495–2502

Wolking S, Schaeffeler E, Lerche H et al (2015) Impact of genetic polymorphisms of ABCB1 (MDR1, P-glycoprotein) on drug disposition and potential clinical implications: update of the literature. Clin Pharmacokinet 54:709–735

Yee SW, Nguyen AN, Brown C et al (2013) Reduced renal clearance of cefotaxime in Asians with a low-frequency polymorphism of OAT3 (SLC22A8). J Pharm Sci 102:3451–3457

Yin J, Wang J (2016) Renal drug transporters and their significance in drug-drug interactions. Acta Pharm Sin B 6:363–373

Role of Clinical Pharmacokinetics Studies in Contemporary Oncology Drug Development

55

Fatih M. Uckun and Sanjive Qazi

Contents

Abstract

Pharmacokinetics (PK) studies enable drug developers to elucidate the relationship of dose to blood concentrations of drugs in various patient populations and determine the need for dose adjustment based on PK differences among demographic subgroups or subgroups with impaired elimination. PK studies also provide the basis for therapeutic drug monitoring in rare patient populations or when effective drugs with very narrow safe therapeutic windows must be used. Population PK studies are aimed at optimizing the dose and schedule by identifying the factors that alter the dose-concentration relationship and determining if such alterations change the therapeutic index using a data-driven approach and integrated sources of information. The clinical importance of identifying and implementing optimum dosing strategies has led to increased application of the population PK strategies in early oncology clinical trials. Multi-scale mechanistic PK models have been developed in an attempt to

F. M. Uckun (✉)
AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA
e-mail: fatih.uckun@att.net; fatih.uckun@aresmit.com

S. Qazi
AresMIT Biomedical Computational Strategies (ABCS), Minneapolis, MN, USA

Ares Pharmaceuticals, St. Paul, MN, USA

Bioinformatics Program, Gustavus Adolphus College, St. Peter, MN, USA
e-mail: sqazi@gustavus.edu

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_24

better predict the clinical performance of the oncology drug candidates. Over the last two decades PK studies have increasingly become an integral part of early clinical development of promising oncology drugs entering the clinical space. Of the total of 4,481 interventional clinical oncology trials with integrated PK studies registered in the clinicaltrials.gov data repository that were initiated within the 24-year time interval between 1994 and 2018, ~60% of the clinical PK studies were initiated within the last 8 years.

Introduction

Pharmacokinetics (PK) is the study of the drug concentrations in the body during a period of time, and it includes the processes by which the drug is absorbed, distributed, metabolized, and excreted (ADME). An ideal drug should have high absolute bioavailability with low variability and exhibit linear PK over therapeutic dose range without significant modulation of the PK by concomitant food or pH-altering medications. An ideal drug should also reach the target site(s) of action promptly at effective/nontoxic concentrations, should not accumulate in nontarget organs, and should not have a narrow therapeutic index. Furthermore, it should not be extensively metabolized by a liver enzyme so that its clearance would not be significantly affected by hepatic dysfunction or by concomitant use of other drugs that affect one or more metabolizing enzymes. However, the PK profiles of most drugs are influenced by their physicochemical properties, product/formulation, administration route, patient's intrinsic and extrinsic factors (e.g., organ dysfunction, diseases, concomitant medications, food). PK studies enable drug developers to elucidate the relationship of dose to blood concentrations of drugs in various patient populations and determining the need for dose adjustment based on PK differences among demographic subgroups or subgroups with impaired elimination (e.g., hepatic or renal disease). Defining the optimum dosing strategy for a population, subgroup, or individual patient requires resolution of the interindividual, kinetic, as well as random variability (Turner et al. 2015; Undevia et al. 2005). Population PK studies are aimed at optimizing the dose and schedule by identifying the factors that alter the dose-concentration relationship and determining if such alterations change the therapeutic index using a data-driven approach and integrated sources of information, as detailed in a 1999 FDA Guidance for Industry that was prepared by the Population Pharmacokinetic Working Group of the Clinical Pharmacology Section of the Medical Policy Coordinating Committee in the Center for Drug Evaluation and Research (CDER) in cooperation with the Center for Biologics Evaluation and Research (CBER) at the FDA (1999). In 2003, the FDA Exposure-Response Working Group under the Medical Policy Coordinating Committee, Center for Drug Evaluation and Research (CDER), in cooperation with the Center for Biologics Evaluation and Research (CBER) at the FDA issued another guidance for industry regarding study design, data analysis, and regulatory applications related to exposure-response relationships (FDA 2003a; Overgaard et al. 2015).

The clinical importance of identifying and implementing optimum dosing strategies has led to increased application of the population PK strategies in early oncology clinical trials. Population PK studies provide actionable safety, efficacy, and dosage optimization information for the drug label because of their early integration with clinical oncology trials. The purpose of this chapter is to review and discuss the increasing role of PK studies in the oncology drug development process.

Contribution of Pharmacokinetics to Clinical Development of Oncology Drugs

Over the last two decades, PK studies have increasingly become an integral part of early clinical development of promising oncology drugs entering the clinical space (Chen et al. 1999; Uckun et al. 1995, 2013, 2015; Ursino et al. 2017; Waller et al. 2018; Wicki et al. 2018). As

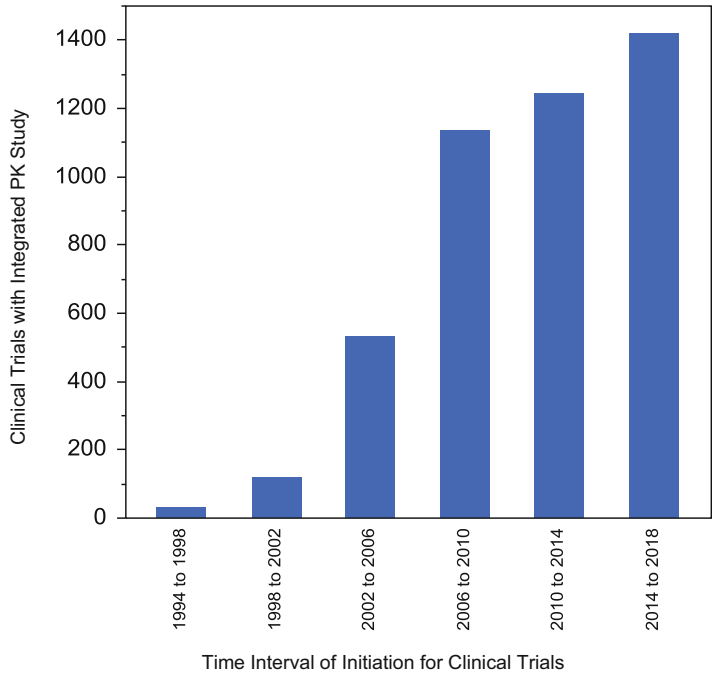


Fig. 1 Clinical trials with integrated PK studies in oncology. We interrogated the clinicaltrials.gov data repository (https://clinicaltrials.gov/) to determine the number of interventional clinical trials that employed PK studies from 1994 to 2018 in 4 year increments. All Interventional trials that were started over the 4-year period were included in the totals. There were a total of 4481 trials counted from 1994 to 2018. A pronounced increase in the use of PK studies was observed after the year 2002. Search terms to identify the trials were "pharmacokinetics," "interventional studies," and "cancer"

shown in Fig. 1, a pronounced and continued increase in the use of integrated PK analyses was observed in clinical trials that started after the year of 1998. Of the total of 4481 interventional clinical oncology trials with integrated PK studies registered in the clinicaltrials.gov data repository that were initiated within the 24-year time interval between August 1994 and July 2018, only 30 (0.7%) were started between August 1994 and July 1998, 121 (2.7%; 3.8-fold increase from previous 4 years) between August 1998 and July 2002, 532 (11.9%; 4.4-fold increase from previous 4 years) between August 2002 and July 2006, 1135 (25.3%; 2.1-fold increase from previous 4-years) between August 2006 and July 2010, 1244 (27.8%; 9.6% increase from previous 4 years) between August 2010 and July 2014, and 1420 (31.7%; 14.1% increase from previous 4 years) between August 2014 and July 2018. Notably, ~60% of the clinical PK studies were initiated within the last 8 years. Hence, PK studies are playing an increasingly important role in the clinical development path of oncology drugs.

PK studies often combined with integrated pharmacodynamics (PD) components play a pivotal role in clinical comparisons of different formulations, prodrugs and dosing schedules aimed at identifying the best way of using a promising new drug at a nontoxic dose level. For example, DTS-201 is a doxorubicin (Dox) prodrug that shows encouraging data in experimental models in terms of both efficacy and safety compared with conventional Dox. Notably, a high equivalent dose of Dox could be delivered without severe drug-related cardiac events. DTS-201 was administered at four dose levels ranging from 80 to 400 mg/m^2, which is equivalent to 45–225 mg/m^2 of conventional Dox (Schöffski et al. 2017). The recommended phase II dose (RP2D) was 400 mg/m^2.

PK modeling and model-informed precision dosing (MIPD) have been explored as tools to optimize treatment outcomes in oncology by maximizing patient safety via overdose protection and by avoiding treatment failures caused by suboptimal drug exposures (Barbolosi et al. 2015). NCT02732275 is a first-in-human phase I study of the epigenetic modulator DS-3201b, a dual inhibitor of enhancer of zeste homolog 1 (EZH1) and EZH2 in patients with relapsed/refractory lymphomas. Recently, a population PK model was developed using the integrated PK data from the study to define the dose-exposure relationships and reported that a 2-compartment PK model with first-order elimination and absorption lag-time best characterized the plasma concentration-time profile of DS-3201a (Atsumi et al. 2017).

Physiologically based pharmacokinetic (PBPK) modeling is a potential tool which can be effectively applied throughout all phases of oncology drug development and allows a more granular prediction of tissue drug exposures (Schwenger et al. 2018; Schultze-Mosgau et al. 2018; Cheeti et al. 2013; Ferl et al. 2016; Saeheng et al. 2018; Rowland 2013; Sager et al. 2015). The number of experimental animals and human participants enrolled in the studies can be reduced using PBPK modeling and PBPK-population-PK modeling. For example, Tsukamoto et al. studied the kinetics of capecitabine and its metabolites. Their PBPK model integrated tissue-specific information about metabolic enzyme activity between tumor and normal cells from in vitro data and enabled the prediction of the therapeutic index in terms of exposure in target organs and toxicity in off-target organs (i.e., gastrointestinal tract toxicity) (Tsukamoto et al. 2001).

Besides the systemic exposure levels of the parent compound and/or its metabolites, several baseline characteristics, including but not limited to age, gender, and race as well as comorbidities of the host also affect the risk of severe side effects and tolerability as well as efficacy of drugs at optimized dose levels (Owonikoko et al. 2018). Therefore, it is very important to identify biomarkers that (i) allow the rational assignment of individual patients to those treatments that are most likely to benefit them and ensure maximized patient safety as well as best survival outcome and likewise and (ii) enroll into a particular study a biomarker-enriched population that is most likely to benefit from the treatment program (Jamal et al. 2017). For example, inhibition of Janus-kinase 1/2 (JAK1/2) is an innovative strategy to treat myeloproliferative neoplasms, but recently this exciting new treatment approach has been shown to be associated with a 15-fold higher risk of development of aggressive B-cell lymphomas. Lymphomas occurring during JAK1/2 inhibitor treatment were preceded by a preexisting B-cell clone in all patients tested. Therefore, detection of a preexisting B-cell clone may identify individuals at risk (Porpaczy et al. 2018).

Unlike small molecules which bind to their molecular targets without significantly affecting the systemic exposure levels, biotherapeutic agents, such as monoclonal antibodies (e.g., the anti-PD1 monoclonal antibody pembrolizumab), bind to their targets with much higher affinity and display a nonlinear "target-mediated drug disposition" (TMDD). The disposition of the drug-target molecular complexes can influence the systemic exposure levels (Ahamadi et al. 2017; Moreau et al. 2012). In addition, the lack of a relationship of pembrolizumab PK and overall survival (OS) in patients with advanced melanoma and non-small cell carcinoma (NSCLC) demonstrates the challenges in determining the RP2D and optimal dosing for monoclonal antibodies and immune-oncology drugs (Turner et al. 2018; Freshwater et al. 2017; Chatterjee et al. 2016; Turan et al. 2018). It is also important to take into consideration the circadian fluctuations of the ADME of oncology drugs (Vérennoneau-Veilleux and Bélair 2017).

Importantly, PK studies provide the basis for therapeutic drug monitoring in rare patient populations or when effective drugs with very narrow safe therapeutic windows must be used (Thomas et al. 2018a). Therapeutic drug monitoring is particularly important for optimized clinical use for certain therapeutics, such as oral antihormonal drugs are essential in the treatment of breast and prostate cancer, that display a high interpatient PK variability, when the treatments

employ fixed doses, which has the associated risks of underdosing as well as overdosing (Groenland et al. 2018; Paci et al. 2014).

When there is compelling evidence from nonclinical studies for an association between systemic exposure levels of a drug or its metabolite and the desired treatment outcomes, PK-guided dose escalation studies utilizing real-time PK measurements to determine the dose cohorts based on the systemic exposure levels could provide the opportunity to determine the maximum tolerated systemic exposure (MTSE) levels and how they compare to the systemic exposure levels proven effective in nonclinical studies.

The Impact of Hepatic and/or Renal Impairment on PK of Oncology Drugs and Patient-Tailored Dosing Schedules

The Cancer Therapy Evaluation Program (CTEP) at the NCI prioritized study of special patient populations with hepatic dysfunction phase I clinical trials (HDCT) to determine safe administration parameters of oncology drugs for subjects with varying degrees of liver dysfunction. HDCT sponsored by CTEP and others have provided clinically useful information on the optimal dosing of oncology drugs in subjects with different degrees of liver test abnormalities that have provided administration guidance in the labels for patients with abnormal organ function. Hepatic dysfunction phase I clinical trials (HDCT) provide safe administration parameters of oncology drugs for subjects with varying degrees of liver dysfunction (Mansfield et al. 2016).

The elimination of several oncology drugs, such as the proteasome inhibitor bortezomib, occurs through metabolism by liver enzymes (Tan et al. 2018). The change in liver function may potentially change the inhibitory and/or inducing potential of the liver metabolizing enzymes, thus the PK and PD in patients with hepatic impairment may differ from patients with normal hepatic function. As cancer patients often have alterations in their liver function due to disease-related reasons (e.g., liver metastases), hepatotoxic treatments (chemotherapy, radiation therapy, treatment with immuno-oncology drugs), and/or other comorbidities they may have, it is important to determine the effects of hepatic impairment on the PK and safety of drugs metabolized by liver enzymes and also, if possible, determine whether dose modification would be necessary in such patients. Patients in these studies are typically assigned to different groups according to their liver function as per NCI and FDA guidance (FDA 2003b). The primary objective of such studies is to evaluate the effect of hepatic impairment on the steady state PK of the respective therapeutic agents in advanced cancer patients. The secondary objectives are to evaluate the effect of hepatic impairment on the safety and antitumor activity of the respective therapeutic agent in advanced cancer patients.

Some of the ongoing studies evaluating the effects of hepatic impairment on the PK and safety of targeted therapeutics include among others NCT01767623 (A Study of The Impact of Severe Hepatic Impairment on the Pharmacokinetics and Safety of Vemurafenib – a BRAF kinase inhibitor – in BRAF V600 Mutation-Positive Cancer Participants), NCT02894385 (Effect of Hepatic and Renal Impairment on the Pharmacokinetics, Safety and Tolerability of BAY1841788/daralutamide – a nonsteroidal antiandrogen), NCT03092999 (Effect of Hepatic Impairment on the Pharmacokinetics, Safety and Tolerability of BAY1002670/Vilaprisan – a steroidal selective progesterone receptor modulator/SPRM), NCT03359850 (Pharmacokinetic and Safety Study of Niraparib – a PARP inhibitor – With Normal or Moderate Hepatic Impairment Patients), NCT03282513 (A Study of AG-120 (Ivosidenib) – an IDH1 inhibitor in Subjects With Mild or Moderate Hepatic Impairment or Normal Hepatic Function), and NCT01429337 (Pharmacokinetics and Safety of Midostaurin – FLT3 inhibitor – in Subjects With Impaired Hepatic Function and Subjects With Normal Hepatic Function).

In a recent Pfizer study (NCT01576406), the effect of hepatic impairment was evaluated on the pharmacokinetics and safety of the ALK-inhibitor crizotinib in patients with advanced cancer. No adjustment to the approved 250 mg twice daily

(BID) dose of crizotinib was recommended for patients with mild hepatic impairment. The recommended dose was 200 mg BID for patients with moderate hepatic impairment, and the dose was recommended not to exceed 250 mg daily for patients with severe hepatic impairment. Adverse events appeared consistent among the hepatic impairment groups (El-Khoueiry et al. 2018). Sonidegib is a potent, selective, and orally bioavailable inhibitor of the Hedgehog signaling pathway, primarily metabolized by the liver. Horsmans et al. assessed the PK and safety of sonidegib in subjects with varying degrees of hepatic function. Sonidegib exposures were similar or decreased in the hepatic impairment groups compared with the normal group, and sonidegib was generally well-tolerated in all subjects. Dose adjustment was not considered necessary for subjects with mild, moderate, or severe hepatic impairment (Horsmans et al. 2018). By comparison, the analysis of the impact of hepatic impairment on the PK and PD of the alkylating agent Trabectedin, that is metabolized by the liver and has been associated with liver toxicities, including including hepatic failure, revealed that Trabectedin treatment of patients with hepatic impairment results in higher plasma exposures but hepatotoxicity in patients with normal liver function can be effectively addressed through dose reductions and delays (Calvo et al. 2018).

It is generally known that renal impairment can affect not only the disposition of drugs that are cleared primarily through the kidney but also other drugs with minimal renal elimination because of the effects of kidney disease on drug-metabolizing enzymes, transporters, and drug-binding proteins. Some drugs such as Udenafil, a phosphodiesterase-5 inhibitor, used to treat erectile dysfunction, are not predominantly eliminated by the kidney but renal impairment can alter its secretion/transport pathways. Significant correlations were observed among the creatinine clearance, oral clearance, and maximum concentration of Udenafil and a dose adjustment of Udenafil would seem warranted in subjects with moderate or severe renal impairment (Cho et al. 2018). Drug PK and safety of oncology drugs must therefore be assessed in subjects with a renal impairment. PK studies in combination with model-based strategies, including population PK and physiologically based PK (PBPK) modeling, have been used to evaluate the impact of renal impairment on dose-exposure relationships and optimize dosing for patients with various degrees of renal impairment (Xiao et al. 2017; Beumer et al. 2016; Tortorici et al. 2012). The insights gained from these studies are used for dose selection/dose adjustment in patient populations with renal impairment to improve the therapeutic index of anti-cancer treatments (EMA 2015; FDA 2010).

Importance of PK in the Changing Regulatory Landscape Regarding Access of Pediatric and Young Adult Patient Populations to New Treatment Platforms

The clinical trial landscape in oncology has traditionally been associated with significant delays in the evaluation of promising new therapies in poor prognosis pediatric cancer patients who are in urgent need for therapeutic innovations (FDA 2018; Freyer et al. 2013; Bleyer et al. 2018; Burke et al. 2007; Uckun and Kenny 2018; Vassal et al. 2015; Veal et al. 2010; Beaver et al. 2017; Thomas et al. 2018b; Chuk et al. 2017; Fern and Taylor 2018). There is growing consensus among pediatric hematologists-oncologists, US Food and Drug Administration (FDA), European Medicines Agency (EMA), coalitions of subject matter experts, support groups, and other stakeholders that these delays have contributed to the unsatisfactory progress in improving the survival outcomes of adolescents with cancer (Kim et al. 2017; FDA 2018; Gaspar et al. 2018; Stark et al. 2016). Both FDA and EMA launched new regulatory initiatives aimed at improving the access of pediatric cancer patients to novel therapies developed for adults with cancer. The European Pediatric Medicine Regulation [(EC)-No1901/2006)] mandated the establishment of the EMA's Pediatric Committee to provide guidance to pharmaceutical companies regarding their Pediatric Investigation Plans (PIPs) for their drugs in pipeline (EC 2006). The multistakeholder

platform ACCELERATE (http://www.accelerate-platform.eu) presented a consensus expert opinion in support of early drug access for adolescents with cancer indicating that enrollment of adolescents of 12 years and over in adult early-phase clinical drug trials would represent a safe and efficient strategy in drug development (Gaspar et al. 2018). Several changes were proposed by ACCELERATE to facilitate that adolescents have access to early drug development programs, including that (i) there should be no set upper or lower age limit criteria for phase II and phase III trials for cancers that are present in both pediatric and adult populations with similar biology and (ii) adolescents over 12 years of age should be included from the onset of the cancer drug development process in adults (Gaspar et al. 2018). In June 2018, FDA issued a draft guidance entitled "Considerations for the Inclusion of Adolescent Patients in Adult Oncology Clinical Trials" (FDA 2018) emphasizing that pediatric oncology drug development should be coordinated with oncology drug development for adults as part of an overall drug development plan and detailing a series of recommendations regarding inclusion of pediatric patients in adult oncology trials in the USA which was based on a previous FDA publication (Chuk et al. 2017) and can be viewed as a strong endorsement of the recent ACCELERATE proposal (Gaspar et al. 2018) in Europe. The new FDA recommendations would certainly expand the options available for adolescent cancer patients who have relapsed after or are refractory to standard therapeutic strategies with no curative options, or for whom no standard therapies with curative intent exist. The draft guidance suggesting that adolescent patients may be enrolled in first-in-human clinical trials after initial adult PK and toxicity data are obtained is aimed at providing significant risk mitigation for adolescents (FDA 2018). Furthermore, the important provisions of the Race for Children Act, which is incorporated as Title V of the FDA Reauthorization Act (FDARA) that was enacted on August 18, 2017 (FD&C Act Sec. 505B (a)(3), 21 USC 355c (a)(3), Public Law 115-52), has created a mechanism to expedite the evaluation of novel medicines with the potential to address the unmet need in the pediatric population by requiring pediatric investigation of appropriate new drugs intended for adults with cancer (Reaman 2018). Specifically, Title V requires evaluation of new molecularly targeted drugs and biologics "intended for the treatment of adult cancers and directed at a molecular target substantially relevant to the growth or progression of a pediatric cancer" in molecularly targeted pediatric cancer investigation to generate clinically meaningful study data, "using appropriate formulations, regarding dosing, safety and preliminary efficacy to inform potential pediatric labeling" by designing and executing earlier rational dose finding and signal seeking trials (Reaman 2018). The Alliance for Childhood Cancer, representing more than 30 national patient advocacy groups and professional medical and scientific organizations invested in advancing the interests of children with cancer, applauded the inclusion of the RACE for Children Act in the FDA Reauthorization Act of 2017 (FDARA) legislation, passed in the Senate and in the House in July of 2017. These new regulatory initiatives by EMA and FDA combined with umbrella clinical trial initiatives aimed at allowing children and adolescents with relapsed or refractory pediatric cancers early access to promising targeted precision medicines have the potential to significantly alter the therapeutic landscape for difficult-to-treat pediatric/adolescent cancers for the benefit of current and future pediatric cancer patients (Uckun and Kenny 2018).

The FDA recommendations in the draft guidance are based in part on the observed similarities in disposition and PK of drugs in adolescents and adults (Thai et al. 2015; Fern and Taylor 2018; Smith et al. 2016; Freyer et al. 2013; FDA 2018; Gaspar et al. 2018; Fouladi et al. 2010; Forrest et al. 2018; Paoletti et al. 2013). Sometimes, the adult PK exposure can be used as target for dose finding in pediatrics. For example, the pediatric sunitinib PK data were adequately predicted from adult data with a mean prediction error of 1.80% (Janssen et al. 2017).

It should be noted, however, that the cited similarities were based on single agent studies with the inherent limitation that a careful

consideration of the PK and PD features of the major metabolites were not included in relationship to other cancer drugs that are typically used in combination trials or comorbidities of patients. The single agent trials have traditionally not included pharmacometabolomics, pharmacogenetics, and pharmacogenomics studies for the parent drug or its metabolites. Modifications of critical proteins through reactive metabolites are thought to be responsible for a number of adverse drug reactions (Krauss et al. 2012; Niu et al. 2017; Kalgutkar and Dalvie 2015; Reis-Mendes et al. 2015; Han et al. 2017; van Andel et al. 2018; Chavan et al. 2018; Sun et al. 2018). Therefore, besides the levels of drug exposure, the generation of chemically reactive metabolites also contributes to drug side effects. The metabolism of some of the anticancer drugs is highly complex due to the engagement of multiple enzymes and transporters and is therefore prone to unintended drug-drug interactions. For example, the standard anticancer drug Irinotecan serves as the prodrug for the 2–3 logs more potent topoisomerase I inhibitor SN-38 that is responsible for the dose-limiting toxicities (DLTs) associated with irinotecan. Single nucleotide polymorphisms in several drug metabolizing enzymes (e.g., uridine diphosphate glucuronosyltransferase [UGT] 1A1, UGT1A7, UGT1A9) and drug transporters (e.g., ATP-binding cassette [ABC] B1, ABCC1) are associated with irinotecan toxicity (de Man et al. 2018). Fluoroacetate is considered one of the major metabolites of 5-fluorouracil responsible for its cardiotoxicity (Reis-Mendes et al. 2015). Several therapeutic and toxic effects of cyclophosphamide are the result of the actions of its active metabolites formed by the hepatic microsomal cytochrome P450 mixed function oxidase system: The active cyclophosphamide metabolites hydroxycyclophosphamide and acrolein are shown to be more cardiotoxic than the parent drug. In human autopsy cardiac tissues of previously doxorubicin (Dox)-treated patients, the cardiac levels of the metabolite doxorubicinol were almost double of the parent compound doxorubicin (Reis-Mendes et al. 2015). Although Paclitaxel cardiotoxicity is usually low and does not seem to be related with the formation of reactive metabolites, its concomitant use with Dox results in significantly increased cardiotoxicity because the pharmacokinetic interactions between paclitaxel and DOX and also because paclitaxel stimulates the NADPH-dependent reduction of Dox into doxorubicinol. The schedule paclitaxel followed by Dox is more cardiotoxic with an incidence of 18–20% of congestive heart failure than in patients with breast cancer given Dox followed by paclitaxel at standard dose levels (Reis-Mendes et al. 2015). Ponatinib is an orally available pan-BCR-ABL tyrosine kinase inhibitor that has been approved for treatment of resistant chronic myeloid leukemia (CML) and Philadelphia chromosome-positive ALL. However, it can cause severe side effects including cardiovascular toxicity with both arterial and venous thromboembolism and severe systemic hypertension, vascular occlusions as well as pancreatitis, and liver toxicity. Although the initial work had suggested CYP3A4 as a major pathway of ponatinib disposition, Lin et al. recently reported that CYP1A1, a highly inducible enzyme that unlike many other P450s can be expressed in most tissues such as lung and lung tumors, is highly active toward this compound and metabolism by CYP1A1 results in the formation of reactive epoxides from ponatinib that likely contribute to the side effects associated with its clinical use (Lin et al. 2017). Epoxides are chemically reactive and can react covalently with both DNA and proteins to cause mutations and toxicity. CYP1A1 levels are constitutively very low but are highly inducible on activation of the aryl hydrocarbon receptor by compounds including polycyclic aromatic hydrocarbons found in cigarette smoke. Notably, hypertension or vaso-occlusive disease observed in ponatinib-treated patients has been associated with smoking. Therefore, it is of vital importance that hybrid adolescent-adult studies incorporate detailed analyses aimed at characterizing the drug-drug interaction at the level of the parent compounds as well as their metabolites both in adolescents and adult populations. Risk mitigation measures aimed at maximizing the safety of adolescent patients enrolled on the hybrid studies should take the data from such analyses into consideration. In view of these regulatory changes, we anticipate a

growing emphasis on PK studies in future clinical trials of oncology drugs with a major focus on the PK profiles of precision medicines in the pediatric/adolescent and young adult patient populations.

Multiscale Mechanistic PK Modeling Platforms

The majority of oncology drug candidates fail in early oncology clinical trials due to excessive toxicity and/or disappointing efficacy (Krauss et al. 2012; Turner et al. 2015; Gadkar et al. 2016). Multiscale mechanistic PK models have been developed in an attempt to better predict the clinical performance of the oncology drug candidates (Rousseau and Marquet 2012; Smith et al. 2017; Darwich et al. 2017; Barbolosi et al. 2015; Bizzotto et al. 2017a, b; Wilkins et al. 2017; Yankeelov et al. 2016). These models take into consideration the complexity of the host response, flux of the drug through different compartments of the host body, nonlinear treatment-emergent responses through drug-induced pertubation of a complex system and account for patient-to-patient differences in regard to drug metabolism and transportation.

Systems level consideration of drug responses in these models attempt to better characterize the hierarchical, nonlinear, dynamic responses at the network level of drug action that may affect both efficacy and toxicity in clinical settings. Systems PK also aims to explain the variations in drug uptake and metabolism by considering (i) drug-specific factors such as physiochemical properties and drug regimens, (ii) patient-specific factors that account for individual differences in the quality (e.g., affinity, catalytic activity of transporters and metabolic enzymes) and quantity of interactions (e.g., synthesis and degradation of drugs entering the body), (iii) epigenetic factors that regulate expression of transporters and metabolism of drug, (iv) tissue organ variabilities in anatomy, size of parenchymal cell numbers, and fluid volumes; and (v) environmental factors that modify the PK via food uptake and nutrition. The software platforms available as analytical tools for these models have also evolved to evaluate a large number of variables, perform clinical simulations along with databases allowing the integration of different knowledge environments. An example of a modeling platform includes "The Drug Disease Model Resources" (DDMoRe; http://www.ddmore.eu/) consortium that aims to improve the accessibility and cost effectiveness of model-informed drug discovery and development (Wilkins et al. 2017) by providing a curated model repository and an interoperability framework to integrate infrastructure for efficient exchange and integration of models across modeling languages (e.g., PFIM, Onolix, Simulx, R, NONMEM7, PsN, WinBUGS, MATLAB, SimCYP). In this contemporary modeling environment, the user is able to interact with the "Interoperability Framework" (IIF) via a graphical front-end interface (MDL-IDE; Bizzotto et al. 2017b; Smith et al. 2017) to enable editing of models written in HTML exchange formats such as PharmML (Swat et al. 2015; Bizzotto et al. 2017a). The advantage of the MDL-IDE workflows is realized by use of scripting in the R statistical computing language which enables full access to thousands of statistical and simulation packages. The IIF can be fully customized for speed, consistency, and fit-for-purpose modeling to better predict the toxicity and efficacy of the drug candidates.

References and Further Reading

Ahamadi M, Freshwater T, Prohn M, Li CH, de Alwis DP, de Greef R, Elassaiss-Schaap J, Kondic A, Stone JA (2017) Model-based characterization of the pharmacokinetics of pembrolizumab: a humanized anti-pd-1 monoclonal antibody in advanced solid tumors. CPT Pharmacometrics Syst Pharmacol 6:49–57

Alsharedi M, Bukamur H, Elhamdani A (2018) Osimertinib for the treatment of patients with EGFR mutation-positive non-small cell lung cancer. Drugs Today (Barc) 54:369–379. https://doi.org/10.1358/dot.2018.54.6.2817668

Atsumi R, Yoshiba S, Maruyama D, Tobinai K, Ishida T, Ishitsuka K, Imaizumi Y, Takeuchi S, Tsukasaki K, Adachi N, Fujitani S, Tachibana M, Yoshihara K, Ishizuka H (2017) Population pharmacokinetic and exposure-response modeling for the EZH1/2 dual inhibitor DS-3201b in patients with non-Hodgkin lymphomas. Blood 130(Suppl 1):2544

Barbolosi D, Ciccolini J, Lacarelle B, Barlési F, André N (2015) Computational oncology – mathematical modelling of drug regimens for precision medicine. Nat Rev Clin Oncol 13:242–254

Beaver JA, Ison G, Pazdur R (2017) Reevaluating eligibility criteria – balancing patient protection and participation in oncology trials. N Engl J Med 376:1504–1505

Beumer JH, Ding F, Tawbi H, Lin Y, Viluh D, Chatterjee I, Rinker M, Chow SL, Ivy SP (2016) Effect of renal dysfunction on toxicity in three decades of cancer therapy evaluation program-sponsored single-agent phase I studies. J Clin Oncol 34:110–116

Bizzotto R, Comets E, Smith G, Yvon F, Kristensen NR, Swat MJ (2017a) PharmML in action: an interoperable language for modeling and simulation. CPT Pharmacometrics Syst Pharmacol 6:651–665. https://doi.org/10.1002/psp4.12213

Bizzotto R, Blaudez E, Borella E, Carrara L, Chan P, Chenel M, Comets E, Gieschke R, Harling K, Harnisch L, Hartung N, Hooker AC, Karlsson MO, Kaye R, Kloft C, Kokash N, Lavielle M, Lestini G, Magni P, Mari A, Mentré F, Muselle C, Nordgren R, Nyberg HB, Parra-Guillén ZP, Pasotti L, Rode-Kristensen N, Sardu ML, Smith GR, Swat MJ, Terranova N, Yngman G, Yvon F, Holford N, DDMoRe Consortium (2017b) Model description language (MDL): a standard for modeling and simulation. CPT Pharmacometrics Syst Pharmacol 6:647–650. https://doi.org/10.1002/psp4.12222

Bleyer A, Tai E, Siegel S (2018) Role of clinical trials in survival progress of American adolescents and young adults with cancer – and lack thereof. Pediatr Blood Cancer 65:e27074. https://doi.org/10.1002/pbc.27074

Burke ME, Albritton K, Marina N (2007) Challenges in the recruitment of adolescents and young adults to cancer clinical trials. Cancer 110:2385–2393

Calvo E, Azaro A, Rodon J, Dirix L, Huizing M, Senecal FM, LoRusso P, Yee L, Poggesi I, de Jong J, Triantos S, Park YC, Knoblauch RE, Parekh TV, Demetri GD, von Mehren M (2018) Hepatic safety analysis of trabectedin: results of a pharmacokinetic study with trabectedin in patients with hepatic impairment and experience from a phase 3 clinical. Trial 36:476–486. https://doi.org/10.1007/s10637-017-0546-9

Chatterjee M, Turner DC, Felip E, Lena H, Cappuzzo F, Horn L, Garon EB, Hui R, Arkenau HT, Gubens MA, Hellmann MD, Dong D, Li C, Mayawala K, Freshwater T, Ahamadi M, Stone J, Lubiniecki GM, Zhang J, Im E, De Alwis DP, Kondic AG, Flotten O (2016) Systematic evaluation of pembrolizumab dosing in patients with advanced non-small-cell lung cancer. Ann Oncol 27:1291–1298. https://doi.org/10.1093/annonc/mdw174

Chavan BB, Tiwari S, Nimbalkar RD, Garg P, R S, Talluri MVNK (2018) In vitro and in vivo metabolic investigation of the Palbociclib by UHPLC-Q-TOF/MS/MS and in silico toxicity studies of its metabolites. J Pharm Biomed Anal 157:59–74. https://doi.org/10.1016/j.jpba.2018.05.008

Cheeti S, Budha NR, Rajan S, Dresser MJ, Jin JY (2013) A physiologically based pharmacokinetic (PBPK) approach to evaluate pharmacokinetics in patients with cancer. Biopharm Drug Dispos 34:141–154. https://doi.org/10.1002/bdd.1830

Chen CL, Levine A, Rao A, O'Neill K, Messinger Y, Myers DE, Goldman F, Hurvitz C, Casper JT, Uckun FM (1999) Clinical pharmacokinetics of the CD19 receptor-directed tyrosine kinase inhibitor B43-Genistein in patients with B-lineage lymphoid malignancies. J Clin Pharmacol 39:1248–1255

Cho YS, Noh YH, Lim HS, Cho SH, Ghim JL, Choe S, Kim SB, Park JS, Lee SK, Yang WS, Chang JW, Bahng MY, Bae KS (2018) Effects of renal impairment on the pharmacokinetics and safety of udenafil. J Clin Pharmacol 58:905–912. https://doi.org/10.1002/jcph.1095

Chuk MH, Mulugeta Y, Cline MR, Mehrotra N, Reaman GH (2017) Enrolling adolescents in disease/target-appropriate adult oncology clinical trials of investigational agents. Clin Cancer Res 23:9–12. https://doi.org/10.1158/1078-0432.CCR-16-1367

Darwich A, Ogungbenro K, Hatley OJ, Rostami-Hodjegan A (2017) Role of pharmacokinetic modeling and simulation in precision dosing of anticancer drugs. Transl Cancer Res 6:S1512–S1529. https://doi.org/10.21037/tcr.2017.09.14

de Man FM, Goey AKL, van Schaik RHN, Mathijssen RHJ, Bins S (2018) Individualization of Irinotecan treatment: a review of pharmacokinetics, pharmacodynamics, and pharmacogenetics. Clin Pharmacokinet. https://doi.org/10.1007/s40262-018-0644-7. [Epub ahead of print]

EC Regulation No 1901/2006 of the European Parliament and of the Council of 12 December 2006 on medicinal products for pediatric use and amending Regulation (EEC) No 1768/92, Directive 2001/20/EC, Directive 2001/83/EC and Regulation (EC) No 726/2004. https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2006_1901/reg_2006_1901_en.pdf

El-Khoueiry AB, Sarantopoulos J, Ciombor KK, Xu H, O'Gorman M, Chakrabarti J, Usari T, El-Rayes BF (2018) Evaluation of hepatic impairment on pharmacokinetics and safety of crizotinib in patients with advanced cancer. Cancer Chemother Pharmacol 81:659–670. https://doi.org/10.1007/s00280-018-3517-8

EMA – European Medicines Agency (2015) Guideline on the evaluation of the pharmacokinetics of medicinal products in patients with decreased renal function. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/02/WC500200841.pdf. Accessed 17 Dec 2015

FDA Guidance for Industry (1999) Population pharmacokinetics. https://www.fda.gov/downloads/drugs/guidances/UCM072137.pdf

FDA Guidance for Industry (2003a) Exposure-response relationships: study design, data analysis, and regulatory applications. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm072109.pdf

FDA Guidance for Industry (2003b) Pharmacokinetics in patients with impaired hepatic function: study design,

data analysis, and impact on dosing and labeling (CDER/CBER – FDA). https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM072123.pdf

FDA Guidance for Industry (2010) Pharmacokinetics in patients with impaired renal function – study design, data analysis, and impact on dosing and labeling. https://www.fda.gov/downloads/drugs/guidances/ucm204959.pdf

FDA Draft Guidance. Considerations for the Inclusion of Adolescent Patients in Adult Oncology Clinical Trials Guidance for Industry (2018) U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER) Oncology Center of Excellence (OCE)

Ferl GZ, Theil FP, Wong H (2016) Physiologically based pharmacokinetic models of small molecules and therapeutic antibodies: a mini-review on fundamental concepts and applications. Biopharm Drug Dispos 37:75–92

Fern LA, Taylor RM (2018) Enhancing accrual to clinical trials of adolescents and young adults with cancer. Pediatr Blood Cancer 11:e27233. https://doi.org/10.1002/pbc.27233

Forrest SJ, Geoerger B, Janeway KA (2018) Precision medicine in pediatric oncology. Curr Opin Pediatr 30:17–24

Fouladi M, Park JR, Stewart CF, Gilbertson RJ, Schaiquevich P, Sun J, Reid JM, Ames MM, Speights R, Ingle AM, Zwiebel J, Blaney SM, Adamson PC (2010) Pediatric phase I trial and pharmacokinetic study of vorinostat: a children's oncology group phase I consortium report. J Clin Oncol 28:3623–3629

Freshwater T, Kondic A, Ahamadi M, Li CH, de Greef R, de Alwis D, Stone JA (2017) Evaluation of dosing strategy for pembrolizumab for oncology indications. J Immunother Cancer 5:43. https://doi.org/10.1186/s40425-017-0242-5

Freyer DR, Felgenhauer J, Perentesis J (2013) COG adolescent and young adult oncology discipline committee. Children's oncology group's 2013 blueprint for research: adolescent and young adult oncology. Pediatr Blood Cancer 60:1055–1058

Gadkar K, Kirouac DC, Mager DE, van der Graaf PH, Ramanujan S (2016) A six-stage workflow for robust application of systems pharmacology. CPT Pharmacometrics Syst Pharmacol 5:235–249

Gaspar N, Marshall LV, Binner D, Herold R, Rousseau R, Blanc P, Capdeville R, Carleer J, Copland C, Kerloeguen Y, Norga K, Pacaud L, Sevaux MA, Spadoni C, Sterba J, Ligas F, Taube T, Uttenreuther-Fischer M, Chioato S, O'Connell MA, Geoerger B, Blay JY, Soria JC, Kaye S, Wulff B, Brugières L, Vassal G, Pearson ADJ, Members of Working Group 1 of the Paediatric Platform of ACCELERATE (2018) Joint adolescent adult early phase clinical trials to improve access to new drugs for adolescents with cancer: proposals from the multi-stakeholder platform-ACCELERATE. Ann Oncol 29(3):766–771. https://doi.org/10.1093/annonc/mdy002

Groenland SL, van Nuland M, Verheijen RB, Schellens JHM, Beijnen JH, Huitema ADR, Steeghs N (2018) Therapeutic drug monitoring of oral anti-hormonal drugs in oncology. Clin Pharmacokinet. https://doi.org/10.1007/s40262-018-0683-0. [Epub ahead of print]

Han X, Zhou Y, Liu W (2017) Precision cardio-oncology: understanding the cardiotoxicity of cancer therapy. NPJ Precis Oncol 1(1):31. https://doi.org/10.1038/s41698-017-0034-x. eCollection 2017

Horsmans Y, Zhou J, Liudmila M, Golor G, Shibolet O, Quinlan M, Emotte C, Boss H, Castro H, Sellami D, Preston RA (2018) Effects of mild to severe hepatic impairment on the pharmacokinetics of sonidegib: a multicenter, open-label, parallel-group study. Clin Pharmacokinet 57:345–354. https://doi.org/10.1007/s40262-017-0560-2

Jamal R, Lapointe R, Cocolakis E, Thébault P, Kazemi S, Friedmann JE, Dionne J, Cailhier JF, Bélanger K, Ayoub JP, Le H, Lambert C, El-Hajjar J, van Kempen LC, Spatz A, Miller WH Jr (2017) Peripheral and local predictive immune signatures identified in a phase II trial of ipilimumab with carboplatin/paclitaxel in unresectable stage III or stage IV melanoma. J Immunother Cancer 5 (83). https://doi.org/10.1186/s40425-017-0290-x

Janssen JM, Zwaan CM, Schellens HM, Beijenen H, Alwin D, Huitema R (2017) Clinical trial simulations in paediatric oncology: a feasibility study from the innovative therapies for children with cancer consortium. Eur J Cancer 85:78–85. https://doi.org/10.1016/j.ejca.2017.07.050

Kalgutkar AS, Dalvie D (2015) Predicting toxicities of reactive metabolite-positive drug candidates. Annu Rev Pharmacol Toxicol 55:35–54

Kim ES, Bruinooge SS, Roberts S, Ison G, Lin NU, Gore L, Uldrick TS, Lichtman SM, Roach N, Beaver JA, Sridhara R, Hesketh PJ, Denicoff AM, Garrett-Mayer E, Rubin E, Multani P, Prowell TM, Schenkel C, Kozak M, Allen J, Sigal E, Schilsky RL (2017) Broadening eligibility criteria to make clinical trials more representative: American Society of Clinical Oncology and Friends of Cancer Research joint research statement. J Clin Oncol 35(33):3737–3744. https://doi.org/10.1200/JCO.2017.73.7916

Krauss M, Schaller S, Borchers S, Findeisen R, Lippert J, Kuepfer L (2012) Integrating cellular metabolism into a multiscale whole-body model. PLoS Comput Biol 8 (10):e1002750

Lin KR, Huang JT, Henderson CJ, Wolf CR (2017) Novel pathways of ponatinib disposition catalyzed by CYP1A1 involving generation of potentially toxic metabolites. J Pharmacol Exp Ther 363:12–19. https://doi.org/10.1124/jpet.117.243246

Mansfield AS, Rudek MA, Vulih D, Smith GL, Harris PJ, Ivy SP (2016) The effect of hepatic impairment on outcomes in phase I clinical trials in cancer subjects. Clin Cancer Res 1–8. https://doi.org/10.1158/1078-0432.CCR-16-0449

Moreau P, Karamanesht II, Domnikova N, Kyselyova MY, Vilchevska KV, Doronin VA, Schmidt A, Hulin C, Leleu X, Esseltine DL, Venkatakrishnan K, Skee D, Feng H, Girgis S, Cakana A, van de Velde H, Deraedt W, Facon T (2012) Pharmacokinetic, pharmacodynamic and covariate analysis of subcutaneous versus intravenous administration of bortezomib in patients with relapsed multiple myeloma. Clin Pharmacokinet 51:823–829. https://doi.org/10.1007/s40262-012-0010-0

Niu J, Scheuerell C, Mehrotra S, Karan S, Puhalla S, Kiesel BF, Ji J, Chu E, Gopalakrishnan M, Ivaturi V, Gobburu J, Beumer JH (2017) Parent-metabolite pharmacokinetic modeling and pharmacodynamics of veliparib (ABT-888), a PARP inhibitor, in patients with BRCA 1/2-mutated cancer or PARP-sensitive tumor types. J Clin Pharmacol 57:977–987. https://doi.org/10.1002/jcph.892

Overgaard RV, Ingwersen SH, Tornoe CW (2015) Establishing good practices for exposure–response analysis of clinical endpoints in drug development. CPT Pharmacometrics Syst Pharmacol 4:565–575

Owonikoko TK, Busari AK, Kim S, Chen Z, Akintayo A, Lewis C, Carthon BC, Alese OB, El-Rayes BF, Ramalingam SS, Harvey RD (2018) Race-, age-, and gender-based characteristics and toxicities of targeted therapies on phase I trials. Oncology 1–9. https://doi.org/10.1159/000488763. [Epub ahead of print]

Paci A, Veal G, Bardin C, Levêque D, Widmer N, Beijnen J, Astier A, Chatelut E (2014) Review of therapeutic drug monitoring of anticancer drugs part 1 – Cytotoxics. Eur J Cancer 50:2010–2019

Paoletti X, Geoerger B, Doz F, Baruchel A, Lokiec F, Le Tourneau C (2013) A comparative analysis of paediatric dose-finding trials of molecularly targeted agent with adults' trials. Eur J Cancer 49:2392–2402

Porpaczy E, Tripolt S, Hoelbl-Kovacic A, Gisslinger B, Bago-Horvath Z, Casanova-Hevia E, Clappier E, Decker T, Fajmann S, Fux DA, Greiner G, Gueltekin S, Heller G, Herkner H, Hoermann G, Kiladjian JJ, Kolbe T, Kornauth C, Krauth MT, Kralovics R, Muellauer L, Mueller M, Prchal-Murphy M, Putz EM, Raffoux E, Schiefer AI, Schmetterer K, Schneckenleithner C, Simonitsch-Klupp I, Skrabs C, Sperr WR, Staber PB, Strobl B, Valent P, Jaeger U, Gisslinger H, Sexl V (2018) Aggressive B-cell lymphomas in patients with myelofibrosis receiving JAK1/2 inhibitor therapy. Blood 132:694–706. https://doi.org/10.1182/blood-2017-10-810739

Reaman G (2018) ASCO pediatric oncology award and lecture, 2018 ASCO annual symposium; relevant molecular targets in pediatric cancers: applicability to pediatric therapeutic investigations required under FDARA 2017

Reis-Mendes AF, Sousa E, de Lourdes Bastos M, Costa VM (2015) The role of the metabolism of anticancer drugs in their induced-cardiotoxicity. Curr Drug Metab 17:75–90

Rousseau A, Marquet P (2012) Application of pharmacokinetic modelling to the routine therapeutic drug monitoring of anticancer drugs. Fundam Clin Pharmacol 16:253–262. https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1472-8206.2002.00086.x

Rowland M (2013) Physiologically-based pharmacokinetic (PBPK) modeling and simulations principles, methods, and applications in the pharmaceutical industry. CPT Pharmacometrics Syst Pharmacol 2:e55

Sager JE, Yu J, Ragueneau-Majlessi I, Isoherranen N (2015) Physiologically based pharmacokinetic (PBPK) modeling and simulation approaches: a systematic review of published models, applications, and model verification. Drug Metab Dispos 43:1823–1837

Saeheng T, Na-Bangchang K, Karbwang J (2018) Utility of physiologically based pharmacokinetic (PBPK) modeling in oncology drug development and its accuracy: a systematic review. Eur J Clin Pharmacol. https://doi.org/10.1007/s00228-018-2513-6. [Epub ahead of print]

Schöffski P, Delord JP, Brain E, Robert J, Dumez H, Gasmi J, Trouet A (2017) First-in-man phase I study assessing the safety and pharmacokinetics of a 1-hour intravenous infusion of the doxorubicin prodrug DTS-201 every 3 weeks in patients with advanced or metastatic solid tumours. Eur J Cancer 86:240–247. https://doi.org/10.1016/j.ejca.2017.09.009

Schultze-Mosgau M-H, Hochel J, Zimmermann T, Brooks A, Bush J, Rottmann A (2018) Characterization of the pharmacokinetics of Vilaprisan: bioavailability, excretion, biotransformation and drug-drug interaction potential. Clin Pharmacokinet 57:1001–1015

Schwenger E, Reddy VP, Moorthy G, Sharma P, Tomkinson H, Masson E, Vishwanathan K (2018) Harnessing meta-analysis to refine an oncology patient population for physiology-based pharmacokinetic modeling of drugs. Clin Pharmacol Ther 103:271–280. https://doi.org/10.1002/cpt.917

Smith AW, Seibel NL, Lewis DR, Albritton KH, Blair DF, Blanke CD, Bleyer WA, Freyer DR, Geiger AM, Hayes-Lattin B, Tricoli JV, Wagner LI, Zebrack BJ (2016) Next steps for adolescent and young adult oncology workshop: an update on progress and recommendations for the future. Cancer 122:988–999

Smith MK, Moodie SL, Bizzotto R, Blaudez E, Borella E, Carrara L, Chan P, Chenel M, Comets E, Gieschke R, Harling K, Harnisch L, Hartung N, Hooker AC, Karlsson MO, Kaye R, Kloft C, Kokash N, Lavielle M, Lestini G, Magni P, Mari A, Mentré F, Muselle C, Nordgren R, Nyberg HB, Parra-Guillén ZP, Pasotti L, Rode-Kristensen N, Sardu ML, Smith GR, Swat MJ, Terranova N, Yngman G, Yvon F, Holford N, DDMoRe consortium (2017) Model description language (MDL): a standard for modeling and simulation. CPT Pharmacometrics Syst Pharmacol 6:647–650. https://doi.org/10.1002/psp4.12222

Stark D, Bielack S, Brugieres L, Dirksen U, Duarte X, Dunn S, Erdelyi DJ, Grew T, Hjorth L, Jazbec J, Kabickova E, Konsoulova A, Kowalczyk JR, Lassaletta A, Laurence V, Lewis I, Monrabal A, Morgan S, Mountzios G, Olsen PR, Renard M, Saeter G, van der Graaf WT, Ferrari A (2016) Teenagers and

young adults with cancer in Europe: from national programmes to a European integrated coordinated project. Eur J Cancer Care (Engl) 25:419–427

Sun Y, Kim JH, Vangipuram K, Hayes DF, Smith EML, Yeomans L, Henry NL, Stringer KA, Hertz DL (2018) Pharmacometabolomics reveals a role for histidine, phenylalanine, and threonine in the development of paclitaxel-induced peripheral neuropathy. Breast Cancer Res Treat 171(3):657–666. https://doi.org/10.1007/s10549-018-4862-3

Swat MJ, Moodie S, Wimalaratne SM, Kristensen NR, Lavielle M, Mari A, Magni P, Smith MK, Bizzotto R, Pasotti L, Mezzalana E, Comets E, Sarr C, Terranova N, Blaudez E, Chan P, Chard J, Chatel K, Chenel M, Edwards D, Franklin C, Giorgino T, Glont M, Girard P, Grenon P, Harling K, Hooker AC, Kaye R, Keizer R, Kloft C, Kok JN, Kokash N, Laibe C, Laveille C, Lestini G, Mentré F, Munafo A, Nordgren R, Nyberg HB, Parra-Guillen ZP, Plan E, Ribba B, Smith G, Trocóniz IF, Yvon F, Milligan PA, Harnisch L, Karlsson M, Hermjakob H, Le Novère N (2015) Pharmacometrics markup language (PharmML): opening new perspectives for model exchange in drug development. CPT Pharmacometrics Syst Pharmacol 4:316–319. https://doi.org/10.1002/psp4.57

Tan CRC, Abdul-Majeed S, Cael B, Barta SK (2018) Clinical pharmacokinetics and pharmacodynamics of bortezomib. Clin Pharmacokinet. https://doi.org/10.1007/s40262-018-0679-9. [Epub ahead of print]

Thai HT, Mazuir F, Cartot-Cotton S, Veyrat-Follet C (2015) Optimizing pharmacokinetic bridging studies in paediatric oncology using physiologically-based pharmacokinetic modelling: application to docetaxel. Br J Clin Pharmacol 80:534–547. https://doi.org/10.1111/bcp.12702

Thomas F, Veal GJ, El Balkhi S, Lafont T, Picard N, Brugières L, Chatelut E, Piguet C (2018a) Therapeutic drug monitoring and dose adaptation of cisplatin in a newborn with hepatoblastoma: a case report. Cancer Chemother Pharmacol. https://doi.org/10.1007/s00280-018-3625-5. [Epub ahead of print]

Thomas SM, Malvar J, Tran H, Shows J, Freyer DR (2018b) A prospective, observational cohort study comparing cancer clinical trial availability and enrollment between early adolescents/young adults and children. Cancer 124:983–990. https://doi.org/10.1002/cncr.31127

Tortorici MA, Cutler D, Zhang L, Pfister M (2012) Design, conduct, analysis, and interpretation of clinical studies in patients with impaired kidney function. J Clin Pharmacol 52(Suppl):109S–118S

Tsukamoto Y, Kato Y, Ura M, Horii I, Ishikawa T, Ishitsuka H, Sugiyama Y (2001) Investigation of 5-FU disposition after oral administration of capecitabine, a triple-prodrug of 5-FU, using a physiologically based pharmacokinetic model in a human cancer xenograft model: comparison of the simulated 5-FU exposures in the tumour tissue between human and xenograft model. Biopharm Drug Dispos 22:1–14

Turan T, Kannan D, Patel M, Matthew Barnes J, Tanlimco SG, Lu R, Halliwill K, Kongpachith S, Kline DE, Hendrickx W, Cesano A, Butterfield LH, Kaufman HL, Hudson TJ, Bedognetti D, Marincola F, Samayoa J (2018) Immune oncology, immune responsiveness and the theory of everything. J Immunother Cancer 6(1):50. https://doi.org/10.1186/s40425-018-0355-5

Turner D, Kondic AG, Anderson KM, Robinson A, Garon EB, Riess JW, Jain L, Mayawala K, Kang J, Ebbinghaus SW, Sinha V, de Alwis DP, Stone JA (2018) Pembrolizumab exposure-response assessments challenged by association of cancer cachexia and catabolic clearance. Clin Cancer Res: clincanres.0415.2018. https://doi.org/10.1158/1078-0432.CCR-18-0415. [Epub ahead of print]

Turner RM, Park BK, Pirmohamed M (2015) Parsing interindividual drug variability: an emerging role for systems pharmacology. Wiley Interdiscip Rev Syst Biol Med 7:221–241

Xiao JJ, Chen JS, Lum B, Graham RA (2017) A survey of renal impairment pharmacokinetic studies for new oncology drug approvals in the USA from 2010 to early 2015: a focus on development strategies and future directions. Anti-Cancer Drugs 28:677–701. https://doi.org/10.1097/CAD.0000000000000513

Uckun FM, Evans WE, Forsyth CJ, Waddick KG, Ahlgren LT, Chelstrom LM, Burkhardt A, Bolen J, Myers DE (1995) Biotherapy of B-cell precursor leukemia by targeting genistein to CD19-associated tyrosine kinases. Science 267:886–891

Uckun FM, Qazi S, Cely I, Sahin K, Shahidzadeh A, Ozercan I, Yin Q, Gaynon P, Termuhlen A, Cheng J, Yiv S (2013) Nanoscale liposomal formulation of a SYK P-site inhibitor against B-precursor leukemia. Blood 121:4348–4354. https://doi.org/10.1182/blood-2012-11-470633

Uckun FM, Myers DE, Qazi S, Ozer Z, Rose R, D'Cruz OJ, Ma H (2015) Recombinant human CD19L-sTRAIL effectively targets B cell precursor acute lymphoblastic leukemia. J Clin Invest 125:1006–1018. https://doi.org/10.1172/JCI76610

Uckun FM, Kenny N (2018) Regulatory reforms offer renewed hope for pediatric cancer Patients who are in urgent need for therapeutic innovations. Clin Res Pediatr 1(2):1–5

Undevia SD, Gomez-Abuin G, Ratain MJ (2005) Pharmacokinetic variability of anticancer agents. Nat Rev Cancer 5:447–458

Ursino M, Zohar S, Lentz F, Alberti C, Friede T, Stallard N, Comets E (2017) Dose-finding methods for phase I clinical trials using pharmacokinetics in small populations. Biom J 59:804–825

van Andel L, Rosing H, Tibben MM, Lucas L, Lubomirov R, Avilés P, Francesch A, Fudio S, Gebretensae A, Hillebrand MJX, Schellens JHM, Beijnen JH (2018) Metabolite profiling of the novel anti-cancer agent, plitidepsin, in urine and faeces in cancer patients after administration of ^{14}C-plitidepsin. Cancer Chemother Pharmacol. https://doi.org/10.1007/s00280-018-3637-1. [Epub ahead of print]

Vassal G, Rousseau R, Blanc P, Moreno L, Bode G, Schwoch S, Schrappe M, Skolnik J, Bergman L, Bradley-Garelik MB, Saha V, Pearson A, Zwierzina H (2015) Creating a unique, multi-stakeholder pediatric oncology platform to improve drug development for children and adolescents with cancer. Eur J Cancer 51:218–224

Veal GJ, Hartford CM, Stewart CF (2010) Clinical pharmacology in the adolescent oncology patient. J Clin Oncol 28:4790–4799

Véronneau-Veilleux F, Bélair J (2017) Modeling circadian fluctuations of pharmacokinetic parameters. Math Model Nat Phenom 12:146–161. https://doi.org/10.1051/mmnp/201712509

Waller CF, Tiessen RG, Lawrence TE, Shaw A, Liu MS, Sharma R, Baczkowski M, Kothekar MA, Micales CE, Barve A, Ranganna GM, Pennella EJ (2018) A pharmacokinetics and pharmacodynamics equivalence trial of the proposed pegfilgrastim biosimilar, MYL-1401H, versus reference pegfilgrastim. J Cancer Res Clin Oncol 144:1087–1095. https://doi.org/10.1007/s00432-018-2643-3

Wicki A, Brown N, Xyrafas A, Bize V, Hawle H, Berardi S, Cmiljanović N, Cmiljanović V, Stumm M, Dimitrijević S, Herrmann R, Prêtre V, Ritschard R, Tzankov A, Hess V, Childs A, Hierro C, Rodon J, Hess D, Joerger M, von Moos R, Sessa C, Kristeleit R (2018) First-in human, phase 1, dose-escalation pharmacokinetic and pharmacodynamic study of the oral dual PI3K and mTORC1/2 inhibitor PQR309 in patients with advanced solid tumors (SAKK 67/13). Eur J Cancer 96:6–16. https://doi.org/10.1016/j.ejca.2018.03.012

Wilkins JJ, Chan P, Chard J, Smith G, Smith MK, Beer M, Dunn A, Flandorfer C, Franklin C, Gomeni R, Harnisch L, Kaye R, Moodie S, Sardu ML, Wang E, Watson E, Wolstencroft K, Cheung S, DDMoRe Consortium (2017) Thoughtflow: standards and tools for provenance capture and workflow definition to support model-informed drug discovery and development. CPT Pharmacometrics Syst Pharmacol 5:285–292

Yankeelov TE, An G, Saut O, Luebeck EG, Popel AS, Ribba B, Vicini P, Zhou X, Weis JA, Ye K, Genin GM (2016) Multi-scale modeling in clinical oncology: opportunities and barriers to success. Ann Biomed Eng 44:2626–2641

Pharmacogenomics in and Its Influence on Pharmacokinetics

56

Guy Montay, Jochen Maas, and Roland Wesch

Contents

Abstract

CYP1A2 is involved to a major extent in the metabolism of several drugs (imipramine, clozapine, fluvoxamine, olanzapine, theophylline, acetaminophen, propranolol, and tacrine) as well as of diet components (methylxanthines), endogenous substrates (estrogens), numerous aryl, aromatic and heterocyclic amines, and polycyclic aromatic hydrocarbons. It is

G. Montay (✉)
GMPK/PK Sanofi-Aventis, Vitry-Sur-Seine, France
e-mail: Guy.Montay@sanofi-aventis.com

J. Maas · R. Wesch
R&D Metabolism and PK Germany, Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany
e-mail: Jochen.Maas@sanofi-aventis.com; Roland.Wesch@sanofi-aventis.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_25

inducible, notably by cigarette smoking, diet habits such as consumption of cruciferous vegetables (e.g., broccoli, watercress, collard greens, Brussels sprouts, and mustard) and of charbroiled meats, some drugs (omeprazole, phenytoin, and rifampicin) and is a target enzyme for the development of some cancers. Up to now, more than 25 CYP1A2 alleles have been detected. Probe drugs for CYP1A2 phenotyping are caffeine and theophylline. For safety concerns and drug availability, the preferred probe is caffeine. Caffeine 3-demethylation is mediated by CYP1A2, and accounts for 80% of caffeine clearance. Caffeine is also a probe drug for N-acetyltransferase and xanthine oxidase (Clin Pharmacol Ther 53:203–514, 1993).

II.T.1	Phase I enzymes
II.T.1.1	CYP1A2
II.T.1.2	CYP2C9
II.T.1.3	CYP2C19
II.T.1.4	CYP2D6
II.T.1.5	CYP3A
II.T.1.6	Other CYPS
II.T.1.6.1	CYP2A6
II.T.1.6.2	CYP2B6
II.T.1.6.3	CYP2C8
II.T.1.6.4	CYP2E1
II.T.2	Phase II enzymes
II.T.2.1	N-acetyltransferases
II.T.2.2	Uridine diphosphate glucuronosyltransferases
II.T.2.3	Methyltransferases
II.T.2.4	Glutathione S-transferases and sulfotransferases

The understanding of the role of pharmacogenetics in drug metabolism expanded greatly in the 1990s. This is mainly due to technological improvements in gene scanning and gene variant identification. The number of variant alleles identified for genes coding for drug metabolizing enzymes (DME) considerably increased in the early 2000s, and continues to increase. The clinical consequences – or at least genotyping–phenotyping relationships – of DME polymorphisms have not been demonstrated for all variants. In the text below, only those DME allele variants will be mentioned for which significant changes in enzyme activity have been found using probe drugs. Comprehensive information on the nomenclature of cytochrome P450 (CYP) alleles can be found at www.imm.ki.se/CYPalleles and Phase I and Phase II DMEs at www.pharmgkb.org/index.jsp.

Phase I Enzymes

CYP1A2

Purpose and Rationale

CYP1A2 is involved to a major extent in the metabolism of several drugs (imipramine, clozapine, fluvoxamine, olanzapine, theophylline, acetaminophen, propranolol, and tacrine) as well as of diet components (methylxanthines), endogenous substrates (estrogens), numerous aryl, aromatic and heterocyclic amines, and polycyclic aromatic hydrocarbons. It is inducible, notably by cigarette smoking, diet habits such as consumption of cruciferous vegetables (e.g., broccoli, watercress, collard greens, Brussels sprouts, and mustard) and of charbroiled meats, some drugs (omeprazole, phenytoin, and rifampicin) and is a target enzyme for the development of some cancers. Up to now, more than 25 CYP1A2 alleles have been detected. Probe drugs for CYP1A2 phenotyping are caffeine and theophylline. For safety concerns and drug availability, the preferred probe is caffeine. Caffeine 3-demethylation is mediated by CYP1A2, and accounts for 80% of caffeine clearance. Caffeine is also a probe drug for N-acetyltransferase and xanthine oxidase (Kalow and Tang 1993).

Procedure

Phenotyping: A fixed or weight-adjusted dose of caffeine (solution, tablet, and coffee) ranging from 1 to 3 mg/kg is administered. Diet requirements have to be respected (stable xanthine-free diet avoiding beverages such as coffee, tea, cola,

chocolate, no food component with CYP1A2-inducing properties) during the test period. As smoking is known to induce CYP1A2, control of stable smoking status is mandatory.

There are two commonly used and robust methods for phenotyping. The first one measures caffeine (1,3,7-methylxanthine) and its N-demethylated metabolite 1,7-dimethylxanthine (paraxanthine) in a plasma or saliva sample collected within 5–7 h post-caffeine dosing (Fuhr and Rost 1994). The second one uses the assay of the metabolites 1-methylurate (1 U), 1-methylxanthine (1X), 5-acetylamino-6-formylamino-3-methyluracil (AFMU), and 1,7-dimethylurate (17 U) levels in urine collected at least for 8 h post-dosing (Campbell et al. 1987; Rostami-Hodjegan et al. 1996).

Commonly used methods for caffeine and metabolite(s) assay in plasma or urine involve an extraction step followed by HPLC with UV detection (Krul and Hageman 1998a; Rasmussen and Bosen 1996; Schreiber-Deturmeny and Bruguerolle 1996). Urine needs to be acidified (pH 3.0–3.5) before sample freezing.

Genotyping: Reduced activity has been reported for CYP1A2*1C and CYP1A2*1F alleles in smoking subjects. Induction of CYP1A2 activity has been associated with these alleles, but the effect of CYP1A2*1F mutation on CYP1A2 activity has not been confirmed (Nordmark et al. 2002). In Caucasians, frequency of the CYP1A2*1C and CYP1A2*1F variants is about 1% and 33%, respectively (Sachse et al. 2003).

Evaluation

Metabolic ratios (MR) used are plasma 17X/137X and urinary (1 U + 1X + AFMU)/17 U.

In controlled conditions, in nonsmoking young and elderly subjects, intraindividual and interindividual variability in 17X/137X MR was about 17% and 47%, respectively, with no effect of age (Simon et al. 2003). A 70-fold range in MR has been observed in smoking and nonsmoking female Caucasian subjects using the urinary MR (Nordmark et al. 1999). Up to 200-fold differences were found using the urinary test. Lower variability is expected using the plasma caffeine test.

Higher CYP1A2 activity in men versus women has been reported, though inconsistently, and in children. Higher MR is usually observed in smokers versus nonsmokers, when population sample size is large. Pregnancy and oral contraceptives intake were found to decrease CYP1A2 activity (Abernathy and Todd 1985; Caubet et al. 2004; Kalow and Tang 1993). CYP1A2 activity was found lower in colorectal patients versus controls (Sachse et al. 2003).

Large variability in CYP1A2 activity explains that its distribution has been described unimodal, bimodal, or trimodal. Poor metabolizers (PM, characterized with a MR <0.12) have been identified in Chinese population and represented about 5% of the population tested, whereas PM could represent 5–10% of Caucasian populations and 14% in Japanese population (Ou-Yang et al. 2000).

Critical Assessment of the Method

Numerous studies have shown good correlation between the 17X/137X plasma MR and caffeine systemic clearance, and plasma MR is considered more robust than the urinary one, since this last one can be affected by the effect of urinary flow on metabolite renal clearances.

Currently, no relationship between CYP1A2 genotype characteristics and CYP1A2 activity, as assessed by the caffeine test, has been usually found. Some associations have been found in specific genetic and environmental conditions (Han et al. 2001). Non-well-controlled conditions for urine sample collection, the effects (induction) linked to environmental factors may overcome the role of CYP1A2 polymorphism, which can explain the paucity of clear associations between CYP1A2 genotyping and phenotyping.

Further investigations are needed to characterize the effect of variants (SNPs, haplotypes) on CYP1A2 activity.

Modifications of the Method

Recent drug assay development involved LC-MS methods (Caubet et al. 2004; Kanazawa et al. 2000). A less practical breath test, using ^{13}C or ^{14}C labeled caffeine, can also be used (Kalow and Tang 1991).

References and Further Reading

Abernathy DR, Todd EL (1985) Impairment of caffeine clearance by chronic use of low-dose oestrogen-containing oral contraceptives. Eur J Clin Pharmacol 28:425–428

Brentano C, Jaillon P (2001) Variability of cytochrome P450 1A2 activity over time in young and elderly healthy volunteers. Br J Clin Pharmacol 52:601–604

Campbell ME, Spielberg SP, Kalow W (1987) A urinary metabolic ratio that reflects systemic caffeine clearance. Clin Pharmacol Ther 42:157–165

Caubet MS, Comte B, Brazier JL (2004) Determination of urinary 13C-caffeine metabolites by liquid chromatography–mass spectrometry: the use of metabolic ratios to assess CYP1A2 activity. J Pharm Sci Biomed Anal 34:379–389

Fuhr U, Rost LK (1994) Simple and reliable CYP1A2 phenotyping by the paraxanthine/caffeine ratio in plasma and saliva. Pharmacogenetics 4:109–116

Han XM, Ou-Yang DS, Lu PX et al (2001) Plasma caffeine metabolite ratio (17X/137X) in vivo associated with G-2964A and C734A polymorphisms of human CYP1A2. Pharmacogenetics 11:429–435

Kalow W, Tang BK (1991) Use of caffeine metabolite ratios to explore CYP1A2 and xanthine oxidase activities. Clin Pharmacol Ther 50:508–519

Kalow W, Tang BK (1993) The use of caffeine for enzyme assays: a critical appraisal. Clin Pharmacol Ther 53:203–514

Kanazawa H, Atsumi R, Matsushima Y et al (2000) Determination of theophylline and its metabolites in biological samples by liquid chromatography–mass spectrometry. J Chromatogr A 870:87–96

Krul C, Hageman G (1998a) Analysis of urinary caffeine metabolites to assess biotransformation enzyme activities by reverse-phase high-performance liquid chromatography. J Chromatogr B 709:24–27

Nordmark A, Lundgren S, Cnattingius S et al (1999) Dietary caffeine as a probe agent for assessment of cytochrome P450 1A2 activity in random urine samples. Br J Clin Pharmacol 47:397–402

Nordmark A, Lundgren S, Ask B et al (2002) The effect of the CYP1A2*F mutation on CYP1A2 inducibility in pregnant women. Br J Clin Pharmacol 54:504–510

Ou-Yang DS, Huang SL, Wang W et al (2000) Phenotypic polymorphism and gender-related differences of CYP1A2 activity in a Chinese population. Br J Clin Pharmacol 49:145–151

Rasmussen BB, Bosen K (1996) Determination of urinary metabolites of caffeine for the assessment of cytochrome P450 1A2, xanthine oxidase and N-acetyltransferase activity in humans. Ther Drug Monitor 18:254–262

Rostami-Hodjegan A, Nurminen S, Jackson PR, Tucker GT (1996a) Caffeine urinary metabolic ratios as markers of enzyme activity: a theoretical assessment. Pharmacogenetics 6:121–149

Sachse S, Bhambra U, Smith G et al (2003) Polymorphisms in the cytochrome P4501A2 gene (CYP1A2) in colorectal cancer patients and control: allele frequencies, linkage disequilibrium and influence on caffeine metabolism. Br J Clin Pharmacol 55:68–76

Schreiber-Deturmeny E, Bruguerolle B (1996) Simultaneous high-performance liquid chromatographic determination of caffeine and theophylline for routine drug monitoring in human plasma. J Chromatogr B 677:305–312

Shirley KL, Hon YY, Penzak S et al (2003) Correlation of cytochrome P450 1A2 activity using caffeine phenotyping and olanzapine disposition in healthy volunteers. Neuropsychopharmacology 25:961–966

Shou M, Korzekwa KR, Brooks EN et al (1997) Role of human hepatic cytochrome P450 1A2 and 3A4 in the metabolic activation of estrone. Carcinogenesis 18:207–214

Simon T, Becquemont L, Hamon B et al (2003) Determination of –3858G-A and 164C-A genetic polymorphisms of CYP1A2 in blood and saliva by rapid alleleic discrimination: large difference in the prevalence of the –3858G-A mutation between Caucasians and Asians. Eur J Clin Pharmacol 59:343–346

CYP2C9

Purpose and Rationale

CYP2C9 is involved in the hydroxylation of about 16% of drugs (Schwarz 2003), including drugs with narrow therapeutic index such as anticoagulants (warfarin, acenocoumarol, and phenprocoumon active S-enantiomers), and anticonvulsivants (phenytoin and hexobarbital), as well as numerous antidiabetic agents (i.e., tolbutamide, glibenclamide, and glipizide), antihypertensive drugs (losartan, irbesartan), nonsteroidal anti-inflammatory agents (i.e., diclofenac, ibuprofen, and celecoxib), diuretic (torsemide), and antirheumatoid agents (leflunomide).

A couple of CYP2C9 variants – mainly CYP2C9*2 and CYP2C9*3 – code for in vivo decreased activity, and two – CYP2C9*6 and CYP2C9*15 – have been reported to be associated with no activity. In Caucasian populations, CYP2C9*2 and CYP2C9*3 are encountered in 20–25% of subjects, while these genotypes have been found in less than 5% of East Asian subjects (Rosemary and Adithan 2007).

Probe drugs regularly used for CYP2C9 phenotyping are tolbutamide, warfarin, phenytoin, and losartan. Diclofenac, flurbiprofen, phenprocoumon, and torsemide have also been used. For safety concerns, the current preferred probe is tolbutamide, despite some risk of hypoglycemia.

Procedure

Phenotyping: The method measures tolbutamide, its CYP2C9-formed 4′-hydroxylated metabolite hydroxytolbutamide and the subsequent carboxytolbutamide metabolite, the latter formed by dehydrogenase enzymes. The urinary excretion of these two metabolites represented more than 85% dose of administered tolbutamide (Veronese et al. 1990, 1993).

Subjects receive a single oral 500 mg tolbutamide tablet in usual Phase I standard controlled conditions, with care to be paid to blood glucose. Urine is collected from drug intake to 8 or 24 h post-dosing.

The assay of tolbutamide and its metabolites is usually performed using HPLC and UV or fluorescence detection (Csillag et al. 1989; Veronese et al. 1990; Kirchheiner et al. 2002a, b; Hansen and Brosen 1999).

Genotyping: About two-third of Caucasian subjects express the wild genotype C9*1/*1. C9*1/*2 and C9*1/*3 heterozygote variants are expressed in 15–25% and 7–16% of Caucasian subjects, whereas the frequency of other variants is lower: 0.5–2.5%, 1–3%, and <1–1.5% for C9*2/*2, C9*2/*3, and C9*3/*3 variants, respectively (Scordo et al. 2001; Lee et al. 2002a, b; Schwarz 2003). More than 95% of Afro-American subjects express the wild genotype C9*1/*1 (Lee et al. 2002a, b). In Asian populations, CYP2C9*1/*3 is expressed in 2–8% subjects, but CYP2C9*2 is absent or extremely rare (Rosemary and Adithan 2007; Schwarz 2003; Xie et al. 2002). Overall, it has been estimated that 0.2–1% and 2–3% of Caucasian and Asian population could be qualified as PM, respectively (Meyer 2000).

Evaluation

The urinary MR (MR, hydroxytolbutamide + carboxytolbutamide)/tolbutamide is generally used. There is a large interindividual variability in MRs in subjects with the same genotype. Different studies performed with different probe drugs (Yasar et al. 2002a, b; Kirchheiner et al. 2002a, b, 2003a; Lee et al. 2002a, b; Miners and Birkett 1998; Morin et al. 2004), highlighted that a PM status could be given to subjects who are homozygous for CYP2C9*3, or expressing CYP2C9*2/*3 variant, but intermediate situations – from extensive to slow metabolizer status – may vary not only among different allele combinations but also with the probe drug used.

Oral contraceptives were found to inhibit CYP2C9 activity using losartan for phenotyping (Sandberg et al. 2004).

Critical Assessment of the Method

The tolbutamide test has the most convincing ability to discriminate between genotype variants and pharmacokinetics. There could be an analytical issue linked to the urine assay precision, as the urinary concentrations of the parent drug are very low in comparison with those of its metabolites.

To date, the CYP2C9*3 variant has been the only one found influencing significantly drug pharmacodynamics for warfarin, acenocoumarol (Sandberg 2003; Morin et al. 2004; Versuyft et al. 2003), glipizide, and glyburide (Kirchheiner et al. 2002a, b) or drug side effects (Sevilla-Mantilla et al. 2004). Inconstant results were found regarding tolbutamide effects (Kirchheiner et al. 2002a, b; Shong et al. 2002). For anticoagulants, the possession of CYP2C9*2 and CYP2C9*3 variants was associated with decreased warfarin dose requirement in patients, and an increased risk of adverse events such as bleeding (Daly and King 2003). An Afro-American subject with only the CYP2C9*6 variant exhibited serious phenytoin side effects associated with a marked impaired elimination of the drug (Kidd et al. 2001).

The variability of CYP2C9 activity observed among ethnic groups cannot be explained with our current knowledge on CYP2C9 variant alleles distribution (Xie et al. 2002).

Modifications of the Method

Losartan (25 mg dose) has been proposed as a safer alternative to tolbutamide. The determination of losartan/E3174 (oxidized metabolite) ratio in 0–8 h urine or in plasma at 6 h post-dosing have been proposed (Yasar et al. 2002a, b; Sekino et al. 2003). However, in a comparative study in 16 subjects, a better correlation between genotyping and phenotyping was found with tolbutamide, as compared to losartan or flurbiprofen, though there was no subject with the C9*2/*3 or C9*3/*3 variants (Lee et al. 2003).

Recently, a 125 mg tolbutamide dose has been validated, with proposal of the use of just one blood sample collected 24 h post-dosing. Its safer use needs the drug to be assayed using LC-MS/MS methodology (Jetter et al. 2004).

References and Further Reading

Csillag K, Vereczkey L, Gachalyi B (1989) Simple high-performance liquid chromatographic method for the determination of tolbutamide and its metabolites in human plasma and urine using photo-diode array detection. J Chromatogr 490:355–363

Daly AK, King BP (2003) Pharmacogenetics of anticoagulants. Pharmacogenetics 13:247–252

Hansen LL, Brosen K (1999) Quantitative determination of tolbutamide and its metabolites in human plasma and urine by high-performance liquid chromatography and UV detection. Ther Drug Monitor 21:664–671

Jetter A, Kinzig-Schippers M, Skott A et al (2004) Cytochrome P450 2C9 phenotyping using low-dose tolbutamide. Eur J Clin Pharmacol 60:165–175

Kidd RS, Curry TB, Gallagher S et al (2001) Identification of a null allele of CYP2C9 in an African-American exhibiting toxicity to phenytoin. Pharmacogenetics 11:803–808

Kirchheiner J, Bauer S, Meineke I et al (2002a) Impact of CYP2C9 and CYP2C19 polymorphisms on tolbutamide kinetics and the insulin and glucose response in healthy volunteers. Pharmacogenetics 12:101–109

Kirchheiner J, Brockmöller J, Meineke I et al (2002b) Impact of CYP2C9 amino acid polymorphisms on glyburide kinetics and on the insulin and glucose response in healthy volunteers. Clin Pharmacol Ther 71:286–296

Kirchheiner J, Störmer E, Meisel C et al (2003a) Influence of CYP2C9 genetic polymorphisms on pharmacokinetics of celecoxib and its metabolites. Pharmacogenetics 13:473–480

Lee CR, Pieper JA, Hinderliter AL et al (2002a) Evaluation of cytochrome P450 2C9 metabolic activity with tolbutamide in CYP2C9*1 heterozygotes. Clin Pharmacol Ther 72:562–571

Lee CR, Goldstein JA, Pieper JA (2002b) Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in vitro and human data. Pharmacogenetics 12:251–263

Lee CR, Pieper JA, Reginald FF et al (2003) Tolbutamide, flurbiprofen and losartan as probes of CYP2C9 activity in humans. J Clin Pharmacol 43:84–91

Meyer UA (2000) Pharmacogenetics and adverse drug reactions. Lancet 356:1667–1671

Miners JO, Birkett DJ (1998) Cytochrome P450 2C9: an enzyme of major importance in human drug metabolism. Br J Clin Pharmacol 45: 525–538

Morin S, Bodin L, Loriot MA et al (2004) Pharmacogenetics of acenocoumarol. Clin Pharmacol Ther 75:403–414

Rosemary J, Adithan C (2007a) The pharmacogenetics of CYP2C9 and CYP2C19: ethnic variation and clinical significance. Curr Clin Pharmacol 2:93–109

Sandberg M, Johansson I, Christensen M et al (2004) The impact of CYP2C9 genetics and oral contraceptives on cytochrome P450 2C9 phenotype. Drug Metab Dispos 32:484–489

Schwarz UI (2003) Clinical relevance of genetic polymorphisms in the human CYP2C9 gene. Eur J Clin Invest 33:23–30

Scordo MG, Aklillu E, Dahl ML et al (2001) Genetic polymorphism of cytochrome P450 2C9 in a Caucasian and a black African population. Br J Clin Pharmacol 52:447–450

Sekino K, Kubota T, Okada Y et al (2003) Effect of the single CYP2C9*3 allele on pharmacokinetics and pharmacodynamics in healthy Japanese subjects. Eur J Clin Pharmacol 59:589–592

Sevilla-Mantilla C, Ortega L, Agundez JA et al (2004) Leflunomide acute hepatitis. Dig Liver Dis 36:82–84

Shong JH, Yoon YR, Kim KA et al (2002) Effects of CYP2C9 and CYP2C19 genetic polymorphisms on the disposition and blood glucose lowering response to tolbutamide in humans. Pharmacogenetics 12: 111–119

Veronese ME, Miners JO, Randles D et al (1990) Validation of the tolbutamide metabolic ratio for population screening with use of sulfaphenazole to produce model phenotypic poor metabolizers. Clin Pharmacol Ther 47:403–411

Veronese ME, Miners JO, Ress DLP, Birkett DJ (1993) Tolbutamide hydroxylation in humans: lack of bimodality in 106 healthy subjects. Pharmacogenetics 3:86–93

Versuyft C, Robert A, Morin S et al (2003) Genetic and environmental risk factors for oral anticoagulant overdose. Eur J Clin Pharmacol 58: 739–745

Xie HG, Prasad HC, Kim RB, Stein CM (2002) CYP2C9 allelic variants: ethnic distribution and functional significance. Adv Drug Deliv Rev 54:1257–1270

Yasar U, Dahl ML, Christensen M, Eliasson E (2002a) Intra-individual variability in urinary losartan oxidation ratio: an in vivo marker of CYP2C9 activity. Br J Clin Pharmacol 54:183–185

Yasar U, Forslund-Bergengren C, Tybring G et al (2002b) Pharmacokinetics of losartan and its metabolite E-3174 in relation to the CYP2C9 genotype. Clin Pharmacol Ther 71:89–98

CYP2C19

Purpose and Rationale

CYP2C19 contributes to the metabolism of about 8% of drugs (Rogers et al. 2002), including S-mephenytoin, proton pump inhibitors (omeprazole, lansoprazole, and pantoprazole), tricyclic antidepressants (amitriptyline, imipramine, clomipramine, and citalopram), benzodiazepines (diazepam and flunitrazepam), torsemide, fluvastatin, and proguanil. Two main variants – CYP2C19*2 and CYP2C19*3 – are coding for in vivo nil activity, as well as CYP2C19*4, *5, *6, *7, and *8 variants. About 15–20% Asians, 4–7% Black Africans, and 3% Caucasians are PM (Scordo et al. 2004).

Probe drugs used for CYP2C19 phenotyping are mephenytoin, omeprazole, and proguanil. The most currently used probe drug is omeprazole.

Procedure

Phenotyping: The method measures omeprazole, and its CYP2C19-formed 5-hydroxylated metabolite in plasma.

Subjects receive a single oral 20 or 40 mg omeprazole capsule in usual Phase I standard controlled conditions. Plasma can be collected from drug intake up to 24 h post-dosing, or only one plasma sample is collected at 2 or 3 h post-dosing.

The assay of omeprazole and its metabolite is usually performed using HPLC and UV detection (Lagerstrom and Persson 1984; Ieri 1996; Yim et al. 2001; Tybring et al. 1997) or LC-MS/MS assay (Kanazawa et al. 2002).

Genotyping: The two alleles CYP2C19*2 and CYP2C19*3 account for quite all PM in Asians (>99%) and Black Africans, but defective alleles have not been fully characterized in 10–15% Caucasians. The CYP2C19*2 allele is the most frequent in Asian populations (30% in Chinese), as well as in Black Africans (about 17%) and in Caucasians (about 15%) (Xie et al. 2001). The CYP2C19*3 accounts for about 25% of inactive forms in Orientals, and is extremely rare in Caucasians (Scordo et al. 2004; Rosemary and Adithan 2007).

Evaluation

The AUC or plasma ratio of omeprazole to 5-hydroxyomeprazole is used.

As expected, homozygous PM subjects have lower metabolic activity as compared to heterozygous PM subjects, and potential interethnic difference has been noticed within a genotype (Yin et al. 2004).

Decreased CYP2C19 activity has been observed with oral contraceptives containing ethinylestradiol (Tamminga et al. 1999; Laine et al. 2000).

Critical Assessment of the Method

Omeprazole hydroxylation rate correlates with S-mephenytoin hydroxylation rate, which was initially the CYP2C19 probe drug (Andersson et al. 1990; Chang et al. 1995; Balian et al. 1995). The alternate pathway – conversion of omeprazole to its sulfone derivative – that is mediated via CYP3A4, does not influence the CYP2C19 pathway of omeprazole (Balian et al. 1995).

Time-dependent kinetics of omeprazole limits its use for phenotyping during chronic therapy (Gafni et al. 2001). CYP2C19 phenotyping with omeprazole may be affected by age, liver disease, and omeprazole therapy (Kimura et al. 1999).

Interethnic differences observed with different CYP2C19 substrates for subjects with same genotype have been attributed to differences in substrate specificity or enzyme isoforms (Bertilsson et al. 1992). The clearance of omeprazole is higher in Caucasian extensive metabolizers (EM) than in Oriental EM, due to a higher proportion of heterozygous EM in this latter population (Ishizaki et al. 1994).

Modifications of the Method

It has been proposed to use omeprazole for both CYP2C19 and CYP3A4 phenotyping (Gonzalez et al. 2003).

References and Further Reading

Andersson T, Regardh CG, Dahl-Puustinen ML, Bertilsson L (1990) Slow omeprazole metabolizers are also poor S-mephenytoin hydroxylators. Ther Drug Monit 12:415–416

Balian JD, Sukhova N, Harris JW et al (1995) The hydroxylation of omeprazole correlates with S-mephenytoin metabolism: a population study. Clin Pharmacol Ther 57:662–669

Bertilsson L, Lou YQ, Du YL et al (1992) Pronounced differences between native Chinese and Swedish populations in the polymorphic hydroxylations of debrisoquin and S-mephenytoin. Clin Pharmacol Ther 51:388–397

Chang M, Dahl ML, Tybring G et al (1995) Use of omeprazole as a probe drug for CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and CYP2C19 genotype. Pharmacogenetics 5:358–363

Gafni I, Nolte H, Tyndale R et al (2001) Resolving the roles of CYP2C19 and CYP3A4 in the metabolism of omeprazole in vivo using chronic omeprazole and ketoconazole. FASEB J 15:A918

Gonzalez HM, Romero EM, Peregrina AA et al (2003) CYP2C19- and CYP3A4-dependent omeprazole metabolism in west Mexicans. J Clin Pharmacol 43:1211–1215

Ishizaki T, Sohn DR, Kobayashi K et al (1994) Interethnic differences in omeprazole metabolism in the two S-mephenytoin hydroxylation phenotypes studied in Caucasians and Orientals. Ther Drug Monit 16:214–215

Kanazawa H, Okada A, Matsushima Y et al (2002) Determination of omeprazole and its metabolites in human plasma by liquid

chromatography–mass spectrometry. J Chromatogr A 949:1–9
Kimura M, Ieri I, Wada Y, Mamyia K (1999) Reliability of the omeprazole hydroxylation index and length of therapy. Br J Clin Pharmacol 47:115–119
Lagerstrom PO, Persson BA (1984) Determination of omeprazole and its metabolites in plasma and urine by liquid chromatography. J Chromatogr A 309:347–356
Laine K, Tybring G, Bertilsson L (2000) No sex-related difference but significant inhibition by oral contraceptives of CYP2C19 activity as measured by the probe drugs mephenytoin and omeprazole in healthy Swedish white subjects. Clin Pharmacol Ther 68:151–159
Rogers JF, Nafziger AN, Bertino JS (2002) Pharmacogenetics affects dosing, efficacy, and toxicity of cytochrome P450-metabolized drugs. Am J Med 113:746–750
Rosemary J, Adithan C (2007b) The pharmacogenetics of CYP2C9 and CYP2C19: ethnic variation and clinical significance. Curr Clin Pharmacol 2:93–109
Scordo MG, Caputi AP, D'Arrigo C et al (2004) Allele and genotype frequencies of CYP2C9, CYP2C19 and CYP2D6 in an Italian population. Pharmacol Res 50:195–200
Tamminga WJ, Werner J, Oosterhuis B et al (1999) CYP2D6 and CYP2C19 activity in a large population of Dutch healthy volunteers: indications for oral contraceptive-related gender differences. Eur J Clin Pharmacol 55:177–184
Tybring G, Bottiger Y, Wyden J, Bertilsson L (1997) Enantioselective hydroxylation of omeprazole catalyzed by CYP2C19. Clin Pharmacol Ther 62:129–137
Xie HG, Kim RB, Wood AJ, Stein CM (2001a) Molecular basis of ethnic differences in drug disposition and response. Annu Rev. Pharmacol Toxicol 41:815–850
Yim DS, Jeong JR, Park JY (2001) Assay of omeprazole and omeprazole sulfone by semi-microcolumn liquid chromatography with mixed-function precolumn. J Chromatogr B 754:487–493
Yin OQP, Tomlinson B, Chow ALH et al (2004) Omeprazole as a CYP2C19 marker in Chinese subjects: assessment of its gene-dose effect and intrasubject variability. J Clin Pharmacol 44:582–589

CYP2D6

Purpose and Rationale

CYP2D6 is involved significantly in the metabolism of drugs mainly used in CNS (antidepressants, i.e., imipramine, paroxetine, citalopram; neuroleptics, i.e., haloperidol, risperidone), or cardiovascular (β-adrenoceptor blockers, i.e., metoprolol; antiarrhythmics, i.e., propafenone and flecainide) disorders. Significant interethnic and interindividual intraethnic differences in CYP2D6 activity have been found. It is found that 5–10% Caucasians, 6–8% Afro-Americans, and only 1% Asians have reduced CYP2D6 activity, and exhibit the PM phenotype. Expression of CYP2D6 has been shown to be polymorphic with up to now more than 80 genetic variants detected for the encoding gene, with more than 15 encoding for inactive enzyme. Probe drugs for CYP2D6 phenotyping are dextromethorphan, debrisoquin, sparteine, and metoprolol. For safety concerns and drug availability, the preferred probe is dextromethorphan (DM) (Schmid et al. 1985).

Procedure

Phenotyping: The method measures DM and its O-demethylated metabolite, dextrorphan (DX), which is formed by CYP2D6. DM and DX, and other metabolites, are excreted in urine, mainly as glucuronide conjugates.

Subjects receive a single oral 10–30 mg DM (generally hydrobromide salt syrup) dose. Urine is collected from drug intake to 8 h post-dosing. Other collection times (0–6, 0–10, 0–12, or 0–24 h) can be used, but short collection intervals might lead to increased intra-subject variability.

Urine is first hydrolyzed with β-glucuronidase. Then, different methods can be used involving DM and DX extraction, followed either by HPLC and fluorescence detection (Chladek et al.

1999; Hoskins et al. 1997) or capillary gas chromatography (Wu et al. 2003).

Genotyping: The incidence of alleles coding for inactive enzymes varies between populations: three "population specific" alleles are CYP2D6*4 in Caucasians, *10 in Asians, and *17 in Africans (Bertilsson et al. 2002). CYP2D6*3, *4, *5, *6 are the main inactive alleles producing the PM phenotype in Caucasians, with CYP2D6*4 most commonly associated with the PM phenotype. By far, the most frequent null allele – not encoding a functional protein product – is CYP2D6*4 with a frequency of 20–25% in Caucasians (Zanger et al. 2004). The frequency of the *17 allele – associated with decreased enzyme activity – is high in Black Africans and in Black Americans, but practically absent in Caucasian populations (Bapiro et al. 2002; Gaedigk et al. 2002; Zanger et al. 2003). Four potential subgroups – ultrarapid metabolizers (UM), extensive metabolizers (EM), intermediate metabolizers (IM), and poor metabolizers (PM) – have been defined based on the genotype–phenotype relationships.

In Caucasian subjects, it has been recommended for "routine test" to genotype for alleles *1, *3, *4, *5, *6 that allow to detect 86–100% of PM (Sachse et al. 1997). To assign correct phenotype in nearly 100% subjects, *9 and *10 variants should also be determined.

Evaluation

Subjects with a DM/DX MR >0.3 are PM. Subjects with DM/DX <0.03 are EM. Those with 0.03 < MR < 0.3 are IM.

No difference or slightly higher CYP2D6 activity in females has been found when comparing to male subjects (Hägg et al. 2001; McCune et al. 2001).

Relationship between phenotyping and genotyping is investigated by plotting log MR versus CYP2D6 allele combinations (Chou et al. 2003).

Critical Assessment of the Method

The method is widely used due to easy and safe administration. High intrasubject variability limits the test for discriminating between EM and UMs (Zanger et al. 2004).

The method is not appropriate in patients with renal impairment, due to reduced renal excretion of DM glucuronide metabolites. Sparteine has been recommended as a probe for this population and to discriminate between the four phenotypes UM, EM, IM, and PM. The DM/DX MR does not allow for consistent differentiation between CYP2D6 EM with one or two active alleles.

Modifications of the Method

Assays have been developed to determine DM and DX in plasma or saliva (Bolden et al. 2002; Hu et al. 1998; Chladek et al. 2000; Härtter et al. 1996). The use of saliva or plasma for CYP2D6 phenotyping has been developed for subject convenience, or for the development of single point methods to be easily incorporated in the "cocktail methods." Good correlation between MRs calculated from plasma, saliva samples and those obtained from urine has been observed.

References and Further Reading

Bapiro TE, Hasler JA, Ridderström M, Masimirembwa CM (2002) The molecular and enzyme kinetic basis for the diminished activity of the cytochrome P450 2D6*17 variant. Biochem Pharmacol 64: 1387–1398

Bertilsson L, Dahl ML, Dalen P, Al-Shurbaji A (2002) Molecular genetics of CYP2D6: clinical relevance with focus on psychotropic drugs. Br J Clin Pharmacol 53:111–122

Bolden RD, Hoke SH II, Eichhold TH et al (2002) Semi-automated liquid–liquid back extraction in a 96-well format to decrease sample preparation time for the determination of dextromethorphan and dextrorphan in human plasma. J Chromatogr B 772:1–10

Chladek J, Zimova G, Martinkova J, Tuma I (1999) Intra-individual variability and influence of urine collection period on dextromethorphan metabolic ratios in healthy subjects. Fundam Clin Pharmacol 13:508–515

Chladek J, Zimova G, Beranek M, Martinkova J (2000) In-vivo indices of CYP2D6 activity: comparison of dextromethorphan metabolic

ratios in 4-h urine and 3-h plasma. Eur J Clin Pharmacol 56:651–657
Chou WH, Yan FX, Robbins-Weilert DK et al (2003) Comparison of two CYP2D6 genotyping methods and assessment of genotype–phenotype relationships. Clin Chem 49:542–551
Gaedigk A, Bradford LD, Marcucci KA, Leeder JS (2002) Unique CYP2D6 activity distribution and genotype–phenotype discordance in black Americans. Clin Pharmacol Ther 72:76–89
Hägg S, Spigset O, Dahlqvist R (2001) Influence of gender and oral contraceptives on CYP2D6 and CYP2C19 activity in healthy volunteers. Br J Clin Pharmacol 51:169–173
Härtter S, Baier D, Dingemanse J et al (1996) Automated determination of dextromethorphan and its main metabolites in human plasma by high-performance liquid chromatography and column switching. Ther Drug Monit 18:297–303
Hoskins JM, Shenfield GM, Gross AS (1997) Modified high-performance liquid chromatographic method to measure both dextromethorphan and proguanil for oxidative phenotyping. J Chromatogr B 696:81–87
Hu OY, Tang HS, Chang WH, Hu TM (1998) Novel single-point plasma or saliva dextromethorphan method for determining CYP2D6 activity. J Pharmacol Exp Ther 285:955–960
Johansson I, Oscarson M, Yue QY et al (1994) Genetic analysis of the Chinese cytochrome P450D locus: characterization of variant CYP2D6 genes present in subjects with diminished capacity for debrisoquin hydroxylation. Mol Pharmacol 46:452–459
McCune J, Lindley C, Decker JL et al (2001) Lack of gender differences and large intrasubject variability in cytochrome P450 activity measured by phenotyping with dextromethorphan. J Clin Pharmacol 41: 723–731
Sachse C, Brockmöller J, Bauer S, Roots I (1997) Cytochromes P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. Am J Hum Genet 60:284–295
Schmid B, Bircher J, Preisig R, Küpfer A (1985) Polymorphic dextromethorphan metabolism: co-segregation of oxidative O-demethylation with debrisoquin hydroxylation. Clin Pharmacol Ther 38:618–624
Wu YJ, Cheng YY, Zeng S, Ma MM (2003) Determination of dextromethorphan and its metabolite dextrorphan in human urine by capillary gas chromatography without derivatization. J Chromatogr B 784:219–224
Zanger UM, Raimindo S, Eichelbaum M (2004) Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. Naunyn-Schmiedeberg's Arch Pharmacol 369:23–37

CYP3A

Purpose and Rationale

CYP3A is the predominant P450 subfamily (CYP3A4, CYP3A5, CYP3A7, and CYP3A43) in the human liver, and contributes significantly to the metabolism of many (at least 50%) drugs in numerous therapeutic classes. CYP3A4 is the major CYP present notably and predominantly in the liver and the small intestine, and interindividual variability in the level of its expression is very high – 20-fold or more (Shimada et al. 1994). CYP3A5 shares rather similar tissue distribution with CYP3A4, but is preferentially expressed in the lung. It represents generally a few percentage of total CYP3A as compared to CYP3A4 (exceptions are esophagus and prostate, specific for CYP3A5, and kidney in which CYP3A5 is predominantly expressed). CYP3A4 and CYP3A5 exhibit overlapping substrate specificity, and there is currently no specific CYP3A5 probe drug. CYP3A7 is primarily the major fetal CYP3A enzyme.

Most of drugs biotransformed with CYP3A are also P-glycoprotein substrates (noticeable exceptions are midazolam and nifedipine). CYP3A and P-glycoprotein contribute substantially to the first-pass elimination of highly cleared CYP3A substrates when orally

administered. However, CYP3A4 and P-glycoprotein activities are not coordinately regulated in the liver and in the intestine (von Richter et al. 2004).

Currently, 40 and 24 alleles have been identified for CYP3A4 and CYP3A5, respectively. Expression of CYP3A5 varies greatly among individuals (Lamba et al. 2002).

Due to multiple confounding factors, such as those involved in endogenous expression of CYP3A regulatory factors, numerous exogenous factors (environment, diet), the interplay between CYP3A and transporters in regulating drug disposition, the establishment of consistent relationships between CYP3A genotype and phenotype is actually a challenge (Wilkinson 2004). Currently, the value of CYP3A genotyping in drug development is far from being clinically useful.

The most used and validated probe drugs for CYP3A phenotyping are midazolam and ^{14}C–erythromycin (Watkins 1994). Alfentanyl, alprazolam, dapsone, DM, lidocaine, nifedipine, omeprazole, quinine, and verapamil have also been used but less frequently, and CYP3A specificity for some of them has been questioned. The "endogenous" 6β-hydroxycortisol test (measurement of 6β-hydroxycortisol: cortisol ratio in urine) is only useful for detecting CYP3A induction, and may be influenced by renal CYP3A activity.

Due to intraindividual differences in the liver and the intestinal CYP3A activity, phenotyping test results are related to the probe drug route of administration.

Procedure

Phenotyping

Midazolam test: Midazolam is primarily metabolized to 1′-hydroxymidazolam by CYP3A. It is rapidly and completely absorbed after oral administration (Gorski et al. 1998). It is the probe of choice to assess intestinal and hepatic or hepatic CYP3A activities only, after oral (Thummel et al. 1996) or intravenous administration, respectively.

Oral test doses are 2, 5, or 7.5 mg (as a solution). IV doses are 0.015, 0.025, or 0.05 mg/kg, or 1 or 2 mg per subject, as a 2–30-min infusion.

Blood samples are collected over a 6-h period. Numerous GC, GC/MS, HPLC/UV, or LC/MS methods have been developed for plasma midazolam assay (Lepper et al. 2004; Frison et al. 2001).

^{14} C-erythromycin breath test or ERMBT: CYP3A4 catalyzes the N-desmethylation of [^{14}C N-methyl] erythromycin. The test consists of the measurement of a single breath expired $^{14}CO_2$ collection obtained at 20 min following the IV administration of a 0.03 mg dose of ^{14}C–erythromycin (2–4 μCi administered) (Watkins 1994). This test is used for assessing hepatic CYP3A activity.

Genotyping: Allelic CYP3A4 gene variants are rare. No impact of the presence of the most common CYP3A4*1B mutation (with a frequency ranging from 0% in Chinese and Japanese to 45% in Afro-Americans) on midazolam, erythromycin, or nifedipine clearance has been evidenced. Most significant mutations are observed for CYP3A5 and CYP3A7. Further information on polymorphic expression of CYP3A5 and CYP3A7 can be found in the review by Lamba et al. (2002).

Evaluation

A complete pharmacokinetic profile is required to assess midazolam clearance, and is therefore more invasive than the ERMBT; however, the latter requires specific logistics for radiolabeled material use. The midazolam or ERMBT phenotype tests are used for dose individualizing of narrow therapeutic index CYP3A-metabolized drugs such as anticancer agents. The ratio 1′-hydroxymidazolam/midazolam has generally been found not useful for phenotyping.

Within a population of similar demographic and health characteristics, a four- to sixfold range in the metabolic clearance of a CYP3A-

drug substrate is usual, with common individual outliers exhibiting high or low activity (Lamba et al. 2002).

Critical Assessment of the Method

Midazolam clearance has been found to correlate with hepatic CYP3A levels (Thummel et al. 1994) as well as ERMBT results (Lown et al. 1992). However weak, inconstant, or lack of correlations between midazolam and ERMBT test results have been observed, which could be explained by binding to different CYP3A active sites. In addition, contrary to the midazolam test, the ERMBT does not capture CYP3A5 activity.

An ethnic difference – that could be drug-specific – in CYP3A4 activity has been observed for few CYP3A4 substrates (alprazolam and nifedipine), with a lower clearance in Asians than in Caucasians (Xie et al. 2001).

CYP3A4 and CYP3A5 genotyping tests could not explain sufficiently the interindividual variability observed in midazolam pharmacokinetics (Eap et al. 2004a).

Modifications of the Method

The combined use of IV midazolam and oral ^{15}N-midazolam or of the ERMBT and oral midazolam tests have been proposed to assess simultaneously the contributions of liver and intestine in CYP3A activity (Gorski et al. 1998; McCrea et al. 1999). The administration of orally given midazolam followed by an intravenous administration has also been validated (Lee et al. 2002c). A low oral 75 μg oral dose has recently been proposed, but needs large-scale validation (Eap et al. 2004b).

Modifications of the ERMBT have been described to improve its predictability in drug clearance estimations in cancer patients (Rivory et al. 2000).

A single blood sample for midazolam assay at 4 h post-dose has been reported as good estimator for IV or oral midazolam clearance determination (Lin et al. 2001).

References and Further Reading

Eap CB, Buclin T, Hustert E et al (2004a) Pharmacokinetics of midazolam in CYP3A4- and CYP3A5-genotyped subjects. Eur J Clin Pharmacol 60:231–236

Eap CB, Buclin T, Cucchia G et al (2004b) Oral administration of midazolam (75 μg) as an in vivo probe for CYP3A activity. Eur J Clin Pharmacol 60:237–246

Frison G, Tedeschi L, Maietti S, Ferrara SD (2001) Determination of midazolam in plasma by solid-phase microextraction and gas chromatography/mass spectrometry. Rapid Commun Mass Spectrom 15:2497–2501

Gorski JC, Jones DR, Haehner-Daniels BD et al (1998) The contribution of intestinal and hepatic CYP3A to the interaction between midazolam and clarithromycin. Clin Pharmacol Ther 64:133–143

Lamba JK, Lin YS, Schuety EG, Thummel KE (2002) Genetic contribution to variable human CYP3A-mediated metabolism. Adv Drug Deliv Rev. 54:1271–1294

Lee JI, Chaves-Gnecco D, Amico JA et al (2002c) Application of simultaneous midazolam administration for hepatic and intestinal cytochrome P450 3A phenotyping. Clin Pharmacol Ther 72:718–728

Lepper ER, Hicks JK, Verweij J et al (2004) Determination of midazolam in human plasma by liquid chromatography with mass-spectrometric detection. J Chromatogr B 806:305–310

Lin YS, Lockwood GF, Graham MA et al (2001) In-vivo phenotyping for CYP3A by single-point determination of midazolam plasma concentration. Pharmacogenetics 11:781–791

Lown K, Kolars J, Turgeon K et al (1992) The erythromycin breath test selectively measures P450IIIA in patients with severe liver disease. Clin Pharmacol Ther 52:229–238

McCrea J, Prueksaritanont T, Gertz BJ et al (1999) Concurrent administration of the erythromycin breath test (EBT) and oral midazolam as in vivo probes for CYP3A activity. J Clin Pharmacol 39:1212–1220

von Richter O, Burk O, Fromm MF et al (2004) Cytochrome P450 3A4 and P-glycoprotein expression in human small intestinal enterocytes and hepatocytes: a comparative analysis in paired tissue specimens. Clin Pharmacol Ther 75:172–183

Rivory LP, Slaviero KA, Seale JP et al (2000) Optimizing the erythromycin breath test for use in cancer patients. Clin Cancer Res 6:3480–3485

Shimada Y, Yamazki H, Mimura M et al (1994) Inter-individual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther 270:413–423

Thummel KE, Shen DD, Podoll TD et al (1994) Use of midazolam as a human cyotochrome P450 3A probe: in vitro–in vivo correlations in liver transplant patients. J Pharmacol Exp Ther 271:549–556

Thummel KE, O'Shea D, Paine MF et al (1996) Oral first-pass elimination of midazolam involves both gastrointestinal and hepatic CYP3A-mediated metabolism. Clin Pharmacol Ther 59:491–502

Watkins PB (1994) Noninvasive tests of CYP3A enzymes. Pharmacogenetics 4:171–184

Wilkinson GR (2004) Genetic variability in cytochrome P450 3A5 and in vivo cytochrome P450 3A activity: some answers but still questions. Clin Pharmacol Ther 76:99–103

Xie HG, Kim RB, Wood AJ, Stein CM (2001b) Molecular basis of ethnic differences in drug disposition and response. Annu Rev. Pharmacol Toxicol 41:815–850

Other CYPS

This section summarizes succinctly the current knowledge on some other CYPs, its role in drug metabolism and its genetic impact have been more recently investigated as compared to other CYPs.

CYP2A6

CYP2A6, primarily expressed in the liver, is the major CYP (the sole at usual low concentrations) involved in nicotine oxidation, and is also involved in the metabolism of carcinogen or procarcinogen compounds (such as nitrosamines and aflatoxins). A couple of drugs is metabolized by CYP2A6: chlormethiazole, coumarin, disulfiram, halothane, valproic acid, and others (Oscarson 2001). CYP2A6 PM is less than 1% in Caucasians but up to 20% in Orientals (Oscarson 2001; Raunio et al. 2001; Xu et al. 2002). The most "in vivo deficient" alleles for PM status are CYP2A6*2 and CYP2A6*4, rather common in Orientals (15% in Chinese, 20% in Japanese). The important role of CYP2A6 in nicotine metabolism was shown in an epidemiological study, revealing that the CYP2A6 genotype was a major determinant for smoking behavior and susceptibility to tobacco-related lung cancer (Fujieda et al. 2004).

Phenotyping has been performed in some countries with coumarin (not available in all countries), despite some limitations with data accuracy obtained with the analytical methods used (Pelkonen et al. 2000; Cok et al. 2001). The test assesses the amount of 7-hydroxycoumarine (free and conjugated) in urine after ingestion of 2–5 mg coumarine by the subjects. Nicotine has also been used as the probe drug for CYP2A6 in vivo activity testing. Recent investigations using pilocarpine as probe demonstrated that PM status was associated with two inactive CYP2A6 alleles, CYP2A6*4A, CYP2A6*7, CYP2A6*9, or CYP2A6*10 (Endo et al. 2008).

References and Further Reading

Cok I, Kocabas NA, Cholerton S et al (2001) Determination of coumarin metabolism in Turkish population. Human Exp Pharmacol 20: 179–184

Endo T, Nakajima M, Fukami T et al (2008) Genetic polymorphisms of CYP2A6 affect the in-vivo pharmacokinetics of pilocarpine. Pharmacogenet Genomics 18:761–772

Fujieda M, Yamazaki H, Saito T et al (2004) Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behaviour and tobacco-related lung cancer risk in male Japanese smokers. Carcinogenesis 25:2451–2458

Oscarson M (2001) Genetic polymorphisms in the cytochrome P450 2A6 (CYP2A6) gene: implications for interindividual differences in nicotine metabolism. Drug Metab Dispos 29:91–95

Pelkonen O, Rautio A, Raunio H, Pasanene M (2000) CYP2A6: a human coumarin 7-hydroxylase. Toxicology 144:139–147

Raunio H, Rautio A, Gullstén H, Pelkonen O (2001) Polymorphisms of CYP2A6 and its practical consequences. Br J Clin Pharmacol 52: 357–363

Xu C, Goodz S, Sellers E, Tyndale RF (2002) CYP2A6 genetic variation and potential consequences. Adv Drug Deliv Rev. 54:1245–1256

CYP2B6

CYP2B6 has been estimated to represent 1–10% of the total hepatic CYP content. It catalyzes bupropion hydroxylation, S-mephenytoin N-demethylation, and is involved in the metabolism of cyclophosphamide, ifosfamide, mianserin, efavirenz, artemisinin, and propofol (Turpeinen et al. 2006). CYP2B6*6 has been associated with reduced bupropion clearance in vitro (Hesse et al. 2004), but not in vivo whereas a moderate clearance increase was observed with CYP2B6*4 (Kirchheiner et al. 2003b). Multiple gene polymorphisms have resulted in phenotypic null alleles (Lang et al. 2004). Pharmacokinetics of the anti-HIV drug efavirenz has been associated with CYP2B6–G516 T polymorphism (Saitoh et al. 2007).

Bupropion (150 mg dose) has been proposed for phenotyping, but it is recommended to administer body weight-adjusted doses (Faucette et al. 2000). Efavirenz may also be a valuable probe for CYP2B6 (Ward et al. 2003).

References and Further Reading

Faucette SR, Hawke L, Lecluyse RL et al (2000) Validation of bupropion hydroxylation as a selective marker of human cytochrome P450 2B6 catalytic activity. Drug Metab Dispos 28:1222–1230

Hesse LM, He P, Krishnaswamy S et al (2004) Pharmacogenetic determinants of interindividual variability in bupropion hydroxylation by cytochrome P450 2B6 in human liver microsomes. Pharmacogenetics 14:225–238

Kirchheiner J, Klein C, Meineke I et al (2003b) Bupropion and 4-OH-bupropion pharmacokinetics in relation to genetic polymorphisms in CYP2B6. Pharmacogenetics 13:619–626

Lang T, Klein K, Richter T et al (2004) Multiple novel nonsynonymous CYP2B6 gene polymorphisms in Caucasians: demonstration of phenotypic null alleles. JPET 311:34–43

Saitoh A, Fletcher C, Brundage R et al (2007) Efavirenz pharmacokinetics in HIV-1-infected children are associated with CYP2B6–G516 T polymorphism. J Acquir Immune Defic Syndr 45:280–285

Turpeinen M, Raunio H, Pelkonen O (2006) The functional role of CYP2B6 in human drug metabolism: substrates and inhibitors in vitro, in vivo and in silico. Curr Drug Metabol 7:705–714

Ward B, Gorski C, Jones D et al (2003) The cytochrome P450 2B6 (CYP2B6) is the main catalyst of efavirenz primary and secondary metabolism: implication for HIV/AIDS therapy and utility of efavirenz as a substrate marker of CYP2B6 catalytic activity. JPET 306:287–300

CYP2C8

CYP2C8 is involved in the metabolism of arachidonic acid, all-trans retinoic acid, paclitaxel, amiodarone, amodiaquine, repaglinide, rosiglitazone, torsemide, troglitazone, and zopiclone. Most of these drugs are also metabolized by CYP3A4. Recently, the potential

contribution of CYP2C8 to the metabolism of NSAIDs in addition to the well-known CYP2C9 role has been highlighted for ibuprofen (Garcia-Martin et al. 2004). The CYP2C8*3 allele (present in 13% and 2% of Caucasians and Afro-American subjects, respectively) has been shown in vitro deficient for paclitaxel and arachidonic acid metabolism (Dai et al. 2001; Bahadur et al. 2002). For the antidiabetic repaglinide, unexpected in vivo lower exposure was observed in subjects with CYP2C8*1/*3 genotype, without any pharmacological consequences (Niemi et al. 2003). For ibuprofen, reduced clearance of the R (−) enantiomer was related to CYP2C8*3 allele, and reduced clearance of the S(+) enantiomer was influenced by CYP2C8*3 and CYP2C9*3 alleles. In subjects homozygous or double heterozygous for these variants (8% of 130 subjects evaluated), the clearances of ibuprofen were only 7–27% of the clearances observed in subjects with no CYP mutations. A strong association between CYP2C8*3 and CYP2C9*2 occurrence has been characterized in a large Swedish population, highlighting linkage between CYP2C8 and CYP2C9 polymorphisms (Yasar et al. 2002c).

Further in vitro/in vivo investigations are needed to assess the relationship between CYP2C8 (and CYP2C9) polymorphisms and drug metabolic clearance, in order to address the clinical relevance of CYP2C8 genotyping.

References and Further Reading

Bahadur N, Leathart JBS, Mutch E et al (2002) CYP2C8 polymorphisms in Caucasians and their relationship with paclitaxel 6-α-hydroxylase activity in human liver microsomes. Biochem Pharmacol 64: 1579–1589

Dai D, Zeldin DC, Blaisdell JA et al (2001) Polymorphisms in CYP2C8 decrease metabolism of the anticancer drug paclitaxel and arachidonic acid. Pharmacogenetics 11:597–607

Garcia-Martin E, Martinez C, Tabarés B et al (2004) Interindividual variability in ibuprofen pharmacokinetics is related to interaction of cytochrome P450 2C8 and 2C9 amino acid polymorphisms. Clin Pharmacol Ther 76:119–127

Niemi M, Leathart JB, Neuvonen M et al (2003) Polymorphisms in CYP2C8 is associated with reduced plasma concentrations of repaglinide. Clin Pharmacol Ther 74:380–387

Yasar U, Lundgren S, Eliasson E et al (2002c) Linkage between CYP2C8 and CYP2C9 genetic polymorphisms. Biochem Biophys Res Commun 299:25–28

CYP2E1

CYP2E1, an ethanol-inducible CYP, activates some procarcinogens such as nitrosamines, is involved in the metabolism of endogenous substrates (steroids and bile acids), alcohols, xanthines, volatile chemicals (toluene, benzene, and halocarbons), but of few drugs (chlorzoxazone, etoposide, dapsone, and high-dose acetaminophen) (Lieber 1997). Seven alleles, 13 genetic mutations have been described, but no genotyping–phenotyping relationships have been well established to date. Based on safe use and CYP selectivity (though CYP1A1, CYP1A2 have been found involved in its biotransformation in vitro), chlorzoxazone is the only in vivo probe drug to phenotype CYP2E1 activity, toward assessment of its 6-hydroxylation (Ono et al. 1995; Lucas et al. 1999; Ernstgard et al. 2004). Due to dose-dependent metabolism, the dose should be preferably administered on a mg/kg basis (10 mg/kg rather than the common 250 or 500 mg doses). Relatively low intraindividual variability in chlorzoxazone metabolism has been observed. Measurement can be done in urine or in plasma, after enzymatic hydrolysis of 6-chlorzoxazone glucuronide, using HPLC and UV detection or LC/MS/MS methods (Frye and Stiff 1996; Frye et al. 1998; Scoot et al. 1999). The use of plasma metabolite ratio determined with only one plasma sample – at 2 h post-dosing – has been recently validated.

References and Further Reading

Ernstgard L, Warholm M, Johanson G (2004) Robustness of chlorzoxazone as an in vivo measure of cytochrome P450 2E1 activity. Br J Clin Pharmacol 58:190–200

Frye RF, Stiff DD (1996) Determination of chlorzoxazone and 6-hydroxy chlorzoxazone in human plasma and urine by high performance liquid chromatography. J Chromatogr B 686:291–296

Frye RF, Adedoyin A, Mauro K et al (1998) Use of chlorzoxazone as an in vivo probe of cytochrome P450 2E1: choice of dose and phenotypic trait measure. J Clin Pharmacol 38:82–89

Kim RB, O'Shea D, Wilkinson GR (1995) Interindividual variability of chlorzoxazone 6-hydroxylation in men and women and its relationship to CYP2E1 genetic polymorphisms. Clin Pharmacol Ther 57:645–655

Lieber CS (1997) Cytochrome P-450 2E1: its physiological and pathological role. Physiol Rev. 77:517–544

Lucas D, Ferrara R, Gonzalez E et al (1999) Chlorzoxazone: a selective probe for phenotyping CYP2E1 in humans. Pharmacogenetics 9:377–388

Ono S, Hatanaka T, Hotta H et al (1995) Chlorzoxazone is metabolized by human CYP1A2 as well as by human CYP2E1. Pharmacogenetics 5:143–150

Scoot RJ, Palmer J, Lewis IA, Pleasance S (1999) Determination of a "GW cocktail" of cytochrome P450 probe substrates and their metabolites in plasma and urine using automated solid phase extraction and fast gradient liquid chromatography tandem mass spectrometry. Rapid Commun Mass Spectrom 13:2305–2319

Phase II Enzymes

With the exception of N-acetyltransferases (detailed below), there are few deficiencies in Phase II drug metabolism enzymes that have resulted in clinically significant effects. Each Phase II enzyme class is most often a superfamily of enzymes, and usually there is large interindividual and interethnic variability in drug conjugations, and overlapping substrate specificity exists for numerous isoenzymes. Despite the crucial role of conjugation enzymes in xenobiotic metabolism, the functional significance of enzyme polymorphism is only known for few substrates. Therefore, with the exception of the caffeine and thiopurine methyltransferase (TPMT) tests (see below), no probe test drug has been yet investigated for in vivo phenotyping and validated to assess phenotyping–genotyping relationships. Nevertheless, some important aspects of enzyme polymorphism on the pharmacokinetics of drugs with narrow therapeutic index are summarized below.

N-Acetyltransferases

Purpose and Rationale

N-acetyltransferases type I (NAT1) and type II (NAT2) catalyze N- and O-acetylation reactions involved in the metabolism of drugs containing arylamino, hydroxyl, sulfhydryl groups and hydrazine structure, and also in environmental carcinogens (such as those present in tobacco smoke, or in diet such as charcoal-broiled food) (Weber and Hein 1985). Pending on the drug, and on the interplay between CYPs and N-acetylases (and other Phase II conjugation enzymes) in xenobiotic metabolism, the impact of subject status "poor acetylator" or "rapid acetylator" on drug activity and/or toxicity may vary, and then is drug specific. NAT1 and NAT2 exhibit a high degree (81%) of amino-acid sequence homology, and share common substrates (Meisel 2002) but coding genes loci are regulated independently. Main NAT2 drug substrates are isoniazid, sulfonamides, procainamide, hydralazine, acebutolol, aminoglutethimide, and dapsone.

Para-aminosalicylic and para-aminobenzoic acids are considered specific substrates for human NAT1, and sulfamethazine, isoniazid, procainamide, and dapsone are considered specific substrates for human NAT2 (Butcher et al. 2002). NAT1 is considered as ubiquitously distributed in the body, whereas NAT2 is expressed in liver and intestinal mucosa.

Polymorphic N-acetylation was first described for isoniazid in the 1950s and is the first example of interindividual pharmacogenetic variability. Until

2007, about 30 and more than 50 variant alleles have been described for NAT1 and NAT2, respectively. At http://N-acetyltransferasenomenclature.louisville.edu overviews on the NAT alleles can be found. The presence of some NAT1 variants, as well as NAT2 variants, has been linked to increased susceptibility to some cancers (notably bladder and colon cancers), and NAT2 polymorphism associated with some drug-induced diseases such as lupus erythematosus (hydralazine and procainamide), Stevens–Johnson or Lyell syndromes (sulfonamides).

Significant interethnic and geographic differences in NAT2 activity have been found. Slow acetylators represent 40–70% Caucasians and 10–20% Asians. High acetylation capacity has been reported in 5% Caucasians (Meyer and Zanger 1997).

Probe drugs for NAT1 phenotyping is PAS, and for NAT2 phenotyping are caffeine, sulfamethazine, procainamide, isoniazid, and dapsone. In vivo testing for NAT2 has been proved useful for drug monitoring to avoid potential side effects generally observed in slow metabolizers (the exception was the anticancer agent amonafide, with myelotoxicity observed in rapid acetylators). The most used test to identify rapid and slow acetylators is the caffeine test, which is described thereafter, though the N-acetylation step takes place after the N-desmethylation of caffeine by CYP1A2 followed by the biotransformation into an unstable intermediate.

Procedure

Phenotyping: Caffeine is metabolized by CYP1A2, NAT2, and xanthine oxidases. The methods could involve the measurement of 5-acetyl-formylamino-3-methyluracil (AFMU), 5-acetyl-amino-3-methyluracil (AAMU, degradation product of AFMU), 1-methyl-xanthine (1MX), and 1-methyluric acid (1MU) in 0–8, 0–12, 0–24 h urine of subjects orally given 200 mg or 2–3 mg/kg caffeine after a xanthine-free regimen. The common MR used is AFMU/1MX, but the AFMU/(AFMU +1MX + 1MU) is more discriminating (Relling et al. 1992; Rostami 1995) and should be used when xanthine-oxidase inhibitors may be present (Fuchs et al. 1999). Other ratios such as AFMU/(1MX + 1MU), or AAMU/1MX, AAMU/(AAMU +1MX + 1MU) have been validated (Tang et al. 1991; Nyeki et al. 2002).

The most common methods to assay caffeine and its metabolite in urine used HPLC with UV detection (Grant et al. 1984; Krul and Hageman 1998b) or mass spectrometry (Baud-Camus et al. 2001).

Genotyping: Mutations of NAT2*5, NAT2*6, NAT2*7, NAT2*14, and NAT2*17 alleles are associated with a slow acetylation phenotype for homozygous subjects (Butcher et al. 2002).

There are large differences among ethnic groups regarding alleles' frequency. High frequency (>28%) of NAT2*5 alleles has been observed in Caucasians and Africans, and of NAT2*7 in Asians (>10%) and of NAT2*14 in Africans (>8%), this last one being <1% in Caucasians and Asians (Meyer and Zanger 1997).

Evaluation

Caffeine test: Subjects with a AFMU/1MX ratio < 0.55 or a AFMU/(AFMU +1MX + 1MU) ratio < 0.26 are slow acetylators (Fuchs et al. 1999). Higher activity has been observed in black as compared to white subjects (Relling et al. 1992), and a gender effect has generally not been observed (Kashuba et al. 1998).

Critical Assessment of the Method

Depending on the probe drug used and on the experimental method, 2 or 3 acetylator types can be described: slow, intermediate, and rapid; the intermediate one being not always distinguished from the rapid one. Phenotype distribution has been considered as a continuous variable (Meisel 2002). Due to slow postnatal maturation of the acetylation enzymatic systems, the acetylation status is evolving in newborns and infants, and depends on the probe drug used (Rane 1999).

Good relationships between genotyping and phenotyping tests have been reported (Meisel et al. 1997; Kita et al. 2001).

The urinary caffeine test is not based on assays of specific substrates and products of NAT2 ("including" other metabolism pathways involving at least xanthine-oxidases), and is affected by diet habits, xanthine-oxidase inhibitors such as

allopurinol (Fuchs et al. 1999), or other drugs (Klebovitch et al. 1995). NAT activities are affected by anti-inflammatory drugs. Of note, acetaminophen is an inhibitor of NAT2 in vivo (Rothen et al. 1998).

Discordances between caffeine and dapsone phenotyping data, and between NAT2 phenotyping status and genotyping have been observed in acutely ill patients infected with HIV (O'Neil et al. 2000), which may be due partly to non-detection of rare NAT2 alleles (Alfirevic et al. 2003).

Modifications of the Method

Some recent references for other used NAT2 phenotyping tests can be found for dapsone in Alfirevic et al. (2003), O'Neil et al. (2000), Queiroz et al. (1997), for sulfamethazine in Hadasova et al. (1996) and Meisel et al. (1997), and for procainamide in Okumura et al. (1997) and Mongey et al. (1999).

References and Further Reading

Alfirevic A, Stalford AC, Vilar FJ et al (2003) Slow acetylator phenotype and genotype in HIV-positive patients with sulphamethoxazole hypersensitivity. Br J Clin Pharmacol 55:158–165

Baud-Camus F, Marquet P, Soursac M et al (2001) Determination of N-acetylation phenotype using caffeine as a metabolic probe and high-performance liquid chromatography with either ultraviolet detection or electrospray mass spectrometry. J Chromatogr B 760:55–63

Butcher NJ, Boukouvala S, Sim E, Minchin RF (2002) Pharmacogenetics of the arylamine N-acetyltransferases. Pharmacogenomics J 2:30–42

Fuchs P, Haefeli WE, Ledermann HR, Wenk M (1999) Xanthine oxidase inhibition by allopurinol affects the reliability of urinary caffeine metabolic ratios as markers for N-acetyltransferase2 and CYP1A2 activities. Eur J Clin Pharmacol 54:869–876

Grant DM, Tang BK, Kalow W (1984) A simple test for acetylator phenotype using caffeine test. Br J Clin Pharmacol 17:459–464

Hadasova E, Franke G, Zschiesche M et al (1996) Debrisoquin 4-hydroxylation and sulphamethazine N-acetylation in patients with schizophrenia and major depression. Br J Clin Pharmacol 41: 428–431

Kashuba ADM, Bertino JS, Kearns GL et al (1998) Quantitation of three-month intraindividual variability and influence of sex and menstrual cycle phase on CYP1A2, N-acetyltransferase-2, and xanthine oxidase activity determined with caffeine phenotyping. Clin Pharmacol Ther 63:540–551

Kita T, Tanigawara Y, Chikizawa S et al (2001) N-acetyltransferase2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients. Biol Pharm Bull 24:544–549

Klebovitch I, Rautio A, Salonpää et al (1995) Antipyrine, coumarin and glipizide affect N-acetylation measured by caffeine test. Biomed Pharmacother 49:225–227

Krul C, Hageman G (1998b) Analysis of urinary caffeine metabolites to assess biotransformation enzyme activities by reverse-phase high-performance liquid chromatography. J Chromatogr B 709:27–34

Meisel P (2002) Arylamine N-acetyltransferases and drug response. Pharmacogenomics 3:349–366

Meisel P, Schroeder C, Wulff K, Siegmund W (1997) Relationship between human genotype and phenotype of N-acetyltransferase (NAT2) as estimated by discriminant analysis and multiple linear regression: 1. Genotype and N-acetylation in vivo. Pharmacogenetics 7:241–246

Meyer UA, Zanger UM (1997) Molecular mechanisms of genetic polymorphisms of drug metabolism. Annu Rev. Pharmacol Toxicol 37:269–296

Mongey AB, Sim E, Risch A, Hess E (1999) Acetylation status is associated with serological change but not clinically significant disease in patients receiving procainamide. J Rheumatol 26:1721–1726

Nyeki A, Buclin T, Biollaz J, Decosterd LA (2002) NAT2 and CYP1A2 phenotyping with caffeine: head-to-head comparison of AFMU

vs AAMU in the urine metabolite ratios. Br J Clin Pharmacol 55:62–67
Okumura K, Kita T, Chikazawa S et al (1997) Genotyping of N-acetylation polymorphism and correlation with procainamide metabolism. Clin Pharm Ther 61:509–517
O'Neil WM, Drobitch RK, MacArthur RD et al (2000) Acetylator phenotype and genotype in patients infected with HIV: discordance between methods for phenotype determination and genotype. Pharmacogenetics 10:171–182
Queiroz RH, Dreossi SA, Carvalho D (1997) A rapid, specific, and sensitive method for the determination of acetylation phenotype using dapsone. J Anal Toxicol 21:203–207
Rane A (1999) Phenotyping of drug metabolism in infants and children: potentials and problems. Pediatrics 104:640–643
Relling MV, Lin JS, Ayers GD, Evans EE (1992) Racial and gender differences in N-acetyltransferase, xanthine oxidase and CYP1A2* activities. Clin Pharmacol Ther 52:643–658
Rostami-Hodjegan A, Nurminen S, Jackson PR, Tucker GT (1996b) Caffeine urinary metabolic ratios as markers of enzyme activity: a theoretical assessment. Pharmacogenetics 6:121–149
Rothen JP, Haefeli WE, Meyer UA et al (1998) Acetominophen is an inhibitor of hepatic N-acetyltransferase 2 in vitro and in vivo. Pharmacogenetics 8:553–559
Tang BK, Kadar D, Qian L et al (1991) Caffeine as a metabolic probe: validation of its use for acetylator phenotyping. Clin Pharmacol Ther 49:648–657
Weber WW, Hein DW (1985) N-acetylation pharmacogenetics. Pharmacol Rev. 37:25–79

Uridine Diphosphate Glucuronosyltransferases

Glucuronidation is a potent detoxification pathway. The uridine diphosphate glucuronosyltransferases (UGTs) are involved in the biotransformation of endogenous substances (bilirubin, biliary acids, and steroid hormones) and numerous drugs and carcinogens. Currently, 20 functional UGTs have been characterized with activity mainly expressed in the liver and the GI tract. There are three subfamilies: UGT1A, UGT2A, and UGT2B, with distinct but broad overlapping substrate specificity existing for the different isoforms of each family. UGT1A1 is the most abundant UGT in the liver. Human diseases related to deficient UGT1A1 alleles are the well-characterized inherited unconjugated hyperbilirubinemias, including the Gilbert's syndrome that affects 6–12% of Caucasian subjects. Exhaustive reviews on roles, tissue patterns of expression, and pharmacogenomics of UGTs can be found in papers from Tukey and Strassburg (2000), Fischer et al. (2001), Guillemette (2003), and Wells et al. (2004).

A decreased clearance has been observed for some drugs metabolized by glucuronidation in patients with Gilbert's syndrome. A clinically significant impact of UGT polymorphism has to date is only demonstrated for some anticancer agents: clearly for irinotecan, and with contradictory results for flavopiridol (Zhai et al. 2003). UGT1A1 and UGT1A9 are involved in the glucuronidation of the active metabolite SN-38 of irinotecan. The presence of the deficient UGT1A1*28 variant (most frequent variant as compared to UGT1A9 variants) has been clinically linked to a decrease in SN-38 glucuronidation rate and to an increased occurrence of serious side effects, mainly severe diarrhea and neutropenia (Ando et al. 1998; Innocenti et al. 2004; Iyer et al. 2002; Paoluzzi et al. 2004). Variants of UGT1A7 were reported to affect SN-38 glucuronidation but only in vitro (Villeneuve et al. 2003). Other factors, such as polymorphism in drug transporter P-glycoprotein and renal excretion, may play a role in the complex disposition pattern of irinotecan.

References and Further Reading

Ando Y, Saka H, Asai G et al (1998) OGT1A1 genotypes and glucuronidation of SN-38, the active metabolite of irinotecan. Ann Oncol 9:845–847
Fischer MB, Paine MF, Strelevitz TJ, Wrighton SA (2001) The role of hepatic and extrahepatic

UDP–glucuronyltransferases in human drug metabolism. Drug Metab Rev. 33:273–297
Guillemette C (2003) Pharmacogenomics of human UDP–glucuronosyltransferase enzymes. Pharmacogenomics J 3:136–158
Innocenti F, Undvia SD, Iyer L et al (2004) Genetic variants in the UDP–glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. J Clin Oncol 8:1382–1388
Iyer L, Das S, Janish L et al (2002) UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. Pharmacogenomics J 2:43–47
Paoluzzi L, Singh AS, Price DK, Danesi R (2004) Influence of genetic variants in UGT1A1 and UGT1A9 in the in vivo glucuronidation of SN-38. J Clin Pharmacol 44:854–860
Tukey RH, Strassburg CP (2000) Human UDP–glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol 40:581–616
Villeneuve L, Girard H, Fortier LC et al (2003) Novel functional polymorphisms in the UGT1A7 and UGT1A9 glucuronidating enzymes in Caucasian and African-American subjects and their impact on the metabolism of 7-ethyl-10-hydroxycamptothecin and flavopiridol anticancer drugs. J Pharmacol Exp Ther 307:117–128
Wells PG et al (2004) Symposium article: glucuronidation and the UDP–glucuronosyltransferases in health and disease. Drug Metab Dispos 32:281–290
Zhai S, Sausville EA, Senderowicz AM et al (2003) Clinical pharmacology and pharmacogenetics of flavopiridol 1-h IV infusion in patients with refractory neoplasms. Anticancer Drugs 14:125–135

Methyltransferases

There are at least four enzymes catalyzing S-, N-, and O-methylation using S-adenosylmethionine, but only TPMT polymorphism has been found to have important clinical consequences. To date, no endogenous substrate of TPMT is known. TPMT is involved in the metabolism of mercaptopurine, azathioprine and thioguanine, narrow therapeutic index drugs in use for the treatment of patients with neoplasia or autoimmune disease, or of transplant recipients. About 0.3% of Caucasian subjects have no detectable enzyme activity and 10% intermediate activity (McLeod and Evans 2001). Four alleles TPMT*2, TPMT*3A, TPMT*3B, and TPMT*3C account for 80–95% of Caucasians with intermediate or low enzyme activities. Patients with low inherent TPMT activity are at great risk for severe potentially life-threatening myelosuppressive toxicity with treatment by the above-mentioned drugs, whereas subjects with very high activity might be underdosed (Zhou 2006). Patients with two nonfunctional variant TPMT alleles should receive 5–10% of drug standard doses. TPMT genotyping has proved its usefulness in individualizing mercaptopurine dose in patients, and can replace the phenotyping test: measurement of the erythrocyte enzyme activity, based on the in vitro conversion of 6-mercaptopurine to 6-methylmercaptopurine or 6-thioguanine to 6-methylthioguanine (Innocenti et al. 2000; Evans 2004). A cut-off concentration of 45.5 nmol thioguanine/gHb h^{-1} for this TPMT phenotyping test has been proposed for assessing the need of the genotyping test (Wusk et al. 2004).

References and Further Reading

Evans WE (2004) Pharmacogenetics of thiopurine S-methyltransferase and thiopurine therapy. Ther Drug Monit 26:186–191
Innocenti F, Iyer L, Ratain MJ (2000) Pharmacogenetics: a tool for individualizing antineoplastic therapy. Clin Pharmacokinet 39: 315–325
McLeod HL, Evans WE (2001) Pharmacogenomics: unlocking the human genome for better drug therapy. Annu Rev. Pharmacol Toxicol 41:101–121
Wusk B, Kullak-Ublick GA, Rammert C et al (2004) Thiopurine S-methyltransferase polymorphisms: efficient screening method for patients considering taking thiopurine drugs. Eur J Clin Pharmacol 60:5–10
Zhou S (2006) Clinical pharmacogenomics of thiopurine S-methyltransferase. Curr Clin Pharmacol 1:119–128

Glutathione S-transferases and Sulfotransferases

Glutathione and sulfatation conjugations are important pathways for generally detoxifying endogenous substrates and xenobiotics (Commandeur et al. 1995). However, some produced metabolites (i.e., mercapturic acids, O-sulfo conjugates) are toxic by different mechanisms, often by reaction with DNA and other cellular nucleophils.

Eight classes of glutathione- S-transferases (GSTs) have been described. The role of the glutathione pathway and the impact of enzyme polymorphism have been highlighted for detoxification and some disease susceptibility, and routine phenotyping of some GSTs exists for clinical safety measurement, but currently there is no evidence of genotyping or phenotyping usefulness for drug dosage adjustment (Hayes and Strange 2000; Tetlow et al. 2004). GSTs are involved in the detoxification of chemotherapeutics, including platinum derivatives. Polymorphisms in the GSTP1 genotype might become a powerful tool to predict oxaliplatin-induced cumulative neuropathy (Lecomte et al. 2006).

Soluble sulfotransferases are involved in the sulfonation of endogenous substrates (notably steroids, neurotransmitters, and eicosanoids) and numerous xenobiotics (i.e., acetaminophen, and organic-platin anticancer agents). The presence of some sulfotransferases variants could be associated with some cancer risk. Phenotyping tests have been developed for some forms (SULT1A and SULT1A3) by measuring platelet sulfotransferase activity (Glatt and Meinl 2004).

References and Further Reading

Commandeur JNM, Stintjes GJ, Vermeulen NPE (1995) Enzymes and transport systems involved in the formation and disposition of glutathione S-conjugates. Pharmacol Rev. 47:271–330

Hayes JD, Strange RC (2000) Glutathione S-transferase polymorphisms and their biological consequences. Pharmacology 61:154–166

Glatt H, Meinl W (2004) Pharmacogenetics of soluble sulfotransferases (SULTs). Naunyn-Schmiedeberg's Arch Pharmacol 369:55–68

Lecomte T, Landi B, Beaune P et al (2006) Glutathione S-transferase P1 polymorphism (Ile ^{105}Val) predicts cumulative neuropathy in patients receiving oxaliplatin-based chemotherapy. Clin Cancer Res 12:3050–3056

Tetlow N, Robinson A, Mantle T, Board P (2004) Polymorphism of human mu class glutathione transferases. Pharmacogenetics 14:359–368

References and Further Reading

Abernathy DR, Todd EL (1985) Impairment of caffeine clearance by chronic use of low-dose oestrogen-containing oral contraceptives. Eur J Clin Pharmacol 28:425–428

Alfirevic A, Stalford AC, Vilar FJ et al (2003) Slow acetylator phenotype and genotype in HIV-positive patients with sulphamethoxazole hypersensitivity. Br J Clin Pharmacol 55:158–165

Andersson T, Regardh CG, Dahl-Puustinen ML, Bertilsson L (1990) Slow omeprazole metabolizers are also poor S-mephenytoin hydroxylators. Ther Drug Monit 12:415–416

Ando Y, Saka H, Asai G et al (1998) OGT1A1 genotypes and glucuronidation of SN-38, the active metabolite of irinotecan. Ann Oncol 9:845–847

Bahadur N, Leathart JBS, Mutch E et al (2002) CYP2C8 polymorphisms in Caucasians and their relationship with paclitaxel 6α-hydroxylase activity in human liver microsomes. Biochem Pharmacol 64:1579–1589

Balian JD, Sukhova N, Harris JW et al (1995) The hydroxylation of omeprazole correlates with S-mephenytoin metabolism: a population study. Clin Pharmacol Ther 57:662–669

Bapiro TE, Hasler JA, Ridderström M, Masimirembwa CM (2002) The molecular and enzyme kinetic basis for the diminished activity of the cytochrome P450 2D6*17 variant. Biochem Pharmacol 64:1387–1398

Baud-Camus F, Marquet P, Soursac M et al (2001) Determination of N-acetylation phenotype using caffeine as a metabolic probe and high-performance liquid chromatography with either ultraviolet detection or electrospray mass spectrometry. J Chromatogr B 760:55–63

Bertilsson L, Lou YQ, Du YL et al (1992) Pronounced differences between native Chinese and Swedish populations in the polymorphic hydroxylations of debrisoquin and S-mephenytoin. Clin Pharmacol Ther 51:388–397

Bertilsson L, Dahl ML, Dalen P, Al-Shurbaji A (2002) Molecular genetics of CYP2D6: clinical relevance

with focus on psychotropic drugs. Br J Clin Pharmacol 53:111–122
Bolden RD, Hoke SHII, Eichhold TH et al (2002) Semi-automated liquid–liquid back extraction in a 96-well format to decrease sample preparation time for the determination of dextromethorphan and dextrorphan in human plasma. J Chromatogr B 772:1–10
Brentano C, Jaillon P (2001) Variability of cytochrome P450 1A2 activity over time in young and elderly healthy volunteers. Br J Clin Pharmacol 52:601–604
Butcher NJ, Boukouvala S, Sim E, Minchin RF (2002) Pharmacogenetics of the arylamine N-acetyltransferases. Pharmacogenomics J 2:30–42
Campbell ME, Spielberg SP, Kalow W (1987) A urinary metabolic ratio that reflects systemic caffeine clearance. Clin Pharmacol Ther 42:157–165
Caubet MS, Comte B, Brazier JL (2004) Determination of urinary 13C-caffeine metabolites by liquid chromatography–mass spectrometry: the use of metabolic ratios to assess CYP1A2 activity. J Pharm Sci Biomed Anal 34:379–389
Chang M, Dahl ML, Tybring G et al (1995) Use of omeprazole as a probe drug for CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and CYP2C19 genotype. Pharmacogenetics 5:358–363
Chladek J, Zimova G, Martinkova J, Tuma I (1999) Intra-individual variability and influence of urine collection period on dextromethorphan metabolic ratios in healthy subjects. Fundam Clin Pharmacol 13:508–515
Chladek J, Zimova G, Beranek M, Martinkova J (2000) In-vivo indices of CYP2D6 activity: comparison of dextromethorphan metabolic ratios in 4-h urine and 3-h plasma. Eur J Clin Pharmacol 56:651–657
Chou WH, Yan FX, Robbins-Weilert DK et al (2003) Comparison of two CYP2D6 genotyping methods and assessment of genotype–phenotype relationships. Clin Chem 49:542–551
Cok I, Kocabas NA, Cholerton S et al (2001) Determination of coumarin metabolism in Turkish population. Human Exp Pharmacol 20:179–184
Commandeur JNM, Stintjes GJ, Vermeulen NPE (1995) Enzymes and transport systems involved in the formation and disposition of glutathione S-conjugates. Pharmacol Rev 47:271–330
Csillag K, Vereczkey L, Gachalyi B (1989) Simple high-performance liquid chromatographic method for the determination of tolbutamide and its metabolites in human plasma and urine using photo-diode array detection. J Chromatogr 490:355–363
Dai D, Zeldin DC, Blaisdell JA et al (2001) Polymorphisms in CYP2C8 decrease metabolism of the anticancer drug paclitaxel and arachidonic acid. Pharmacogenetics 11:597–607
Daly AK, King BP (2003) Pharmacogenetics of anticoagulants. Pharmacogenetics 13:247–252
Eap CB, Buclin T, Hustert E et al (2004a) Pharmacokinetics of midazolam in CYP3A4- and CYP3A5-genotyped subjects. Eur J Clin Pharmacol 60:231–236
Eap CB, Buclin T, Cucchia G et al (2004b) Oral administration of midazolam (75 μg) as an in vivo probe for CYP3A activity. Eur J Clin Pharmacol 60:237–246
Endo T, Nakajima M, Fukami T et al (2008) Genetic polymorphisms of CYP2A6 affect the in-vivo pharmacokinetics of pilocarpine. Pharmacogenet Genomics 18:761–772
Ernstgard L, Warholm M, Johanson G (2004) Robustness of chlorzoxazone as an in vivo measure of cytochrome P450 2E1 activity. Br J Clin Pharmacol 58:190–200
Evans WE (2004) Pharmacogenetics of thiopurine S-methyltransferase and thiopurine therapy. Ther Drug Monit 26:186–191
Faucette SR, Hawke L, Lecluyse RL et al (2000) Validation of bupropion hydroxylation as a selective marker of human cytochrome P450 2B6 catalytic activity. Drug Metab Dispos 28:1222–1230
Fischer MB, Paine MF, Strelevitz TJ, Wrighton SA (2001) The role of hepatic and extrahepatic UDP–glucuronyltransferases in human drug metabolism. Drug Metab Rev 33:273–297
Frison G, Tedeschi L, Maietti S, Ferrara SD (2001) Determination of midazolam in plasma by solid-phase microextraction and gas chromatography/mass spectrometry. Rapid Commun Mass Spectrom 15:2497–2501
Frye RF, Stiff DD (1996) Determination of chlorzoxazone and 6-hydroxy chlorzoxazone in human plasma and urine by high performance liquid chromatography. J Chromatogr B 686:291–296
Frye RF, Adedoyin A, Mauro K et al (1998) Use of chlorzoxazone as an in vivo probe of cytochrome P450 2E1: choice of dose and phenotypic trait measure. J Clin Pharmacol 38:82–89
Fuchs P, Haefeli WE, Ledermann HR, Wenk M (1999) Xanthine oxidase inhibition by allopurinol affects the reliability of urinary caffeine metabolic ratios as markers for N-acetyltransferase2 and CYP1A2 activities. Eur J Clin Pharmacol 54:869–876
Fuhr U, Rost LK (1994) Simple and reliable CYP1A2 phenotyping by the paraxanthine/caffeine ratio in plasma and saliva. Pharmacogenetics 4:109–116
Fujieda M, Yamazaki H, Saito T et al (2004) Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behaviour and tobacco-related lung cancer risk in male Japanese smokers. Carcinogenesis 25:2451–2458
Gaedigk A, Bradford LD, Marcucci KA, Leeder JS (2002) Unique CYP2D6 activity distribution and genotype–-phenotype discordance in black Americans. Clin Pharmacol Ther 72:76–89
Gafni I, Nolte H, Tyndale R et al (2001) Resolving the roles of CYP2C19 and CYP3A4 in the metabolism of omeprazole in vivo using chronic omeprazole and ketoconazole. FASEB J 15:A918
Garcia-Martin E, Martinez C, Tabarés B et al (2004) Interindividual variability in ibuprofen

pharmacokinetics is related to interaction of cytochrome P450 2C8 and 2C9 amino acid polymorphisms. Clin Pharmacol Ther 76:119–127
Glatt H, Meinl W (2004) Pharmacogenetics of soluble sulfotransferases (SULTs). Naunyn Schmiedeberg's Arch Pharmacol 369:55–68
Gonzalez HM, Romero EM, Peregrina AA et al (2003) CYP2C19- and CYP3A4-dependent omeprazole metabolism in west Mexicans. J Clin Pharmacol 43:1211–1215
Gorski JC, Jones DR, Haehner-Daniels BD et al (1998) The contribution of intestinal and hepatic CYP3A to the interaction between midazolam and clarithromycin. Clin Pharmacol Ther 64:133–143
Grant DM, Tang BK, Kalow W (1984) A simple test for acetylator phenotype using caffeine test. Br J Clin Pharmacol 17:459–464
Guillemette C (2003) Pharmacogenomics of human UDP–glucuronosyltransferase enzymes. Pharmacogenomics J 3:136–158
Hadasova E, Franke G, Zschiesche M et al (1996) Debrisoquin 4-hydroxylation and sulphamethazine - N-acetylation in patients with schizophrenia and major depression. Br J Clin Pharmacol 41:428–431
Hägg S, Spigset O, Dahlqvist R (2001) Influence of gender and oral contraceptives on CYP2D6 and CYP2C19 activity in healthy volunteers. Br J Clin Pharmacol 51:169–173
Han XM, Ou-Yang DS, Lu PX et al (2001) Plasma caffeine metabolite ratio (17X/137X) in vivo associated with G-2964A and C734A polymorphisms of human CYP1A2. Pharmacogenetics 11:429–435
Hansen LL, Brosen K (1999) Quantitative determination of tolbutamide and its metabolites in human plasma and urine by high-performance liquid chromatography and UV detection. Ther Drug Monitor 21:664–671
Härtter S, Baier D, Dingemanse J et al (1996) Automated determination of dextromethorphan and its main metabolites in human plasma by high-performance liquid chromatography and column switching. Ther Drug Monit 18:297–303
Hayes JD, Strange RC (2000) Glutathione S-transferase polymorphisms and their biological consequences. Pharmacology 61:154–166
Hesse LM, He P, Krishnaswamy S et al (2004) Pharmacogenetic determinants of interindividual variability in bupropion hydroxylation by cytochrome P450 2B6 in human liver microsomes. Pharmacogenetics 14:225–238
Hoskins JM, Shenfield GM, Gross AS (1997) Modified high-performance liquid chromatographic method to measure both dextromethorphan and proguanil for oxidative phenotyping. J Chromatogr B 696:81–87
Hu OY, Tang HS, Chang WH, Hu TM (1998) Novel single-point plasma or saliva dextromethorphan method for determining CYP2D6 activity. J Pharmacol Exp Ther 285:955–960
Innocenti F, Iyer L, Ratain MJ (2000) Pharmacogenetics: a tool for individualizing antineoplastic therapy. Clin Pharmacokinet 39:315–325
Innocenti F, Undvia SD, Iyer L et al (2004) Genetic variants in the UDP–glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. J Clin Oncol 8:1382–1388
Ishizaki T, Sohn DR, Kobayashi K et al (1994) Interethnic differences in omeprazole metabolism in the two S-mephenytoin hydroxylation phenotypes studied in Caucasians and Orientals. Ther Drug Monit 16:214–215
Iyer L, Das S, Janish L et al (2002) UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. Pharmacogenomics J 2:43–47
Jetter A, Kinzig-Schippers M, Skott A et al (2004) Cytochrome P450 2C9 phenotyping using low-dose tolbutamide. Eur J Clin Pharmacol 60:165–175
Johansson I, Oscarson M, Yue QY et al (1994) Genetic analysis of the Chinese cytochrome P450D locus: characterization of variant CYP2D6 genes present in subjects with diminished capacity for debrisoquin hydroxylation. Mol Pharmacol 46:452–459
Kalow W, Tang BK (1991) Use of caffeine metabolite ratios to explore CYP1A2 and xanthine oxidase activities. Clin Pharmacol Ther 50:508–519
Kalow W, Tang BK (1993) The use of caffeine for enzyme assays: a critical appraisal. Clin Pharmacol Ther 53:203–514
Kanazawa H, Atsumi R, Matsushima Y et al (2000) Determination of theophylline and its metabolites in biological samples by liquid chromatography–mass spectrometry. J Chromatogr A 870:87–96
Kanazawa H, Okada A, Matsushima Y et al (2002) Determination of omeprazole and its metabolites in human plasma by liquid chromatography–mass spectrometry. J Chromatogr A 949:1–9
Kashuba ADM, Bertino JS, Kearns GL et al (1998) Quantitation of three-month intraindividual variability and influence of sex and menstrual cycle phase on CYP1A2, N-acetyltransferase-2, and xanthine oxidase activity determined with caffeine phenotyping. Clin Pharmacol Ther 63:540–551
Kidd RS, Curry TB, Gallagher S et al (2001) Identification of a null allele of CYP2C9 in an African-American exhibiting toxicity to phenytoin. Pharmacogenetics 11:803–808
Kim RB, O'Shea D, Wilkinson GR (1995) Interindividual variability of chlorzoxazone 6-hydroxylation in men and women and its relationship to CYP2E1 genetic polymorphisms. Clin Pharmacol Ther 57:645–655
Kimura M, Ieri I, Wada Y, Mamyia K (1999) Reliability of the omeprazole hydroxylation index and length of therapy. Br J Clin Pharmacol 47:115–119
Kirchheiner J, Bauer S, Meineke I et al (2002a) Impact of CYP2C9 and CYP2C19 polymorphisms on tolbutamide kinetics and the insulin and glucose response in healthy volunteers. Pharmacogenetics 12:101–109
Kirchheiner J, Brockmöller J, Meineke I et al (2002b) Impact of CYP2C9 amino acid polymorphisms on glyburide kinetics and on the insulin and glucose response in healthy volunteers. Clin Pharmacol Ther 71:286–296

Kirchheiner J, Störmer E, Meisel C et al (2003a) Influence of CYP2C9 genetic polymorphisms on pharmacokinetics of celecoxib and its metabolites. Pharmacogenetics 13:473–480

Kirchheiner J, Klein C, Meineke I et al (2003b) Bupropion and 4-OH-bupropion pharmacokinetics in relation to genetic polymorphisms in CYP2B6. Pharmacogenetics 13:619–626

Kita T, Tanigawara Y, Chikizawa S et al (2001) N-acetyltransferase2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients. Biol Pharm Bull 24:544–549

Klebovitch I, Rautio A, Salonpää P et al (1995) Antipyrine, coumarin and glipizide affect N-acetylation measured by caffeine test. Biomed Pharmacother 49:225–227

Krul C, Hageman G (1998a) Analysis of urinary caffeine metabolites to assess biotransformation enzyme activities by reverse-phase high-performance liquid chromatography. J Chromatogr B 709:24–27

Krul C, Hageman G (1998b) Analysis of urinary caffeine metabolites to assess biotransformation enzyme activities by reverse-phase high-performance liquid chromatography. J Chromatogr B 709:27–34

Lagerstrom PO, Persson BA (1984) Determination of omeprazole and its metabolites in plasma and urine by liquid chromatography. J Chromatogr A 309:347–356

Laine K, Tybring G, Bertilsson L (2000) No sex-related difference but significant inhibition by oral contraceptives of CYP2C19 activity as measured by the probe drugs mephenytoin and omeprazole in healthy Swedish white subjects. Clin Pharmacol Ther 68:151–159

Lamba JK, Lin YS, Schuety EG, Thummel KE (2002) Genetic contribution to variable human CYP3A-mediated metabolism. Adv Drug Deliv Rev 54:1271–1294

Lang T, Klein K, Richter T et al (2004) Multiple novel nonsynonymous CYP2B6 gene polymorphisms in Caucasians: demonstration of phenotypic null alleles. JPET 311:34–43

Lecomte T, Landi B, Beaune P et al (2006) Glutathione S-transferase P1 polymorphism (Ile ^{105}Val) predicts cumulative neuropathy in patients receiving oxaliplatin-based chemotherapy. Clin Cancer Res 12:3050–3056

Lee CR, Pieper JA, Hinderliter AL et al (2002a) Evaluation of cytochrome P450 2C9 metabolic activity with tolbutamide in CYP2C9*1 heterozygotes. Clin Pharmacol Ther 72:562–571

Lee CR, Goldstein JA, Pieper JA (2002b) Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in vitro and human data. Pharmacogenetics 12:251–263

Lee JI, Chaves-Gnecco D, Amico JA et al (2002c) Application of simultaneous midazolam administration for hepatic and intestinal cytochrome P450 3A phenotyping. Clin Pharmacol Ther 72:718–728

Lee CR, Pieper JA, Reginald FF et al (2003) Tolbutamide, flurbiprofen and losartan as probes of CYP2C9 activity in humans. J Clin Pharmacol 43:84–91

Lepper ER, Hicks JK, Verweij J et al (2004) Determination of midazolam in human plasma by liquid chromatography with mass-spectrometric detection. J Chromatogr B 806:305–310

Lieber CS (1997) Cytochrome P-450 2E1: its physiological and pathological role. Physiol Rev 77:517–544

Lin YS, Lockwood GF, Graham MA et al (2001) In-vivo phenotyping for CYP3A by single-point determination of midazolam plasma concentration. Pharmacogenetics 11:781–791

Lown K, Kolars J, Turgeon K et al (1992) The erythromycin breath test selectively measures P450IIIA in patients with severe liver disease. Clin Pharmacol Ther 52:229–238

Lucas D, Ferrara R, Gonzalez E et al (1999) Chlorzoxazone: a selective probe for phenotyping CYP2E1 in humans. Pharmacogenetics 9:377–388

McCrea J, Prueksaritanont T, Gertz BJ et al (1999) Concurrent administration of the erythromycin breath test (EBT) and oral midazolam as in vivo probes for CYP3A activity. J Clin Pharmacol 39:1212–1220

McCune J, Lindley C, Decker JL et al (2001) Lack of gender differences and large intrasubject variability in cytochrome P450 activity measured by phenotyping with dextromethorphan. J Clin Pharmacol 41:723–731

McLeod HL, Evans WE (2001) Pharmacogenomics: unlocking the human genome for better drug therapy. Annu Rev Pharmacol Toxicol 41:101–121

Meisel P (2002) Arylamine N-acetyltransferases and drug response. Pharmacogenomics 3:349–366

Meisel P, Schroeder C, Wulff K, Siegmund W (1997) Relationship between human genotype and phenotype of N-acetyltransferase (NAT2) as estimated by discriminant analysis and multiple linear regression: 1. Genotype and N-acetylation in vivo. Pharmacogenetics 7:241–246

Meyer UA (2000) Pharmacogenetics and adverse drug reactions. Lancet 356:1667–1671

Meyer UA, Zanger UM (1997) Molecular mechanisms of genetic polymorphisms of drug metabolism. Annu Rev Pharmacol Toxicol 37:269–296

Miners JO, Birkett DJ (1998) Cytochrome P450 2C9: an enzyme of major importance in human drug metabolism. Br J Clin Pharmacol 45:525–538

Mongey AB, Sim E, Risch A, Hess E (1999) Acetylation status is associated with serological change but not clinically significant disease in patients receiving procainamide. J Rheumatol 26:1721–1726

Morin S, Bodin L, Loriot MA et al (2004) Pharmacogenetics of acenocoumarol. Clin Pharmacol Ther 75:403–414

Niemi M, Leathart JB, Neuvonen M et al (2003) Polymorphisms in CYP2C8 is associated with reduced plasma concentrations of repaglinide. Clin Pharmacol Ther 74:380–387

Nordmark A, Lundgren S, Cnattingius S et al (1999) Dietary caffeine as a probe agent for assessment of cytochrome P450 1A2 activity in random urine samples. Br J Clin Pharmacol 47:397–402

Nordmark A, Lundgren S, Ask B et al (2002) The effect of the CYP1A2*F mutation on CYP1A2 inducibility in pregnant women. Br J Clin Pharmacol 54:504–510

Nyeki A, Buclin T, Biollaz J, Decosterd LA (2002) NAT2 and CYP1A2 phenotyping with caffeine: head-to-head comparison of AFMU vs AAMU in the urine metabolite ratios. Br J Clin Pharmacol 55:62–67

O'Neil WM, Drobitch RK, MacArthur RD et al (2000) Acetylator phenotype and genotype in patients infected with HIV: discordance between methods for phenotype determination and genotype. Pharmacogenetics 10:171–182

Okumura K, Kita T, Chikazawa S et al (1997) Genotyping of N-acetylation polymorphism and correlation with procainamide metabolism. Clin Pharm Ther 61:509–517

Ono S, Hatanaka T, Hotta H et al (1995) Chlorzoxazone is metabolized by human CYP1A2 as well as by human CYP2E1. Pharmacogenetics 5:143–150

Oscarson M (2001) Genetic polymorphisms in the cytochrome P450 2A6 (CYP2A6) gene: implications for interindividual differences in nicotine metabolism. Drug Metab Dispos 29:91–95

Ou-Yang DS, Huang SL, Wang W et al (2000) Phenotypic polymorphism and gender-related differences of CYP1A2 activity in a Chinese population. Br J Clin Pharmacol 49:145–151

Paoluzzi L, Singh AS, Price DK, Danesi R (2004) Influence of genetic variants in UGT1A1 and UGT1A9 in the in vivo glucuronidation of SN-38. J Clin Pharmacol 44:854–860

Pelkonen O, Rautio A, Raunio H, Pasanene M (2000) CYP2A6: a human coumarin 7-hydroxylase. Toxicology 144:139–147

Queiroz RH, Dreossi SA, Carvalho D (1997) A rapid, specific, and sensitive method for the determination of acetylation phenotype using dapsone. J Anal Toxicol 21:203–207

Rane A (1999) Phenotyping of drug metabolism in infants and children: potentials and problems. Pediatrics 104:640–643

Rasmussen BB, Bosen K (1996) Determination of urinary metabolites of caffeine for the assessment of cytochrome P450 1A2, xanthine oxidase and N-acetyltransferase activity in humans. Ther Drug Monitor 18:254–262

Raunio H, Rautio A, Gullstén H, Pelkonen O (2001) Polymorphisms of CYP2A6 and its practical consequences. Br J Clin Pharmacol 52:357–363

Relling MV, Lin JS, Ayers GD, Evans EE (1992) Racial and gender differences in N-acetyltransferase, xanthine oxidase and CYP1A2* activities. Clin Pharmacol Ther 52:643–658

Rivory LP, Slaviero KA, Seale JP et al (2000) Optimizing the erythromycin breath test for use in cancer patients. Clin Cancer Res 6:3480–3485

Rogers JF, Nafziger AN, Bertino JS (2002) Pharmacogenetics affects dosing, efficacy, and toxicity of cytochrome P450-metabolized drugs. Am J Med 113:746–750

Rosemary J, Adithan C (2007) The pharmacogenetics of CYP2C9 and CYP2C19: ethnic variation and clinical significance. Curr Clin Pharmacol 2:93–109

Rostami-Hodjegan A, Nurminen S, Jackson PR, Tucker GT (1996) Caffeine urinary metabolic ratios as markers of enzyme activity: a theoretical assessment. Pharmacogenetics 6:121–149

Rothen JP, Haefeli WE, Meyer UA et al (1998) Acetominophen is an inhibitor of hepatic N-acetyltransferase 2 in vitro and in vivo. Pharmacogenetics 8:553–559

Sachse C, Brockmöller J, Bauer S, Roots I (1997) Cytochromes P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. Am J Hum Genet 60:284–295

Sachse S, Bhambra U, Smith G et al (2003) Polymorphisms in the cytochrome P4501A2 gene (CYP1A2) in colorectal cancer patients and control: allele frequencies, linkage disequilibrium and influence on caffeine metabolism. Br J Clin Pharmacol 55:68–76

Saitoh A, Fletcher C, Brundage R et al (2007) Efavirenz pharmacokinetics in HIV-1-infected children are associated with CYP2B6–G516T polymorphism. J Acquir Immune Defic Syndr 45:280–285

Sandberg M, Johansson I, Christensen M et al (2004) The impact of CYP2C9 genetics and oral contraceptives on cytochrome P450 2C9 phenotype. Drug Metab Dispos 32:484–489

Schmid B, Bircher J, Preisig R, Küpfer A (1985) Polymorphic dextromethorphan metabolism: co-segregation of oxidative O-demethylation with debrisoquin hydroxylation. Clin Pharmacol Ther 38:618–624

Schreiber-Deturmeny E, Bruguerolle B (1996) Simultaneous high-performance liquid chromatographic determination of caffeine and theophylline for routine drug monitoring in human plasma. J Chromatogr B 677:305–312

Schwarz UI (2003) Clinical relevance of genetic polymorphisms in the human CYP2C9 gene. Eur J Clin Invest 33:23–30

Scoot RJ, Palmer J, Lewis IA, Pleasance S (1999) Determination of a "GW cocktail" of cytochrome P450 probe substrates and their metabolites in plasma and urine using automated solid phase extraction and fast gradient liquid chromatography tandem mass spectrometry. Rapid Commun Mass Spectrom 13:2305–2319

Scordo MG, Aklillu E, Dahl ML et al (2001) Genetic polymorphism of cytochrome P450 2C9 in a Caucasian and a black African population. Br J Clin Pharmacol 52:447–450

Scordo MG, Caputi AP, D'Arrigo C et al (2004) Allele and genotype frequencies of CYP2C9, CYP2C19 and CYP2D6 in an Italian population. Pharmacol Res 50:195–200

Sekino K, Kubota T, Okada Y et al (2003) Effect of the single CYP2C9*3 allele on pharmacokinetics and pharmacodynamics in healthy Japanese subjects. Eur J Clin Pharmacol 59:589–592

Sevilla-Mantilla C, Ortega L, Agundez JA et al (2004) Leflunomide acute hepatitis. Dig Liver Dis 36:82–84

Shimada Y, Yamazki H, Mimura M et al (1994) Interindividual variations in human liver cytochrome

P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther 270:413–423

Shirley KL, Hon YY, Penzak S et al (2003) Correlation of cytochrome P450 1A2 activity using caffeine phenotyping and olanzapine disposition in healthy volunteers. Neuropsychopharmacology 25:961–966

Shong JH, Yoon YR, Kim KA et al (2002) Effects of CYP2C9 and CYP2C19 genetic polymorphisms on the disposition and blood glucose lowering response to tolbutamide in humans. Pharmacogenetics 12:111–119

Shou M, Korzekwa KR, Brooks EN et al (1997) Role of human hepatic cytochrome P450 1A2 and 3A4 in the metabolic activation of estrone. Carcinogenesis 18:207–214

Simon T, Becquemont L, Hamon B et al (2003) Determination of –3858G-A and 164C-A genetic polymorphisms of CYP1A2 in blood and saliva by rapid alleleic discrimination: large difference in the prevalence of the –3858G-A mutation between Caucasians and Asians. Eur J Clin Pharmacol 59:343–346

Tamminga WJ, Werner J, Oosterhuis B et al (1999) CYP2D6 and CYP2C19 activity in a large population of Dutch healthy volunteers: indications for oral contraceptive-related gender differences. Eur J Clin Pharmacol 55:177–184

Tang BK, Kadar D, Qian L et al (1991) Caffeine as a metabolic probe: validation of its use for acetylator phenotyping. Clin Pharmacol Ther 49:648–657

Tetlow N, Robinson A, Mantle T, Board P (2004) Polymorphism of human mu class glutathione transferases. Pharmacogenetics 14:359–368

Thummel KE, Shen DD, Podoll TD et al (1994) Use of midazolam as a human cyotochrome P450 3A probe: in vitro–in vivo correlations in liver transplant patients. J Pharmacol Exp Ther 271:549–556

Thummel KE, O'Shea D, Paine MF et al (1996) Oral first-pass elimination of midazolam involves both gastrointestinal and hepatic CYP3A-mediated metabolism. Clin Pharmacol Ther 59:491–502

Tukey RH, Strassburg CP (2000) Human UDP–glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol 40:581–616

Turpeinen M, Raunio H, Pelkonen O (2006) The functional role of CYP2B6 in human drug metabolism: substrates and inhibitors in vitro, in vivo and in silico. Curr Drug Metabol 7:705–714

Tybring G, Bottiger Y, Wyden J, Bertilsson L (1997) Enantioselective hydroxylation of omeprazole catalyzed by CYP2C19. Clin Pharmacol Ther 62:129–137

Veronese ME, Miners JO, Randles D et al (1990) Validation of the tolbutamide metabolic ratio for population screening with use of sulfaphenazole to produce model phenotypic poor metabolizers. Clin Pharmacol Ther 47:403–411

Veronese ME, Miners JO, Ress DLP, Birkett DJ (1993) Tolbutamide hydroxylation in humans: lack of bimodality in 106 healthy subjects. Pharmacogenetics 3:86–93

Versuyft C, Robert A, Morin S et al (2003) Genetic and environmental risk factors for oral anticoagulant overdose. Eur J Clin Pharmacol 58:739–745

Villeneuve L, Girard H, Fortier LC et al (2003) Novel functional polymorphisms in the UGT1A7 and UGT1A9 glucuronidating enzymes in Caucasian and African-American subjects and their impact on the metabolism of 7-ethyl-10-hydroxycamptothecin and flavopiridol anticancer drugs. J Pharmacol Exp Ther 307:117–128

von Richter O, Burk O, Fromm MF et al (2004) Cytochrome P450 3A4 and P-glycoprotein expression in human small intestinal enterocytes and hepatocytes: a comparative analysis in paired tissue specimens. Clin Pharmacol Ther 75:172–183

Ward B, Gorski C, Jones D et al (2003) The cytochrome P450 2B6 (CYP2B6) is the main catalyst of efavirenz primary and secondary metabolism: implication for HIV/AIDS therapy and utility of efavirenz as a substrate marker of CYP2B6 catalytic activity. JPET 306:287–300

Watkins PB (1994) Noninvasive tests of CYP3A enzymes. Pharmacogenetics 4:171–184

Weber WW, Hein DW (1985) N-acetylation pharmacogenetics. Pharmacol Rev 37:25–79

Wells PG et al (2004) Symposium article: glucuronidation and the UDP–glucuronosyltransferases in health and disease. Drug Metab Dispos 32:281–290

Wilkinson GR (2004) Genetic variability in cytochrome P450 3A5 and in vivo cytochrome P450 3A activity: some answers but still questions. Clin Pharmacol Ther 76:99–103

Wu YJ, Cheng YY, Zeng S, Ma MM (2003) Determination of dextromethorphan and its metabolite dextrorphan in human urine by capillary gas chromatography without derivatization. J Chromatogr B 784:219–224

Wusk B, Kullak-Ublick GA, Rammert C et al (2004) Thiopurine S-methyltransferase polymorphisms: efficient screening method for patients considering taking thiopurine drugs. Eur J Clin Pharmacol 60:5–10

Xie HG, Kim RB, Wood AJ, Stein CM (2001) Molecular basis of ethnic differences in drug disposition and response. Annu Rev Pharmacol Toxicol 41:815–850

Xie HG, Prasad HC, Kim RB, Stein CM (2002) CYP2C9 allelic variants: ethnic distribution and functional significance. Adv Drug Deliv Rev 54:1257–1270

Xu C, Goodz S, Sellers E, Tyndale RF (2002) CYP2A6 genetic variation and potential consequences. Adv Drug Deliv Rev 54:1245–1256

Yasar U, Dahl ML, Christensen M, Eliasson E (2002a) Intra-individual variability in urinary losartan oxidation ratio: an in vivo marker of CYP2C9 activity. Br J Clin Pharmacol 54:183–185

Yasar U, Forslund-Bergengren C, Tybring G et al (2002b) Pharmacokinetics of losartan and its metabolite E-3174 in relation to the CYP2C9 genotype. Clin Pharmacol Ther 71:89–98

Yasar U, Lundgren S, Eliasson E et al (2002c) Linkage between CYP2C8 and CYP2C9 genetic polymorphisms. Biochem Biophys Res Commun 299:25–28

Yim DS, Jeong JR, Park JY (2001) Assay of omeprazole and omeprazole sulfone by semi-microcolumn liquid chromatography with mixed-function precolumn. J Chromatogr B 754:487–493

Yin OQP, Tomlinson B, Chow ALH et al (2004) Omeprazole as a CYP2C19 marker in Chinese subjects: assessment of its gene-dose effect and intrasubject variability. J Clin Pharmacol 44:582–589

Zanger UM, Raimindo S, Eichelbaum M (2004) Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. Naunyn Schmiedeberg's Arch Pharmacol 369:23–37

Zhai S, Sausville EA, Senderowicz AM et al (2003) Clinical pharmacology and pharmacogenetics of flavopiridol 1-h IV infusion in patients with refractory neoplasms. Anti-Cancer Drugs 14:125–135

Zhou S (2006) Clinical pharmacogenomics of thiopurine S-methyltransferase. Curr Clin Pharmacol 1:119–128

PK/PD Approaches

57

Yichao Yu, Diether Rüppel, Willi Weber, and Hartmut Derendorf

Contents

Abstract

The success of the drug development program heavily relies on the rational drug design with appropriate choice of drug and dosing regimen. This requires a good understanding of both drug delivery mechanism and drug response mechanism. Two of the most important pharmacologic disciplines, namely, pharmacokinetics (PK) and pharmacodynamics (PD), can be linked together by PK/PD approach, which has tremendous potential to influence decision-making through modeling and simulation. With its nature of an interdisciplinary science,

Willi Weber has retired.

Y. Yu · H. Derendorf (✉)
Department of Pharmaceutics College of Pharmacy, University of Florida, Gainesville, FL, USA
e-mail: yyu2013@ufl.edu; hartmut@ufl.edu

D. Rüppel
PKDM/TMED, Sanofi-Aventis Germany, Frankfurt, Germany
e-mail: diether.rueppel@sanofi.com

W. Weber
Frankfurt, Germany

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_26

this state-of-art strategy can leverage different kinds of preclinical and clinical data through mathematical and statistical models. This framework is powerful to assist researchers with better understanding of drug behavior and effectiveness, disease progression, and the impact of demographic characteristics on a subpopulation or individual patients. The aim of this chapter is to provide an overview of basic concepts in PK and PD and discuss various approaches in PK/PD modeling and simulation, together with its applications in antibiotic drug development as it is very well established in this field. The PK/PD concepts and theories presented here are not limited to antibiotics only but can also be broadly applied to drug development in other therapeutic areas.

Purpose and Rationale

Pharmacokinetics (PK) describes the time course of the drug concentration in the body, normally in blood or plasma. Pharmacodynamics (PD) describes the relationship between drug concentration and response, desired or not desired. PK/PD is the link between PK and PD describing the time course of drug effect.

Reviews on PK/PD are regularly published (Derendorf and Meibohm 1999; Derendorf et al. 2000; Csajka and Verotta 2006), and the more specialized literature is becoming more and more abundant. The book of Gabrielsson and Weiner (2016) is an excellent introduction to PK/PD and modeling. This chapter is a comprehensive summary of the state-of-the art with references to basic and recent literature. PK/PD modeling is assessed by reviewing and discussing its application in antibiotic drug development.

Mathematical Models in Biology

Biology is complex, much too complex to be completely described in equations. And it is extremely difficult to obtain all the data needed for a detailed description. The target exposure or concentration may not be easy to be determined with the limited raw data from observation. Models are reductions of reality to a mathematical system that can be handled, for example, by a computer. Models are informatic descriptions of a true biological process. They are often the simplified representations but differ in the degree of simplification. Since the nature of simplification is based on the intended use of the model, therefore there is no right model but rather judged by "fitness for purpose" (Mould and Upton 2012). Model cannot be better than the data without using previous knowledge. A dataset will never be complete. Models will differ depending on the available data. And each model is part of the truth. The situation is not new. It is the morale of the famous parable of "the blind men and the elephant" probably originating from Asia a long time ago:

> Six blind men were asked to determine what an elephant looked like by feeling different parts of the elephant's body. The blind man who feels a leg says the elephant is like a pillar; the one who feels the tail says the elephant is like a rope; the one who feels the trunk says the elephant is like a tree branch; the one who feels the ear says the elephant is like a hand fan; the one who feels the belly says the elephant is like a wall; and the one who feels the tusk says the elephant is like a solid pipe. And they started arguing and fighting and could not find out what was the truth. Finally, a wise man explains to them: "All of you are right. The reason every one of you is telling it differently is because each one of you touched a different part of the elephant. So, actually the elephant has all the features you mentioned."

Depending on the data and the task, different parts of a system are described, and different conclusions can be drawn. The question about the use of a model under development decides on the data needed. What is the resolution in time, in space, and in concentration required? Will the model be used for interpolation only, or is it to simulate new situations by extrapolation? Pure empiric models can be used to interpolate in the limits of data used to develop the model. More mechanistic models are more difficult, but more useful when cautious extrapolation beyond the limits of observed data is required. Finally, models may change if additional data are available.

Modeling in Pharmacometrics

The modeler can use two principal types of models. One possibility is to fit a simpler empirical model with large sample size or exposure ranges. Such approach is useful for describing the underlining relationship. However, further extrapolation outside of the study population would not be straightforward if the parameters are not biologically interpretable. Hence empirical models are usually descriptive models which would provide adequate information to understand the response difference in the data. The other type of the model incorporates the essential pharmacological and physiological information into the model but still keeps it as simple as possible, which is also classified as predictive models. The modeling approach can start with a complex model where the complete physiological knowledge is expressed in mathematical equations and then simplify this complicated model to the level that it can be processed in a computer and the parameters are identifiable. This type of models can be intended to extrapolate the same relationship to other populations from whom data has not been used to derive the model. Stronger assumptions are usually made for this type of model. Therefore, model validation is particularly important to ensure the confidence of model application for its intended use.

To assess the credibility of modeling and simulation, both the fidelity and robustness of the model need to be considered. The fidelity of the model can be gauged by comparing the model structure and parameters to the real and important biological basis. And the robustness of the model can be addressed by parameter sensitivity analysis, variance estimation, and bootstrap resampling. With the evolving computational sciences and unprecedented availability of data in pharmaceutical sciences, building credible models to quantitatively assess and evaluate drug therapies becomes feasible. Model-based drug development, mainly initiated by population PK/PD, provides a more rational and efficient approach for dose selection and optimization.

One of the biggest challenges for data interpretation is the variability in exposure and response between individual subjects. The variability comes from the individual's physiologic characteristics such as age, weight, gender, etc. Understanding the impact of these covariates on PK parameters is particularly important in model development. The data collected at different time points for the same individual are correlated with each other; ignoring this correlation would lead to inflated variability in PK/PD model and biased covariate-parameter relationship. In order to ensure the robustness of parameter estimation, mixed-effect modeling is often used to appropriately controlling variability in model fitting. Mixed-effect modeling does not only calculate the PK/PD parameters but also their statistical distribution in the population and is therefore called population PK/PD in this context.

Procedure

Basic Concepts of Pharmacokinetics

Pharmacokinetics (PK) is also referred as "what the body does to the drug" including the processes of absorption, distribution, metabolism, and excretion that govern the concentration-time course of a drug in blood or plasma. In principle, non-compartmental, compartmental, and physiologically based models can be used for PK modeling. Most PK models use compartments as the building block to describe the concentration-time course, in most cases in the central compartment. The "compartment," a very common abstract concept in PK models, is assumed as a region of the body where the drug is homogenously mixed and shares similar kinetics. The central compartment usually represents plasma or the systemic circulation. However, the central compartment is not necessarily the site of action since most drug-target action occurs in tissues. The drug concentration at the site of action may be accessible to measurements, but very often it is not and has to be calculated in the PK models. For that reason, the PK models generally consist of a central compartment and one or two peripheral compartments linked by distribution rate constants.

Drug distribution into tissue depends on both plasma protein and tissue binding. In plasma, drugs can bind to proteins such as albumin, α-1 acid glycoproteins, and lipoproteins. Only the free, unbound drug can distribute into the tissue via passive diffusion and subsequently reaches its site of action in the tissue for pharmacologic effects. Therefore, the frequently measured total plasma concentration may not be an ideal measure, but the free concentration at the target site, if possible, should be used to derive PK models. Achieving and maintaining goal tissue concentration near the target site is critical to maximize drug effectiveness and minimize drug toxicity. One of the most frequently used methods to sample unbound drug concentration in the interstitial fluid (ISF) of various tissues is microdialysis (Chaurasia et al. 2007). Rather than obtaining average total tissue concentration after homogenization of tissue samples, microdialysis technique provides a less invasive and more direct way to continuously measure the unbound concentration in the tissue. In brief, a microdialysis probe is implanted into the tissue and continuously perfused with a perfusate. By measuring the dialysate sample and correlating the concentration in dialysate and tissue, the free drug concentration in the ISF can thus be calculated.

To clearly explain how clinical microdialysis works, an example from Schuck et al. is illustrated here (Schuck et al. 2005). Microdialysis technique was used to study the effect of simulated microgravity on the tissue distribution of ciprofloxacin. Six healthy volunteers participated in a crossover study after a single 250-mg oral ciprofloxacin in normal gravity and simulated microgravity. Blood, urine, and *in vivo* microdialysis samples were obtained from thigh muscle in each subject. After probe implantation, probe calibration by retrodialysis was performed before microdialysis started. Probe was perfused with 0.1 mg/L ciprofloxacin solution, and drug concentration in dialysate was measured. *In vivo* recovery of ciprofloxacin was determined by computing the disappearance rate through the semipermeable membrane at the tip of microdialysis probe using the following equation:

$$\text{Recovery}\ (\%) = \left(1 - \frac{\text{Concentration}_{\text{dialysate}}}{\text{Concentration}_{\text{perfusate}}}\right) \times 100 \quad (1)$$

After the completion of probe calibration and washout period, ciprofloxacin was then administered, and microdialysis sample collection started and continues with the predetermined 30-min time interval for each collection. Under the assumption of quantitatively equal diffusion process in both directions through semipermeable membrane, the same *in vivo* recovery was used to calculate free concentration in the tissue:

$$\text{Concentration}_{\text{tissue}} = \frac{\text{Concentration}_{\text{dialysate}}}{\text{Recovery}(\%)} \times 100 \quad (2)$$

The mean PK profile of total plasma and free tissue concentrations is presented in Fig. 1. It shows very similar total ciprofloxacin concentration in plasma under normal gravity and microgravity; however, slightly lower but not significantly different tissue concentration was observed in simulated microgravity. If total plasma concentration was used to estimate the therapeutic outcome of ciprofloxacin, the drug efficacy may be overestimated. With the successful measurement of free concentration in tissue by clinical microdialysis, this study suggested a slightly impaired tissue penetration of ciprofloxacin in microgravity. The free concentrations at the target site are better predictors of therapeutic outcome since these are the direct link between PK and PD.

Basic Concepts of Pharmacodynamics

The drug at the target site interacts with the target and hence results in active pharmacological action. The term pharmacodynamics (PD) describes "what the drug does to the body." In general, a drug effect E is depending on time t, drug concentrations C in the past until present, and additional covariates co through a function F:

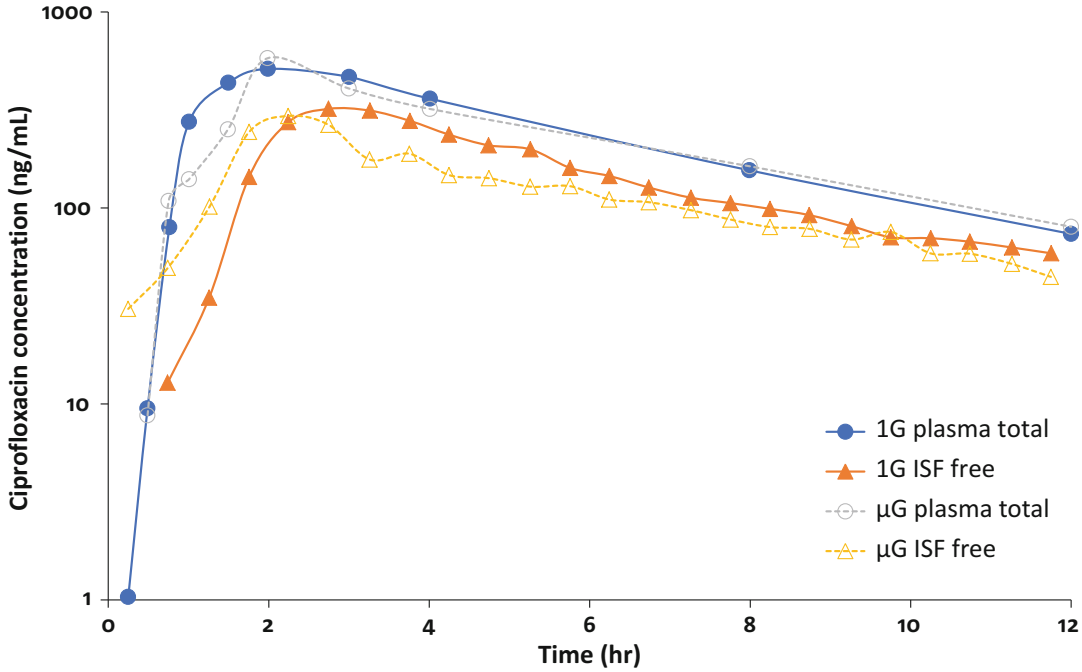


Fig. 1 Total plasma (circle) and free interstitial (triangle) concentrations of ciprofloxacin under Earth's gravity (solid) and simulated microgravity (open). Points represent the means of six subjects. (Image adapted from Schuck et al. (2005) and used with permission)

$$E(t) = F(t, C(-\infty,t), \mathrm{co}(-\infty,t)) \quad (3)$$

This is the description of a general dynamic system. In a static system, the effect depends only on the current time without any memory:

$$E(t) = F(t, C(t), \mathrm{co}(t)) \quad (4)$$

The system may be time variant or invariant. For an invariant system, the response to a drug concentration is always the same and does not depend on the time when it is applied. In a time-variant system, the response is depending on the time, for example, through time changing covariates co(t).

A relaxed system is at rest ($F = 0$) for drug concentrations $C = 0$:

$$F(t,0,\mathrm{co}(t)) \quad (5)$$

for all times $t \in [-\infty, \infty]$ and is excited ($F > 0$) by drug concentrations $C > 0$. For a linear system, the effect of the sum of two concentrations is the sum of the effect of each concentration. If the relationship between drug concentration and pharmacologic effect is linear, it can be expressed as a straight line:

$$E = S \bullet C + E_0 \quad (6)$$

with effect E, the slope of the line S, drug concentration C, and the effect in the absence of drug E_0. However, the relationship between concentration and effect may not be linear for the whole concentration range in most cases, especially at very low or very high drug concentrations. The linearly increase in drug effect also contradicts with the actual biologic response where a maximum response will be reached with ever-increasing drug concentration. Since many effects are non-linear, so this simple additive procedure cannot be applied in most situations.

A well-known model that can simply and adequately describe the concentration-effect relationship is the $E_{\max}$ model:

$$E = \frac{E_{\max} \bullet C}{\mathrm{EC}_{50} + C} \tag{7}$$

where $E_{\max}$ is the maximum drug effect and EC_{50} is the concentration yields 50% of the maximum effect. This model can mimic the increasing trend of effect as drug concentration increased and is also able to predict the saturation effect at very high concentration levels. It is also important to notice that this model can be simplified to a linear relationship for small $C \ll \mathrm{EC}_{50}$ such that $S = E_{\max}/\mathrm{EC}_{50}$ and to an inverse proportional relationship for large $C \gg \mathrm{EC}_{50}$ such that $\left(E = E_{\max} \times \left(1 - \frac{\mathrm{EC}_{50}}{C}\right)\right)$.

In order to adjust the shape of effect-concentration relationship, a Hill factor is often incorporated into the $E_{\max}$ model that becomes the so-called sigmoid $E_{\max}$ model:

$$E = \frac{E_{\max} \bullet C^{\gamma}}{\mathrm{EC}_{50}{}^{\gamma} + C^{\gamma}} \tag{8}$$

where γ affects the slope of the middle part of the curve with flattened curve if $\gamma < 1$ and steeper curve if $\gamma > 1$. As γ increases, the effect becomes more sensitive to the change of drug concentration. When γ is very large (>5), the effect E turns out to be 0 if below the threshold EC_{50} and $E_{\max}$ if above EC_{50}. The concentration-effect relationship is almost like an on-off switch in such threshold model. Although this parameter rarely has mechanistic meanings, this value may help characterize and classify drug effect (Goutelle et al. 2008). For example, it has been shown that a high Hill coefficient and low maximum kill rate are observed for time-dependent antibiotics, while a low Hill coefficient and high maximum kill rate are seen among concentration-dependent antibiotics (Czock and Keller 2007).

Types of Pharmacokinetic Models

The most important component of model-based drug development is the pharmacokinetic model, which has different types: non-compartmental analysis, compartmental models, and physiological models. Non-compartmental analysis uses concentration-time data to estimate essential PK parameters such as AUC, CL, $t_{1/2}$, $C_{\max}$, $T_{\max}$, etc. with fewer assumptions than other model-based approaches. The accuracy of non-compartmental analysis increases with more time points in concentration-time data. Even though non-compartmental analysis has less predictability of PK profiles with different dosing regimen that has not been investigated, this approach has frequently been applied for allometric scaling and is also acceptable for bioequivalence studies.

The majority of PK models are compartmental models which rely on linear or nonlinear differential equations derived from mass balance to describe drug kinetics. Instead of assuming the body as one homogenous compartment, compartmental modeling divides the whole body into several interconnected compartments with each consisting of organs or tissues that are kinetically homogenous. It also assumes that the rate of drug distribution between compartments follows first-order kinetics. This approach is critically important since it can predict concentration-time profiles of alternative dosing regimen from simulation. Thus this simple modeling technique is quite useful in PK/PD approach to relate drug response mechanism with drug delivery mechanism. The most widely used modeling technique in PK/PD approach is population modeling, which will be discussed with more details later in this chapter.

The biggest limitation of conventional compartmental models is that different compartments may not have a clear physiological significance but are abstract mathematical constructs. To overcome this limitation and extrapolate PK to different physiological conditions or alternatively to drugs has similar property; physiologically based pharmacokinetic (PBPK) models were developed based on actual physiological and biological meanings for drug and are expected to be a simple and direct approach to relate the observed drug response to target tissue exposure. Both organ physiology (weight, blood flow, enzyme expression, etc.) and drug-specific physical-chemical property (solubility, protein binding, tissue to plasma partition coefficient, etc.) are critical prior information for robust PBPK model development. A validated PBPK model can predict the

quantitative behavior of similar drugs or extrapolate PK with altered physiology which is beyond the range of investigated experimental conditions.

Types of Pharmacodynamic Models

Limited conclusions can be drawn from PK alone; PK/PD approach bridges the gap between the time course of drug concentration and therapeutic response. PD response usually has several dependent variables as clinical endpoints, surrogates, and biomarkers. In most cases, only one is modeled at a time. It is an important decision to choose an appropriate dependent variable which is meaningful, measurable, and appropriate for modeling. Dependent variables may be continuous (blood pressure), categorical (several score levels), or binary (alive, dead).

Continuous Response Variables

The simplest PD model is a direct response model where the effect is directly related to the concentration in central compartment. Theoretically, the therapeutic response is directly triggered by drug concentration at effect site. The direct correlation between response and plasma concentration may result from rapid drug distribution and instant pharmacological response or the equilibrium between plasma and effect site achieved at steady state (Derendorf et al. 2000). This type of model can often be expressed using Eq. 8.

A delayed effect may be observed when there is a temporal dissociation between the time course of blood/plasma concentration and the effect of the pharmacological agent. This dissociation may occur due to the delayed distribution between central compartment and target site, or indirect mechanisms such as time-consuming synthesis or degradation of endogenous substance, or a more complex receptor-mediated effect (Derendorf et al. 2000). Indirect models are often used to describe this type of model where the relationship between response and concentration is not one-to-one but rather shown in a counterclockwise hysteresis loop.

In case of delayed effect, it is often that either an effect compartment model or an indirect response model can be used to describe it. An effect compartment is a hypothetical compartment characterized by the time course of concentration at the effect site. It is linked to the PK compartment but with no intercompartment mass transfer. The indirect response mechanism can be described using indirect response models where the drug affects a precursor and subsequently influences the PD response. The rate of change of the response over time in the absence of drug can be expressed as:

$$\frac{dR}{dt} = k_{in} - k_{out}R \tag{9}$$

where R is the measured response, k_{in} is the apparent zero-order rate constant for response production, and k_{out} is the first-order rate constant for response dissipation. Four basic models were proposed that the drug can either inhibit or stimulate the production or loss of the response (Sharma and Jusko 1996). Both the inhibition and stimulation functions can be incorporated into k_{in} or k_{out} in Eq. 9 with expression of $\left(I(C_p) = 1 - \frac{I_{max} \bullet C_p}{IC_{50} + C_p}\right)$ for inhibition and $\left(S(C_p) = 1 + \frac{E_{max} \bullet C_p}{EC_{50} + C_p}\right)$ for stimulation.

One of the possible reasons for the lag of drug effects may come from signal transduction controlled by secondary messengers (Mager et al. 2003). Other mechanistic PD models based on irreversible effects (e.g., cell or target inactivation, enzyme inactivation) and tolerance mechanism (e.g., counter-regulation, precursor pool depletion) are also important components and can play a big role in PK/PD approach.

Noncontinuous Response Variables

The majority of PD models describe continuous response variables. However, clinical data are not always continuous but sometimes categorical variables such as the severity of a disease (ordinal scaled), or binary variables (dead, alive), or time-to-event (censored) data. Such responses are more clinically relevant to drug efficacy and safety and can be described as the probability of an event occur using logistic models or survival models.

Logistic regression is suitable to use for the prediction of probability change with predictors when the outcome is a binary response. The logit

transformation, the link function in logistic regression, is defined as the logarithm of the odds (the probability of an effect occurs divided by the probability of the effect not occur):

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = L(x) \tag{10}$$

where $L(x)$ is a linear function of predictors such as $L(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$ where x_i represents the predictors (e.g., drug concentration, exposure, time). The logit transformation expands the ends of scale by mapping the interval (0, 1) onto $(-\infty, \infty)$, such that the small difference in p would reflect as a larger difference on logit scale. Instead of just p, logit(p) is used as the response in this regression. Its inverse function, also known as the logistic function, is as following:

$$p(x) = \frac{\exp(\text{logit})}{1+\exp(\text{logit})} = \frac{\exp(L(x))}{1+\exp(L(x))} \quad \text{or}$$
$$p(x) = \frac{1}{1+\exp(-\text{logit})} = \frac{1}{1+\exp(-L(x))} \tag{11}$$

Both the logit and its inverse functions are monotonically increasing; the confidence interval for the probability of an event occurs (p) can be extrapolated from the confidence interval for logit (p). The application of logistic regression can also be extended in cases where the dependent variable has more than two outcome categories or when the multiple categories are ordered.

For the analysis of time-to-event data (such as time from origin of treatment to disease progression or duration of survival), they are often "censored" due to loss of follow-up (right censoring) or delayed entry (left censoring). Survival analysis aims to avoid this bias of incomplete longitudinal data. The survival function is defined as the probability that the event has not occurred by certain duration such that $S(t) = \Pr(T > t)$, where T denotes the time of the event occur. Also, notice that $F(t) = 1 - S(t)$ where $F(t)$ is the cumulative distribution function; this leads to the relationship of $-S'(t) = f(t)$, the probability density function of the lifetimes T. The hazard function $h(t)$ is the probability of an event occur immediately following time t given that the event has not occurred at time t:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T < t + \Delta t | \, T > t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{12}$$

The cumulative hazard $H(t)$ is the area under the hazard function which relates to the survival function in the relationship as:

$$H(t) = -\ln S(t) \;\text{ or }\; S(t) = \exp[-H(t)] \tag{13}$$

The survival function $S(t)$ is monotonically decreasing, while the cumulative hazard function $H(t)$ is monotonically increasing. When no event time is censored, the survival and hazard can be easily estimated by assuming parametric survival distributions (e.g., exponential, Weibull, log-logistic, etc.)

In case of data with right censoring or left truncation, nonparametric method should be used. Kaplan-Meier estimator and Nelson-Aalen estimator provide nonparametric estimates for survival function and cumulative hazard rate function, respectively. The Cox regression model, also known as the proportional hazards regression, is a semi-parametric model with an unspecified baseline hazard function (nonparametric) and a parametric component:

$$h(t|X) = h_0(t)\exp(\beta X) \tag{14}$$

where $X = (X_1, \ldots, X_n)$ is the covariate vector and $\beta = (\beta_1, \ldots, \beta_k)'$ is the coefficient vector with each β_i the log hazard ratio associated with one-unit change in X_i with the rest of covariates remaining constant. The baseline hazard only depends on t but not on any covariates X_i, while the hazard ratio $\exp(\beta X)$ only depends on the covariates but not on time t (independent of time, so-called proportional hazard). However, time-dependent covariates are more commonly seen in pharmacometrics setting since the hazard

is associated with drug concentration which varies over time. In this case, the Cox regression model can still be used but the hazard ratio depends on time as well:

$$h(t|X) = h_0(t)\exp(\beta X(t)) \qquad (15)$$

Gieschke et al. applied both logistic regression and Cox regression model to explore the relationship between systemic capecitabine exposure and its safety/efficacy outcomes (Gieschke et al. 2003). Holford discussed the details of the link between basic concepts of PK/PD and time-to-event analysis and how this approach can reveal more information for the prediction of therapeutic effects (Holford 2013). Gong et al. proposed the application of machine learning method for time-to-event data analysis, and their results indicated that machine learning-based methods provide better performance than the traditional Cox model (Gong et al. 2018).

Evaluation

Plotting effects versus various covariates like time, concentrations, and demographic variables will help to generate hypothesis about a future model. Is the effect time invariant? Is a hysteresis observed? It is Clockwise or counterclockwise? A strategy to decide about using an effect compartment or a direct response model has been discussed by Felmlee et al. (2012). Knowledge about the underlying process can also help to decide on this question and about the related physiology and will also help to explore more mechanistic models.

The following items should be clearly addressed in a PK/PD model development:

1. Problem and purpose of the model
2. Assumptions in the modeling process (explicit and implicit assumptions)
3. Rationale of model development (Why is a model considered to be better than another?)
4. Validation strategy using internal or external data

General Approaches of Pharmacometrics

All models are mathematical representations of the data. Often, the main objective of developing a PK/PD model is to describe the data, to predict unknown situations, and to explain underlying mechanisms. The application of using mathematics and statistics to understand data during drug development originates from simple descriptive summary of data. With the fast development in this field for the last 30 years, a lot of improvements have been seen in model efficiency and capability, especially after the inception of population modeling.

Traditional Approach

The simplest modeling approach to evaluate data from multiple subjects or animals is "naïve pooled approach," where data from all individuals are pooled first and then fit. The mean response estimated using this approach is always biased, and the interindividual differences in exposure and response are also ignored. Consequently, this is rarely used for PK/PD data analysis. Before the inception of population approach, the traditional method that modelers used is "two-stage approach." The individual's data are fit separately first. Subsequently, individual parameter estimates are combined, and descriptive summary statistics are calculated including mean, variance, and covariates on each parameter. The mean estimates of parameters obtained from this approach are generally good, but the estimates of interindividual variability are biased and imprecise (Sheiner and Beal 1980). In addition, the two-stage approach highly relies on the richness of data for each individual. Hence this approach may not be applicable in the situation of sparse data where individual parameters are not easy to estimate. To overcome the limitation of both earlier approaches, Sheiner et al. developed a new approach which allowed dealing with sparse data to estimate population mean parameters and interindividual variability and incorporate covariate effects (Sheiner et al. 1972). With the advancement in computing power, this valid and robust

population approach has been widely used in PK/PD modeling to facilitate current drug development.

Population Approach

The population approach, also known as the mixed-effect modeling, consists of both fixed effect parameters (population mean values) and random effect parameters (variability within the population). In general, all PK and PD parameters have certain distributions in population. Nonlinear mixed-effect modeling can simultaneously calculate parameters and their distributions from the full set of individual data. In this way, information from each individual are gathered and contributed to the covariate determination and the corresponding variability quantification. A general population model consists of three components: a structural model, a covariate model, and a stochastic model (Mould and Upton 2012). Structural model adopts the classical compartmental model to describe the time course of concentration profile (PK) or measured response (PD). Covariate model characterizes the relationship between PK/PD response to demographic covariates (such as weight, height, gender, etc.). The covariate identification and covariate model development are very critical as it supports labeling for special population based on their demographic information such as kidney/liver function, metabolic status, etc. Stochastic model describes the unexplainable variability in the data which includes inter- and intraindividual variability, residual, etc. More detailed discussion of population PK can be found in a separated chapter in this book. The application of population approach in PK/PD modeling provides clinical pharmacologist with better understanding of underlining mechanisms and assists rational dose adjustment for subpopulations of patients.

Learning and Confirming Circle

Drug development usually involves several iterations of model-informed learning and confirming where learning answers "how much/what" questions and confirming answers "yes/no" questions. Classical clinical studies are self-consistent, that is, a hypothesis is accepted or refused with the information coming from the study and nothing else. The learn and confirm approach (Sheiner 1997) consists of alternating learning and confirming cycles: a study to generate a hypothesis and a subsequent study to confirm the hypothesis (or to improve the model), followed by further hypothesis generating based on the previous finding and confirming (or not confirming) studies. Phase 1 and Phase 2B/3 are considered as learning stages, while Phase 2A and Phase 3/4 are seen as confirming stages based on the objectives in each stage. However, learning is still a very important subsidiary in the confirmation stages since learning while confirming would help keep the knowledge updated for confirming questions.

Software

Given that pharmacometrics is an interdisciplinary science fused with pharmacology, physiology, computer science, and mathematical/statistical modeling, PK/PD modeling and simulation heavily relies on software to make clinical trials more efficient. With the rapid evolving computation efficiency, a variety of efficient, flexible, and user-friendly software are available recently. For non-compartmental analysis, WinNonlin (Phoenix) has been widely used in pharmaceutical industry. Other software such as PK packages in R, Kinetica, Scientist, and PKSolver can also perform non-compartmental analysis. As for nonlinear mixed-effect modeling, NONMEM was the first and most commonly used software for population PK and PK/PD modeling and simulation. Ever since the release of the first version, NONMEM has been continuously updated with new statistical methods and estimation algorithms. Because NONMEM does not have a very user-friendly interface, "front-end" and "back-end" software (such as Pirana, R, SAS) have been incorporated with NONMEM to overcome this limitation. Other software which has been widely used for mixed-effect modeling include MONOLIX, a software based on the stochastic approximation expectation maximization (SAEM) algorithm for reliable convergence and ADAPT, which is based on Monte Carlo

parametric expectation maximization. Mixed-effect modeling is also possible in R with packages such as NLME, RxODE, mrgsolve, saemix, etc. The Pmetrics package in R allows nonparametric modeling of population PK using nonparametric adaptive grid (NPAG) algorithm. In addition, packages like PKgraph and ggplot provide graphical user interface for model diagnosis. The biggest advantages of R are the various statistical packages available and the possibilities of data management and visualization and that many scientists working in population PK/PD are programming with R and are publishing solutions for various problems. Berkeley Madonna also serves as a fast ordinary differential equation solver with a visualization interface. Besides NONMEM, Bayesian pharmacometrics modeling can also be performed in BUGS and Stan. SimCyp, PK-Sim, and GastroPlus are commonly used for PBPK modeling. Versions and owners of software change rapidly. Actual information about these commercial or free software packages can be found on the corresponding web sites on the Internet.

Critical Assessment of the Method

PK/PD Concepts in Antimicrobials

Establishing PK/PD relationship of drug candidates is critical during the whole process of drug development. Bridging the PK/PD information from preclinical to clinical studies and optimizing dosage regimen for special population is very crucial to ensure drug efficacy and to minimize toxicity. In this section, we will introduce important concepts of PK/PD approach in antimicrobial drug development since the PK/PD approach in this area has been well developed. The determination of correct dose and dosing interval can be evaluated via PK/PD approach (Sy et al. 2016). The discussion of PK/PD approach in antibiotic dose optimization is mainly separated into three approaches, namely, the minimum inhibitory concentration (MIC)-based, the time course-based, and *in vivo* animal model-based approaches. The PK/PD concepts and theories presented for antibiotics are also broadly applicable to drug development in other therapeutic areas.

MIC-Based Approaches

MIC, defined as the minimum inhibitory concentration, is a PD surrogate index of the susceptibility of certain pathogen in the presence of a specific antimicrobial treatment. This simple index does not reflect the time course of pharmacological effect but provides very useful information on antibiotic efficacy. From *in vitro* experiment, MIC can be easily determined as the lowest antibiotic concentration that inhibits the visible growth of microorganism in the end of a 16–20-h incubation period (Sy and Derendorf 2016). The most common MIC-based PD indices (Fig. 2) that quantitatively link the drug exposure and microbiological outcomes for *in vivo* efficacy prediction are the percentage of a dosing interval in which the drug concentration is above the MIC ($T > MIC$, time above MIC), the ratio between the peak concentration and MIC (C_{max}/MIC), and the 24-h area under the concentration-time curve divide by MIC (AUC/MIC). Since only the unbound proportion of drug is pharmacologically active, an italicized prefix *f*, implying the free drug concentration, is often seen with these indices.

The PK/PD index that associated with specific antibiotic agent is determined from either the dose fractionation study in the rodent or the time-kill kinetic studies (Sy et al. 2016). The microbiological outcome, which is quantified by the change in log10 colony-forming unit (CFU), usually correlated very well with at least one of these MIC-based PD surrogate indices. If the PK/PD relationship is best characterized by $T > MIC$, implying time-dependent bacterial killing, i.e., once the free drug concentration is above the MIC, increasing the antibiotic concentration will not further increase the killing rate, but extending this period would ensure drug efficacy. One classic example of "time-dependent" killing is β-lactam antibiotics as the efficacy is enhanced with longer exposure times at which the free concentration maintained above MIC. In the case of antibiotics exhibiting concentration-dependent killing, the antimicrobial killing rate increases as the concentration of antibiotic increases. Therefore, the aim for this

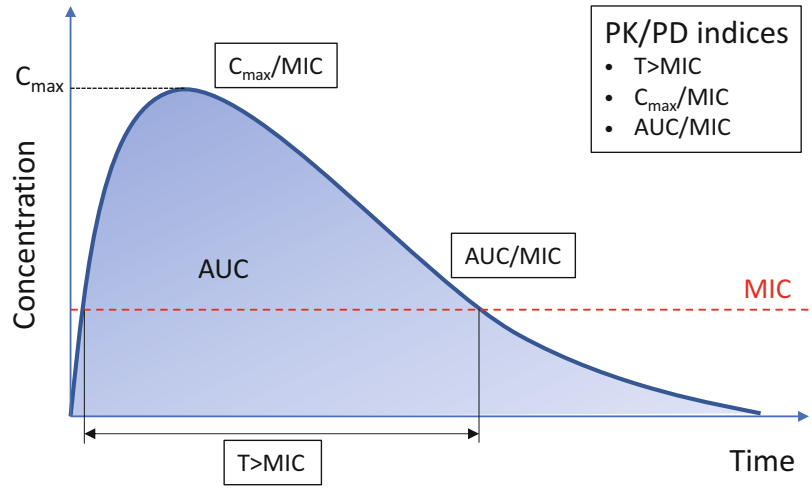


Fig. 2 Pharmacokinetic-pharmacodynamic indices as targets for achieving antimicrobial efficacy. T > MIC, the percentage of a dosing interval in which the drug concentration is above the MIC; C_{max}/MIC, the ratio of the maximal drug concentration to the MIC; AUC/MIC, the ratio of the area under the drug concentration-time curve to the MIC

type of antibiotic is to attain the magnitude of fC_{max}/MIC ratio, which is proportional to the initial killing rate. The killing pattern of aminoglycosides correlates well with this ratio. For drugs which have concentration-independent killing but with persistent effects after drug exposure, their killing can typically be characterized by fAUC/MIC ratio. The bacterial killing rate for this type is related to both the duration when free drug concentration is above the MIC and the total exposure of the antibiotic. A good example that belongs to this group is vancomycin due to its extended postantibiotic effects that inhibit bacteria regrowth even free drug concentration dropped below MIC.

Time Course-Based Approaches

One of the limitations with the simple MIC-based approach is that the complex interactions between the host, the pathogen, and the drug itself are difficult to be captured by this static *in vitro* parameter. Different combinations of bacteria growth and killing kinetics in response of antimicrobial agent may lead to the same MIC in the end of incubation period. In addition, the MIC determined by twofold dilution is a relatively crude index, rather than an accurate estimation. Furthermore, the antibiotic concentration can remain constant during the time course of *in vitro* MIC measurement, or keep changing to mimic the dynamic change of free concentration at the actual target site *in vivo* . Therefore, another approach that provides more detailed PK/PD information to evaluate time course of bacterial response with dynamic exposure of the anti-infective agent is the *in vitro* time-kill assays.

The time-kill experiments can be either static (fixed antibiotic concentration) or dynamic (constantly changing antibiotic concentration). During static time-kill study, a sample was taken at each prespecified time, and bacterial colonies were counted after incubation. Although this experiment is simple, static studies are labor intensive as the drug concentration ranges from 0.25-fold up to 16-fold based on twofold dilution, and the duration of each experiment varies from 6 to 72 h (Sy et al. 2016). For dynamic time-kill study, the drug concentration in the medium is changing by replacing with the fresh ones without antibiotic contained, rather than having the same medium throughout the experiment. The dynamic time-kill experiments are usually conducted via a hollow-fiber system where the model is continuously flushed with the fresh medium using a flow rate which is similar as the half-life of the investigated agent. Therefore, the *in vivo* drug kinetics can be simulated by this *in vitro* system.

Animal Models

The *in vitro* setting conditions, for either MIC determination or time-kill assays, are still highly distinct from the kinetic situation *in vivo* at the site of infection. Instead of aerobic and protein-free *in vitro* environment, the *in vivo* condition is rather

anaerobic, acidic, and prone to protein binding (Levison 2004). Also, one of the drawbacks of hollow-fiber model is that the diffusion blockage may occur due to bacteria clusters adhering along the capillary walls, which may ultimately alter the flow rate in the system (Sy and Derendorf 2016). To overcome the limitation of *in vitro* PK/PD evaluation, such as lack of immune system and limited nutrients available, *in vivo* evaluation of PK/PD relationship is also feasible. One possibility would be to use animal models such as rats or mice.

Drug distribution at the target site are most commonly studied in thigh or lung infection models usually carried out in neutropenic rodents to avoid the variability of different immunity levels. Neutropenia can be induced by administration of cyclophosphamide (Zuluaga et al. 2006). Due to faster drug elimination in animals than humans, uranyl nitrate administration was performed before treatment to induce transient renal impairment, therefore delaying drug elimination in animals (Craig et al. 1991). PK samples were collected using microdialysis technique to measure the unbound drug concentration in the ISF at the target tissue. The bacterial density change from the starting inoculum was evaluated at the end of treatment period. These PD results were correlated with the PK/PD indices (i.e., $fT > MIC$, $fC_{max} > MIC$, and $fAUC/MIC$) in dose fractionation studies. Animal models provide a more pertinent approach to evaluate humanlike PK/PD; however, one of the biggest disadvantages is this procedure is not able to mimic human PK for drugs with extensive hepatic metabolism (Sy et al. 2016).

Model-Based Drug Development

Establishing PK/PD relationship using modeling and simulation is very critical for dosing regimen selection of antimicrobial agents, as the delineation of such relationship can greatly help selection of dosing regimen that has a high probability to overcome bacterial resistance and optimize clinical outcome (Drusano 2004). These model-based approaches provide useful information on the PK/PD index that best characterize the antimicrobial activity, as well as the translational value based on *in vitro* time-kill kinetics or animal PK data. In this section, we will discuss and illustrate the application of pharmacometrics in antimicrobial drug development. For MIC-based PK/PD indices, the dose and dosing interval can be determined through simulation of human PK and the desired target value of PK/PD index. When it comes to time-kill-based or animal PK/PD data, different dosing regimens can be evaluated via simulation of bacterial responses based on semi-mechanistic models.

Monte Carlo PK/PD Simulations

Given robust descriptive PK model and PD index determined from preclinical experiments, simulation of virtual clinical trials with different dosing regimens and targets can be done for dose optimization. Monte Carlo simulation is a computer-based mathematical technique that helps answer many "what if" questions before conducting expensive clinical trials in patient population (Roberts et al. 2011). Utilizing results from PK/PD simulation provides better confidence to the dosing regimen selected for clinical trials. The major components of Monte Carlo simulation include essential PK parameters and their corresponding interindividual variability, predefined antibiotic PD index, or parameters from semi-mechanistic PD models. For MIC-based approach, the simulated individual concentration-time profile can be evaluated against the prespecified PK/PD index. If PK/PD model was used for simulation, optimal dose selection would be based on simulated data of bacterial response over time. However, simulation results may be invalid due to potential confounding factors or small sample size in PK data used for model development, or model development is based on total concentration rather than free concentration. In these situations, the simulation results should be interpreted with caution.

Probability of Target Attainment

For MIC-based approach, the percentage of the virtual subjects that achieve the target PK/PD index under certain dosing regimen can be

computed from simulated individual PK profiles. This is defined as the probability of target attainment (PTA). PTA is usually determined from 1000 to 10,000 individual concentration-time profiles simulated from population PK model, considering the interindividual viability. Under each dosing regimen, the likelihood of achieving target attainment at a prespecified MIC value was calculated based on the distribution of PK/PD index (e.g., $fT > MIC$, fC_{max}/MIC, $fAUC/MIC$) from simulated PK profiles. This specific target is often associated with 1- or 2- log10 reduction of bacterial count from animal studies. After repeating the above likelihood calculation with a range of increasing MIC values, PTA and MIC can be plotted with x-axis as the MIC values and y-axis as the probability. This plot indicates the trend of less probability of successful antimicrobial achievement with increasing MIC. A probability greater or equal to 90% is usually accepted.

To illustrate the application of this approach, the example used here is from Singh and his colleagues (2017). PTA analysis of tigecycline was performed on traditional PK/PD target AUC/MIC >6.96 h and the new target of $fAUC/MIC$ >2.05 h, accounting for atypical nonlinear plasma protein binding. The simulation was based on 10,000 individuals with body weight, creatinine clearance, and gender incorporated into the individual clearance. Individual AUC at steady state for dosing intervals ($AUC_{ss(0–24h)}$) was calculated from daily doses of 100, 150, 200, 250, and 300 mg intravenously. At each MIC level in a range of 0.064–64 mg/L, individual AUC/MIC ratio was calculated using each $AUC_{ss(0–24h)}$ divided by the corresponding MIC value. If the virtual subject has an AUC/MIC ratio greater than 6.96 h, the clinical outcome on this subject is considered as successful. Subsequently, the overall percentage of virtual subjects was computed at each MIC level under each dosing regimen (Fig. 3). Applying similar approach but combined with a protein binding model, $fAUC/MIC$ ratio of each individual was calculated, and the percentage of virtual subjects achieving $fAUC/MIC$ at least 2.05 h at different MICs with different tigecycline daily doses is also shown in Fig. 3.

If the target is based on AUC/MIC, a significant increase of target achievement is clearly shown with increasing daily dose. In contrast, dose increment does not significantly alter target attainment if it is based on $fAUC/MIC$. This finding implies the importance of using free concentration and exposure in PK/PD index for drug efficacy and the impact of protein binding on clinical breakpoints. PTA analysis was also used in this example to evaluate the effect of plasma protein binding on tigecycline clinical breakpoint selection. As Fig. 4 shows, the PTA analysis results at a daily dose of 100 mg; the clinical breakpoint for tigecycline against *E. coli* is 0.5 mg/L without consideration of protein binding but is 0.25 mg/L regarding the free AUC. Since clinical breakpoint is a very critical criterion to stratify patients into different susceptibility phenotypes, ignoring the protein binding could put more patients into risk of clinical failure.

Another approach to utilize Monte Carlo simulation results for clinical outcome prediction is to compute the expected PTA for a given microorganism population with a specific dosing regimen (Asin-Prieto et al. 2015; Mouton et al. 2005). This is the so-called cumulative fraction of response (CFR), which can be calculated using the previous PTA values and MIC distribution in the equation of $\left(\mathrm{CFR}(\%) = \sum_{i=1}^{n} \mathrm{PTA}_i \times F_i\right)$, where PTA_i is the PTA at specific MIC and F_i is the bacterial isolate frequency at that corresponding MIC level. In this case, CFR is particularly useful when the exact pathogen susceptibility is unknown because the success of clinical outcome is predicted based on the whole population rather than a single MIC value. The population MIC distribution can be obtained from European Committee on Antimicrobial Susceptibility Testing (EUCAST) or from a particular healthcare facility. It is important to notice that the pathogen susceptibility can vary between different locations and over time (Roberts et al. 2011; Asin-Prieto et al. 2015). Therefore, CFR is often specific as its calculation is based on the MIC distribution in a specific facility at a particular time.

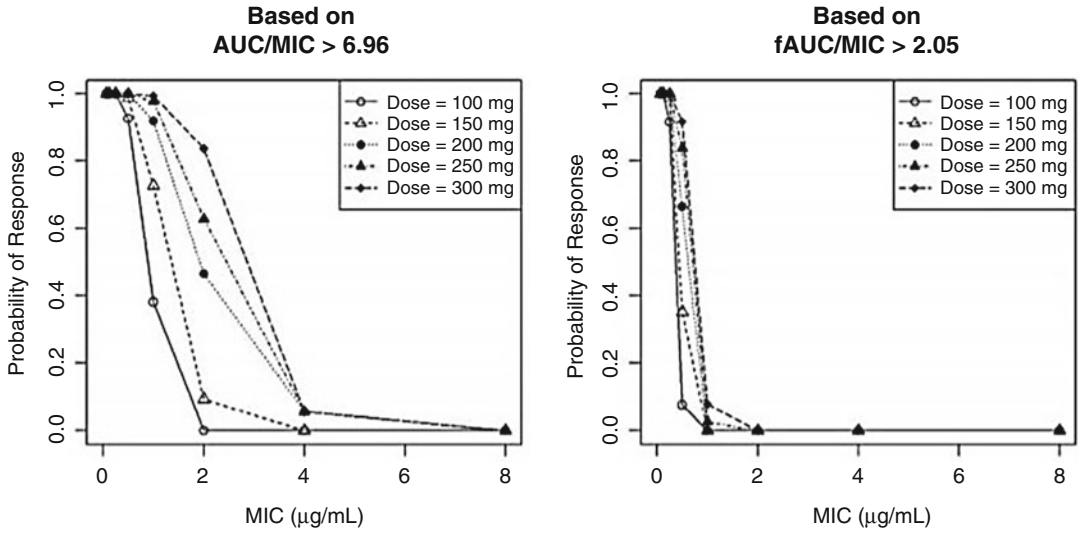


Fig. 3 Probability of target attainment as a function of minimum inhibitory concentration (MIC) of tigecycline against *Escherichia coli* at different doses with target AUC/MIC >6.96 h (left panel) and with target *f* AUC/MIC >2.05 h (right panel). (Image adapted from Singh et al. (2017) and used with permission)

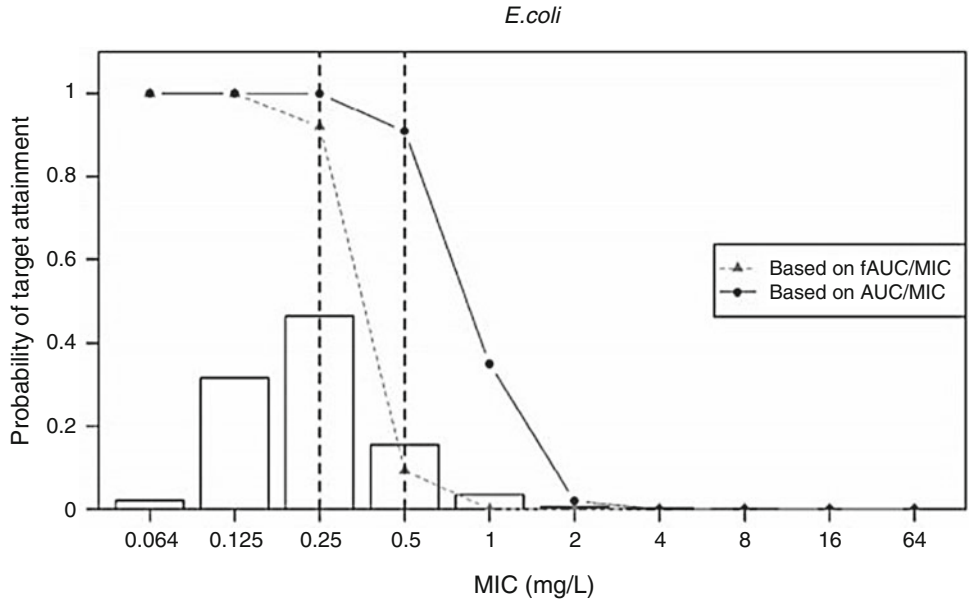


Fig. 4 Probability of target attainment at different minimum inhibitory concentrations (MICs) of tigecycline against *Escherichia coli* with total AUC/MIC (target 6.96 h) and ƒAUC/MIC (target 2.05 h). The distribution of MICs was obtained from the European Committee on Antimicrobial Susceptibility Testing (EUCAST). (Image adapted from Singh et al. (2017) and used with permission)

Semi-mechanistic PK/PD Model

If the PD data is from time-kill curve experiments, semi-mechanistic PK/PD model is commonly used to establish the relationship between the bacterial colony population change and antimicrobial concentration over time. The information used for PD model development is often borrowed from static time-kill

experiments, while the dynamic time-kill data provides additional information on PK/PD relationship and thus can be used for model validation. Rather than complicated mechanistic models considering the bacterial growth cycle, states of bacterial susceptibility, drug-receptor information, and the mechanisms of drug action, semi-mechanistic models are empirical models with simpler terms but still able to capture the bacterial response on antibiotics and monitor the development of bacterial resistance (Sy et al. 2016).

In the absence of antimicrobial intervention, the population number of bacteria in an inoculum over time represents the net result of bacterial natural self-replication and degradation (Jusko 1971):

$$\frac{\mathrm{dN}}{\mathrm{dt}} = k_{\mathrm{growth}}N - k_{\mathrm{death}}N \tag{16}$$

where N is the bacterial count and k_{growth} and k_{death} are the first-order rate constant for bacterial growth and death, respectively. This model assumes constant growth and death rate; however, it is difficult to separately define each constant as the observations are often based on $k_{\mathrm{net}} = k_{\mathrm{growth}} - k_{\mathrm{death}}$. Also notice that in both static and dynamic time-kill curve studies, bacteria grow exponentially but then reach a plateau or stationary level when the net growth is zero (Nielsen et al. 2011b). A logistic growth model (Tam et al. 2005) can better describe this self-limiting growth pattern:

$$\frac{\mathrm{dN}}{\mathrm{dt}} = k_{\mathrm{net}}\left(1 - \frac{N}{N_{\max}}\right)N \tag{17}$$

where $N_{\max}$ is the carrying capacity or the maximum population size in the system. From the analytical solution of Eq. 17, one important characteristic of this model is that bacteria population approaches the carrying capacity as time goes to infinity, i.e., $\left(\lim_{n\to\infty} N(t) = N_{\max}\right)$. In other words, bacterial population achieves stationary phase when N approaches $N_{\max}$.

When the *in vitro* system is exposed with antimicrobial agents, the drug effect can be incorporated into the logistic growth model such that:

$$\frac{\mathrm{dN}}{\mathrm{dt}} = k_{\mathrm{net}}\left(1 - \frac{N}{N_{\max}}\right)N - f_{\mathrm{kill}}(\mathrm{drug}) \tag{18}$$

where the f_{kill}(drug) is the drug effect (Nolting et al. 1996; Mouton and Vinks 2005; Liu et al. 2005; Treyaprasert et al. 2007), which is often represented by a sigmoid $E_{\max}$ or simple $E_{\max}$ model as

$$f_{\mathrm{kill}}(\mathrm{drug}) = \frac{E_{\max}C^{\gamma}}{\mathrm{EC_{50}}^{\gamma} + C^{\gamma}}N \tag{19}$$

where C is the drug concentration, $E_{\max}$ is the maximum killing effect, and $\mathrm{EC_{50}}$ is the concentration at which half of the maximum drug effect is achieved. The shape parameter γ, also known as the Hill factor, is 1 in simple $E_{\max}$ model.

Modification on Eq. 18, such as incorporating an adaptation or delay function, allows the logistic growth model to adapt to bacterial regrowth. With an adaptation function on $\mathrm{EC_{50}}$, the drug effect is modified as $\left(f_{\mathrm{kill}}(\mathrm{drug}) = \frac{E_{\max}C^{\gamma}}{(\alpha \bullet \mathrm{EC_{50}})^{\gamma} + C^{\gamma}}N\right)$, wherein $\alpha = 1 + \beta(1 - e^{-C\tau t})$ with β the maximal adaptation and τ the rate of adaptation factor (Tam et al. 2008). By using this adaptation function, the decline in the kill rate over time was successfully modeled. Subsequently, the model can well predict the microbial responses to both gentamicin and amikacin. A delay term, usually expressed as $1 - e^{-\mathrm{kt}}$, ranges between 0 and 1 from time zero to infinity. This function can be incorporated on the bacterial growth phase and/or anti-infective drug effect (Nolting et al. 1996; Liu et al. 2005; Treyaprasert et al. 2007). Treyaprasert et al. tried different types of delay function in PD model of azithromycin against four bacterial strains and found that incorporating delay term on both term can best describe the antibiotic response (Treyaprasert et al. 2007).

An alternative method to describe the self-limiting growth capacity is to implement the idea of bacterial phenotypic switching between

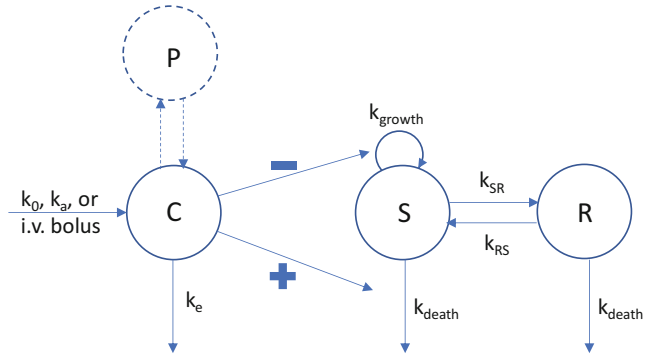


Fig. 5 Schematic illustration of the pharmacokinetic/pharmacodynamic model. *C* central compartment; *P* peripheral compartment; *S* proliferating and drug-sensitive bacteria; *R* resting and drug-insensitive bacteria; k_0 drug infusion rate constant; k_a drug absorption rate constant; k_e drug elimination rate constant; k_{growth} and k_{death} rate constants for multiplication and degradation of bacteria, respectively; k_{SR} rate constant for transformation from the growing, sensitive stage to the resting stage; k_{RS} rate constant for transformation from the resting stage to sensitive stage

susceptible (S) normally growing cells and persistent resting (R) cells with reduced growth rates (Balaban et al. 2004). In this case, the total number of bacteria would be the sum of both subpopulations (S + R) as it is illustrated in Fig. 5. In this model, assumptions were made as the majority of the bacteria in the early growth phase are in susceptible state and bacteria in persistent state do not respond to antimicrobials. As the bacteria population in the system increases, bacteria in growing stage are gradually transformed into a resting stage, leading to the stationary phase. The kinetic behavior of each subpopulation without antibiotic exposure can be described in the following equations:

$$\frac{\mathrm{dS}}{\mathrm{dt}} = k_{\text{growth}}S - k_{\text{death}}S - k_{\text{SR}}S + k_{\text{RS}}R \quad (20)$$

$$\frac{\mathrm{dR}}{\mathrm{dt}} = k_{\text{SR}}S - k_{\text{RS}}R - k_{\text{death}}R \quad (21)$$

where k_{SR} and k_{RS} are the transfer rates between susceptible population and resting population. Since persistent population is unlikely to return it back to susceptible state, thus the term k_{RS} can often be set as 0. The antimicrobial effect, with same sigmoid $E_{\max}$ model in Eq. 19, can either decrease bacterial growth rate or increase bacterial death rate and incorporated as an additive or proportional effect:

$$\frac{\mathrm{dS}}{\mathrm{dt}} = k_{\text{growth}}(1 - f(\text{drug}))S - k_{\text{death}}S - k_{\text{SR}}S + k_{\text{RS}}R \quad (22)$$

$$\frac{\mathrm{dS}}{\mathrm{dt}} = k_{\text{growth}}S - (k_{\text{death}} + f(\text{drug}))S - k_{\text{SR}}S + k_{\text{RS}}R \quad (23)$$

$$\frac{\mathrm{dS}}{\mathrm{dt}} = k_{\text{growth}}S - k_{\text{death}}(1 + f(\text{drug}))S - k_{\text{SR}}S + k_{\text{RS}}R \quad (24)$$

Nielsen et al. used this semi-mechanistic model to simultaneously fit *in vitro* static and dynamic time-kill data of *Streptococcus pyogenes* exposed to five different antibiotics (Nielsen et al. 2007, 2011b). The model was modified with an addition of an effect compartment to describe the time delay of drug effect. The investigators later on extended this model with a binding model implemented to describe the on and off adaptive resistance for gentamicin (Nielsen et al. 2011a). Simulation was performed on this model. The dose fractionation study indicated that the PK/

PD indices can be identified from *in silico* predictions based on this semi-mechanistic PK/PD model. This study implied the power of applying PK/PD model derived from *in vitro* studies to describe the antimicrobial effect and select the optimal dosing regimen for clinical studies.

Models of Combination Therapies

Combination therapy, a drug intervention consisting of at least two therapeutic agents for the treatment of same condition, has gain popularity in recent years to overcome the emergence of bacterial resistance. The effect of combination therapy can also be characterized by logistic growth model utilizing Loewe additivity to evaluate synergistic effect of drug combination when each component has its intrinsic antimicrobial activity (Greco et al. 1995). Under the assumption that each therapeutic agent cannot interact with itself, the $E_{\max}$ model to evaluate the drug combination of two agents can be expressed as:

$$E_{\max} = \frac{k_{\max}\left(\frac{C_1}{\alpha_1 \mathrm{EC}_{50,1}} + \frac{C_2}{\alpha_2 \mathrm{EC}_{50,2}} + \frac{\gamma C_1 C_2}{\alpha_1 \alpha_2 \mathrm{EC}_{50,1} \mathrm{EC}_{50,2}}\right)^k}{1 + \left(\frac{C_1}{\alpha_1 \mathrm{EC}_{50,1}} + \frac{C_2}{\alpha_2 \mathrm{EC}_{50,2}} + \frac{\gamma C_1 C_2}{\alpha_1 \alpha_2 \mathrm{EC}_{50,1} \mathrm{EC}_{50,2}}\right)^k} \tag{25}$$

where $k_{\max}$ is the initial killing rate, α_i refers to the same adaptation function defined previously, and γ indicated the Loewe synergism ($\gamma > 0$) or Loewe antagonism ($\gamma < 0$). A semi-mechanistic PK/PD model incorporating Loewe additivity has been used to successfully describe the combination effect of vertilmicin (an aminoglycoside) and ceftazidime (a β-lactam) against *Pseudomonas aeruginosa* (Zhuang et al. 2015). In this case, both vertilmicin and ceftazidime have antimicrobial activities with their own mechanism of action. Characterizing the potential synergistic effect in PK/PD modeling and simulation would assist dose selection of this therapeutic combination.

When it comes to the combination of β-lactam (BL) and β-lactamase inhibitor (BLI), such as avibactam and aztreonam, ceftazidime and avibactam, and meropenem and vaborbactam, the BLI has limited intrinsic antimicrobial activity but restores the antimicrobial activity of BL, leading to an enhanced spectrum of activity with this type of drug combination. Given that the microorganism's susceptibility to BL enhanced with increasing concentration of BLI and the degradation of BL also depends on bacterial density, it is reasonable to incorporate a drug degradation model into the semi-mechanistic PK/PD model. Sy et al. have developed such models for ceftazidime/avibactam combination to predict the exposure-response behavior of both agents (Sy et al. 2017, 2018). The schematic representation of the model structure is illustrated in Fig. 6.

Similar as other semi-mechanistic models of logistic growth equation, the antimicrobial effect of BL can be described in a sigmoidal $E_{\max}$ model. However, instead of using a simple

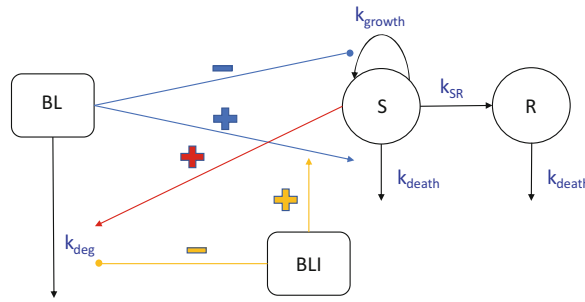


Fig. 6 Schematic representation of the combined effects of β-lactam antibiotic (BL) and β-lactamase inhibitor (BLI) on the growth and killing of pathogens in a closed system including bacterial density-dependent BL degradation. (Image adapted from Sy et al. (2018) and used with permission)

EC_{50} parameter for drug potency, a bi-exponential function of BLI concentration is applied in Eq. 19 such that BL EC_{50} shifts toward lower values with increasing BLI concentrations. Therefore, the function of drug effect has been modified to:

$$f_{\text{kill}}(\text{drug}) = \frac{E_{\max} \bullet BL^{\gamma}}{(A\mathrm{e}^{-\alpha \bullet \text{BLI}} + B\mathrm{e}^{-\beta \bullet \text{BLI}})^{\gamma} + BL^{\gamma}} N \tag{26}$$

where A and B are constants with their sum determines the initial value of EC_{50} when BLI is absent. α and β are the exponents that describe the relationship of BLI concentration and BL potency. The bi-exponential model was used since BL MIC drops rapidly from monotherapy to combination therapy with even the lowest BLI concentration, and then BL MIC continues decreasing but in a slower pace as BLI concentration increases. Given that MIC correlated to EC_{50} in the logistic growth model (Mouton and Vinks 2005), this bi-exponentially decreasing function is suitable to model the mechanism that BLI restores BL susceptibility to bacteria. As previously discussed, a delay function can also be implemented on the growth and/or kill term in this semi-mechanistic model to modulate the initial decline of bacteria and delay in later regrowth in the presence of BLI.

During the time-kill studied by Sy et al., degradation of both BL and BLI was also monitored. A drastic degradation rate was seen when bacterial population is high in the system and BL degraded in a concentration-dependent manner of BLI. Hence, a saturable Michaelis-Menten-type equation was applied to BL degradation model with degradation rate depending on bacteria population as well as the concentration of both BL and BLI:

$$\frac{\mathrm{dBL}}{\mathrm{dt}} = -\frac{\mathrm{Deg}_{\max} \bullet N^{\varphi}}{K_m{}^{\varphi} + N^{\varphi}} \left(1 - \frac{\mathrm{BLI}}{\mathrm{IC}_{50} + \mathrm{BLI}}\right) \mathrm{BL} \tag{27}$$

where $\mathrm{Deg}_{\max}$ is the maximum degradation rate constant, K_m is the bacteria density at which degradation rate is half of the maximum value, φ is the Hill coefficient that determines the shape of the function, and IC_{50} is the BLI concentration that yields 50% decrease of the BL degradation rate.

By applying this semi-mechanistic model, considering bacteria-mediated BL degradation and inhibition of degradation of BLI, Sy et al. were able to develop the PK/PD model from static time-kill data to simultaneously describe the dynamic change of multidrug resistant *Pseudomonas aeruginosa*, ceftazidime degradation, and the inhibition effect from avibactam (Sy et al. 2018). This model was further validated using dynamic time-kill data as well as data from animal models. By incorporating the mechanism of drug resistance, the model was able to give more detailed prediction of the bacterial dynamics in response to BL/BLI combination. It has the potential to expedite the BL/BLI combination drug development by confident simulation of clinical trials.

Special Population

Since most of the time, the initial PK model development is based on data from healthy adult volunteers in Phase 1. Drug PK may not be the same for patients or special populations such as obese population, geriatric population, pediatric population, etc. due to the different ontogeny, physiology, and pathophysiology conditions in these special populations. Elderly patients are particularly subject to drug toxicity including antibiotics due to their diminished physiology reserve and the frequent polypharmacy, therefore requiring dose adjustment of some antimicrobial agents (Benson 2017). For pediatric population, allometric extrapolation of clearance can give reasonable pediatric dose for similar exposure in children if using a drug-specific allometric exponent rather than using the fixed exponent of 0.75 (Mahmood 2007). It was also found that an age-dependent exponent in allometric scaling model can well predict the first-in-children dose (Mahmood et al. 2014). Dosing in obesity is also complicated not only due to the physiological alternation and comorbidity but also the lack of standardized measure of creatinine clearance and the

variability of types of body weight used for dose adjustment (Meng et al. 2017). Renal impairment has a big impact on the PK of drugs which are extensively eliminated by kidney (≥30%). Particularly, for patients with end-stage renal disease (ESRD, renal function is less than 10% of the normal capacity), whether dose needs to be adjusted and how the dose should be adjusted becomes a critical issue as ESRD patients also routinely receive hemodialysis to assist drug removal. Pharmacometrics is critically valuable for dose optimization in special population. The dosing finding can be based on both approaches as we discussed previously: (1) PTA based on PK modeling and simulation and PK/PD indices and (2) simulation based on semi-mechanistic PK/PD model.

An example is illustrated here on the application of PK/PD approach for gentamicin dosing strategy in ESRD patients receiving hemodialysis (Zhuang et al. 2016). Similar approach can also be utilized for dose adjustment in other special populations. A one-compartment model was able to describe gentamicin PK adequately. An additional clearance during hemodialysis was incorporated into the model as $CL_{TOT} = CL_{NR} + CL_R + HEMO \cdot CL_{HD}$ such that the total clearance is the sum of nonrenal clearance, renal clearance, and hemodialysis clearance when dialysis is on (HEMO is an indicator with value 1 when hemodialysis is on and 0 when hemodialysis is off). PD model was developed based on data from *in vitro* static and dynamic time-kill studies against three bacterial strains. An adaptive factor was incorporated onto EC_{50} as previously described.

Based on this semi-mechanistic PK/PD model, Monte Carlo simulation was performed with two dosing regimens: (i) 120 mg after hemodialysis, which is the recommendation from gentamicin label, and (ii) 240 mg 1 h before hemodialysis, which has been used in several literatures. The concentration-time profiles for 1000 subjects are showing in Fig. 7 (upper) with the lines indicating mean values and shaded bands implying the 50% (dark) and 95% CI (light). The bacterial killing over time in ESRD patients undergoing both dosing regimens was also predicted (Fig. 7 lower). Simulation results suggested that PTA of fC_{max}/MIC > 8 are 10% and 100% for the first and second dosing regimen, respectively. Since gentamicin toxicity is associated with the trough concentration, therefore PTA of fC_{trough}/MIC < 2 was also calculated for both dose regimens, yielding 80% and 25%, respectively. PTA results implied the first dosing regimen has better safety but lower efficacy than the second. Predictions of bacterial response from the semi-mechanistic PK/PD model suggested that the second dosing regimen is only slightly better than the first one with both dosing regimens displaying bacterial density reduction of $>1 \times 10^2$. Therefore, the author concludes that the first dosing regimen provides a well-balanced benefit/risk profile than the second regimen in ESRD patients.

Modifications of the Method

Pharmacometrics has become a very important component in drug development to maximize the clinical potential of drugs. During the recent 30 years, PK/PD approaches have been rapidly evolving and widely spread in academia, pharmaceutical industry, and regulatory agency. The principles and applications presented here illustrate the opportunities in model-informed drug development. Books from Ette and Williams (2007) and Schmidt and Derendorf (2014) provide excellent overview on the basic principles of pharmacometrics and numerous examples of its applications to expedite successful drug development.

FDA and EMA issued guidelines on population PK (FDA 1999; EMA 2007) and other related topics such as pediatrics (EMA 2006), exposure-response relationships (FDA 2003), or QTc interval prolongation (FDA 2005). Both agencies have emphasized the power of PK/PD approach in effective data leveraging and risk/benefit balancing for rational dose selection and response prediction. Currently, FDA is conducting model-informed drug development pilot program to facilitate decision-making and improve trial success probability for drug development.

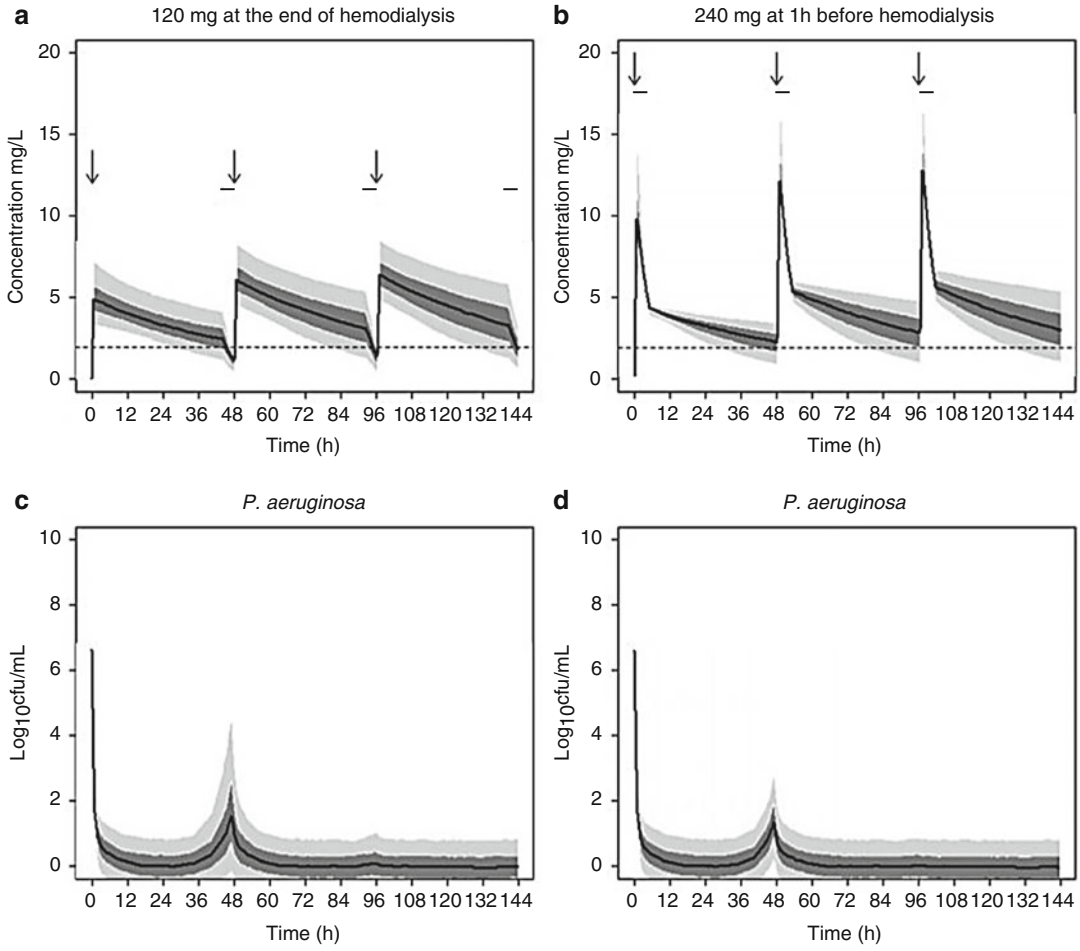


Fig. 7 Prediction of antibacterial activity of gentamicin against *P. aeruginosa* under two dosing regimens. (**a**) Concentration-time profile with 120 mg at the end of hemodialysis; (**b**) concentration-time profile with 240 mg 1 h before hemodialysis; (**c**) bacterial response-time profile with 120 mg at the end of hemodialysis; (**d**) bacterial response-time profile with 240 mg 1 h before hemodialysis. Dashed lines represent the safety threshold (2 mg/L). Arrows represent dosing time. Tick error bars represent the hemodialysis time interval. (Image adapted from Zhuang et al. (2016) and used with permission)

References and Further Reading

Asin-Prieto E, Soraluce A, Troconiz IF, Campo Cimarras E, Saenz de Ugarte Sobron J, Rodriguez-Gascon A, Isla A (2015) Population pharmacokinetic models for cefuroxime and metronidazole used in combination as prophylactic agents in colorectal surgery: model-based evaluation of standard dosing regimens. Int J Antimicrob Agents 45(5):504–511

Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S (2004) Bacterial persistence as a phenotypic switch. Science 305(5690):1622–1625

Benson JM (2017) Antimicrobial pharmacokinetics and pharmacodynamics in older adults. Infect Dis Clin N Am 31(4):609–617

Chaurasia CS, Muller M, Bashaw ED, Benfeldt E, Bolinder J, Bullock R, Bungay PM, DeLange EC, Derendorf H, Elmquist WF, Hammarlund-Udenaes M, Joukhadar C, Kellogg DL Jr, Lunte CE, Nordstrom CH, Rollema H, Sawchuk RJ, Cheung BW, Shah VP, Stahle L, Ungerstedt U, Welty DF, Yeo H (2007) AAPS-FDA workshop white paper: microdialysis principles, application, and regulatory perspectives. J Clin Pharmacol 47(5):589–603

Craig WA, Redington J, Ebert SC (1991) Pharmacodynamics of amikacin in vitro and in mouse thigh and lung

infections. J Antimicrob Chemother 27(Suppl C):29–40
Csajka C, Verotta D (2006) Pharmacokinetic-pharmacodynamic modelling: history and perspectives. J Pharmacokinet Pharmacodyn 33(3):227–279
Czock D, Keller F (2007) Mechanism-based pharmacokinetic-pharmacodynamic modeling of antimicrobial drug effects. J Pharmacokinet Pharmacodyn 34(6):727–751
Derendorf H, Meibohm B (1999) Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives. Pharm Res 16(2):176–185
Derendorf H, Lesko LJ, Chaikin P, Colburn WA, Lee P, Miller R, Powell R, Rhodes G, Stanski D, Venitz J (2000) Pharmacokinetic/pharmacodynamic modeling in drug research and development. J Clin Pharmacol 40(12 Pt 2):1399–1418
Drusano GL (2004) Antimicrobial pharmacodynamics: critical interactions of 'bug and drug'. Nat Rev Microbiol 2(4):289–300
EMA (2006) Guidance on the role of pharmacokinetics in the development of medicinal products in the pediatric population. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003066.pdf
EMA (2007) Guidance on reporting the results of population pharmacokinetic analyses. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003067.pdf
Ette EI, Williams PJ (2007) Pharmacometrics: the science of quantitative pharmacology. Wiley, Hoboken
FDA (1999) Guidance for industry population pharmacokinetics. https://www.fda.gov/downloads/drugs/guidances/UCM072137.pdf
FDA (2003) Guidance for industry exposure-response relationships – study design, data analysis, and regulatory applications. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm072109.pdf
FDA (2005) Guidance for industry E14 clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for nonantiarrhythmic drugs. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073153.pdf
Felmlee MA, Morris ME, Mager DE (2012) Mechanism-based pharmacodynamic modeling. Methods Mol Biol 929:583–600
Gabrielsson J, Weiner D (2016) Pharmacokinetic and pharmacodynamic data analysis: concepts and applications, 5th edn., revised and expanded. Apotekarsocieteten, Stockholm
Gieschke R, Burger HU, Reigner B, Blesch KS, Steimer JL (2003) Population pharmacokinetics and concentration-effect relationships of capecitabine metabolites in colorectal cancer patients. Br J Clin Pharmacol 55 (3):252–263
Gong X, Hu M, Zhao L (2018) Big data toolsets to pharmacometrics: application of machine learning for time-to-event analysis. Clin Transl Sci 11:305–311
Goutelle S, Maurin M, Rougier F, Barbaut X, Bourguignon L, Ducher M, Maire P (2008) The Hill equation: a review of its capabilities in pharmacological modelling. Fundam Clin Pharmacol 22(6):633–648
Greco WR, Bravo G, Parsons JC (1995) The search for synergy: a critical review from a response surface perspective. Pharmacol Rev 47(2):331–385
Holford N (2013) A time to event tutorial for pharmacometricians. CPT Pharmacometrics Syst Pharmacol 2: e43
Jusko WJ (1971) Pharmacodynamics of chemotherapeutic effects: dose-time-response relationships for phase-nonspecific agents. J Pharm Sci 60(6):892–895
Levison ME (2004) Pharmacodynamics of antimicrobial drugs. Infect Dis Clin N Am 18(3):451–465. vii
Liu P, Rand KH, Obermann B, Derendorf H (2005) Pharmacokinetic-pharmacodynamic modelling of antibacterial activity of cefpodoxime and cefixime in in vitro kinetic models. Int J Antimicrob Agents 25(2):120–129
Mager DE, Wyska E, Jusko WJ (2003) Diversity of mechanism-based pharmacodynamic models. Drug Metab Dispos 31(5):510–518
Mahmood I (2007) Prediction of drug clearance in children: impact of allometric exponents, body weight, and age. Ther Drug Monit 29(3):271–278
Mahmood I, Staschen CM, Goteti K (2014) Prediction of drug clearance in children: an evaluation of the predictive performance of several models. AAPS J 16(6):1334–1343
Meng L, Mui E, Holubar MK, Deresinski SC (2017) Comprehensive guidance for antibiotic dosing in obese adults. Pharmacotherapy 37(11):1415–1431
Mould DR, Upton RN (2012) Basic concepts in population modeling, simulation, and model-based drug development. CPT Pharmacometrics Syst Pharmacol 1:e6
Mouton JW, Vinks AA (2005) Pharmacokinetic/pharmacodynamic modelling of antibacterials in vitro and in vivo using bacterial growth and kill kinetics: the minimum inhibitory concentration versus stationary concentration. Clin Pharmacokinet 44(2):201–210
Mouton JW, Dudley MN, Cars O, Derendorf H, Drusano GL (2005) Standardization of pharmacokinetic/pharmacodynamic (PK/PD) terminology for anti-infective drugs: an update. J Antimicrob Chemother 55(5):601–607
Nielsen EI, Viberg A, Lowdin E, Cars O, Karlsson MO, Sandstrom M (2007) Semimechanistic pharmacokinetic/pharmacodynamic model for assessment of activity of antibacterial agents from time-kill curve experiments. Antimicrob Agents Chemother 51(1):128–136
Nielsen EI, Cars O, Friberg LE (2011a) Pharmacokinetic/pharmacodynamic (PK/PD) indices of antibiotics predicted by a semimechanistic PKPD model: a step toward model-based dose optimization. Antimicrob Agents Chemother 55(10):4619–4630
Nielsen EI, Cars O, Friberg LE (2011b) Predicting in vitro antibacterial efficacy across experimental designs with

a semimechanistic pharmacokinetic-pharmacodynamic model. Antimicrob Agents Chemother 55(4):1571–1579

Nolting A, Dalla Costa T, Rand KH, Derendorf H (1996) Pharmacokinetic-pharmacodynamic modeling of the antibiotic effect of piperacillin in vitro. Pharm Res 13(1):91–96

Roberts JA, Kirkpatrick CM, Lipman J (2011) Monte Carlo simulations: maximizing antibiotic pharmacokinetic data to optimize clinical practice for critically ill patients. J Antimicrob Chemother 66(2):227–231

Schmidt S, Derendorf H (2014) Applied pharmacometrics. AAPS advances in the pharmaceutical sciences series: 14. Springer, New York

Schuck EL, Grant M, Derendorf H (2005) Effect of simulated microgravity on the disposition and tissue penetration of ciprofloxacin in healthy volunteers. J Clin Pharmacol 45(7):822–831

Sharma A, Jusko WJ (1996) Characterization of four basic models of indirect pharmacodynamic responses. J Pharmacokinet Biopharm 24(6):611–635

Sheiner LB (1997) Learning versus confirming in clinical drug development. Clin Pharmacol Ther 61(3):275–291

Sheiner LB, Beal SL (1980) Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm 8(6):553–571

Sheiner LB, Rosenberg B, Melmon KL (1972) Modelling of individual pharmacokinetics for computer-aided drug dosage. Comput Biomed Res 5(5):411–459

Singh RSP, Mukker JK, Drescher SK, Deitchman AN, Derendorf H (2017) A need to revisit clinical breakpoints of tigecycline: effect of atypical non-linear plasma protein binding. Int J Antimicrob Agents 49(4):449–455

Sy SKB, Derendorf H (2016) Pharmacokinetics I: PK-PD approach, the case of antibiotic drug development. In: Müller M (ed) Clinical pharmacology: current topics and case studies, 2nd edn. Springer, Cham, pp 185–217

Sy SK, Zhuang L, Derendorf H (2016) Pharmacokinetics and pharmacodynamics in antibiotic dose optimization. Expert Opin Drug Metab Toxicol 12(1):93–114

Sy S, Zhuang L, Xia H, Beaudoin ME, Schuck VJ, Derendorf H (2017) Prediction of in vivo and in vitro infection model results using a semimechanistic model of avibactam and aztreonam combination against multidrug resistant organisms. CPT Pharmacometrics Syst Pharmacol 6(3):197–207

Sy SKB, Zhuang L, Xia H, Beaudoin ME, Schuck VJ, Nichols WW, Derendorf H (2018) A mathematical model-based analysis of the time-kill kinetics of ceftazidime/avibactam against *Pseudomonas aeruginosa*. J Antimicrob Chemother 73(5):1295–1304

Tam VH, Schilling AN, Nikolaou M (2005) Modelling time-kill studies to discern the pharmacodynamics of meropenem. J Antimicrob Chemother 55(5):699–706

Tam VH, Ledesma KR, Vo G, Kabbara S, Lim TP, Nikolaou M (2008) Pharmacodynamic modeling of aminoglycosides against *Pseudomonas aeruginosa* and *Acinetobacter baumannii*: identifying dosing regimens to suppress resistance development. Antimicrob Agents Chemother 52(11):3987–3993

Treyaprasert W, Schmidt S, Rand KH, Suvanakoot U, Derendorf H (2007) Pharmacokinetic/pharmacodynamic modeling of in vitro activity of azithromycin against four different bacterial strains. Int J Antimicrob Agents 29(3):263–270

Zhuang L, Sy SK, Xia H, Singh RP, Mulder MB, Liu C, Derendorf H (2015) Evaluation of in vitro synergy between vertilmicin and ceftazidime against *Pseudomonas aeruginosa* using a semi-mechanistic pharmacokinetic/pharmacodynamic model. Int J Antimicrob Agents 45(2):151–160

Zhuang L, He Y, Xia H, Liu Y, Sy SK, Derendorf H (2016) Gentamicin dosing strategy in patients with end-stage renal disease receiving haemodialysis: evaluation using a semi-mechanistic pharmacokinetic/pharmacodynamic model. J Antimicrob Chemother 71(4):1012–1021

Zuluaga AF, Salazar BE, Rodriguez CA, Zapata AX, Agudelo M, Vesga O (2006) Neutropenia induced in outbred mice by a simplified low-dose cyclophosphamide regimen: characterization and applicability to diverse experimental models of infectious diseases. BMC Infect Dis 6:55

General Principles of Pharmacovigilance in Clinical Development

58

Rainer Heissing and Anne-Ruth van Troostenburg

Contents

R. Heissing (✉)
Heissing Pharmacovigilance Consulting GmbH,
Partenheim, Germany
e-mail: rheissing@heissingpvconsulting.de

A.-R. van Troostenburg
Gilead Sciences International Ltd, Cambridge, UK

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_63

Abstract

Pharmacovigilance is a broad field that spans across all stages of the life cycle from preclinical drug development, clinical development, marketing approval, and post-marketing use. This chapter will focus on the pharmacovigilance aspects of interventional clinical trials. It provides a brief overview over the key elements of protecting patients in clinical trials as well as collecting and reporting safety information for the purposes of developing the safety profile of an investigational medicinal product. Regulations and requirements across the globe are complex and national, while certain international standards through the ICH guidelines form a common basic platform through which multinational clinical trials can harmonize.

Introduction and Scope

Pharmacovigilance / Drug safety is a broad field that spans across all stages of the life cycle from preclinical drug development, clinical development, marketing approval, and post-marketing use. This chapter will focus on the drug safety aspects of interventional clinical trials; other studies such as observational or epidemiological studies and non-study pharmacovigilance of drug use after marketing approval are beyond the scope.

The following pages will initially look at the regulatory requirements – taking the ICH guidelines (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, ICH) first and then adding a brief view on the legal framework and guidance in three major regions in the world: the US Food and Drug Administration (FDA), the European Union's European Medicines Agency (EMA), and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA).

Thereafter key definitions and terminology are introduced to provide the basis for understanding of the following discussions around planning a clinical trial, the collection and reporting of adverse events and the tools and requirements for monitoring of patient safety in clinical trials, and, lastly, actions to take to actively protect patients.

The rules and regulations that govern drug safety in clinical trials are complex and, as they are within the jurisdiction of each individual country's health authority, differ across the world. Therefore, this chapter will focus on the underlying fundamental principles and rely on the definitions and basic rules agreed in the ICH guidelines, which form the basis from which individual country regulations have further evolved. Wherever and whenever an actual clinical trial is being planned, it is critical to ensure that the specific rules of the country or countries where this clinical trial will be conducted are followed.

When designing a clinical trial, a clear research question needs to translate into defined objectives for the study and specific data collected to meet the primary (and secondary) objectives of the trial. In the typical efficacy clinical trial, these will be a small and very well-defined set of data points. However, the purpose of drug safety in clinical trials is broader and beyond the narrowly defined efficacy data set, and the vast majority of data collected is in support of the drug safety requirements.

Drug safety in clinical trials has a twofold purpose and in the order listed firstly to protect the subject/patient who is participating in the clinical trial and secondly to understand the general drug safety and tolerability of the drug being studied for the protection of patients who would be exposed to the drug after its approval in general

use and under less well-controlled conditions than those within a clinical trial. The purpose and priorities have their origin in the Declaration of Helsinki and form the fundamental basis for clinical trials and the ethics governing scientific research.

The Interests of the Individual Patient in the Study Have Precedence over the Interests of Society at Large

This principle is taken forward in the practical guidelines on GCP (ICH E6) (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2016).

Regulatory Requirements

ICH

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guideline E6 "Good Clinical Practice" (GCP) (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2016) was finalized in 1996. It is the international basis describing the responsibilities and expectations of all stakeholders in the conduct of clinical trials in Europe, Japan, the USA, and beyond.

The clinical safety-related guidelines (as opposed to the ICH safety guidelines, which concern preclinical requirements) are presented in ICH E2A-F and the Medical Dictionary for Regulatory Affairs (MedDRA) under the ICH multidisciplinary guidelines. Those specific to safety in clinical trials are:

- E2A – Clinical Safety Data Management: Definitions and Standards for Expedited Reporting (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 1994)
- E2B(R3) (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2013a) and E2B(R3) IWG (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2013b) – Clinical Safety Data Management: Data Elements for Transmission of Individual Case Safety Reports (ICSR) and Implementation: Electronic Transmission of ICSRs
- E2F – Developmental Safety Update Report (DSUR) (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2010)

USA

The Food and Drug Administration (FDA) is the supervising authority for investigational new drugs and interventional clinical trials in the USA.

Interventional clinical trials performed in the USA are regulated by the code of federal regulations CFR (Code of Federal Regulations (CFR) – Title 21 food and drugs – Chapter I food and drug administration (FDA), Department of health and human services – Subchapter D – Drugs for human use – Part 312 – Investigational new drug application (n.d.)). In CFR Title 21, Chapter I, subchapter D, part 312 "Investigational New Drug Application," the general rules and detailed requirements are described, from setup and application until the closure of a clinical trial.

Definitions and specifics about the collection, ongoing analysis, and reporting of safety information are outlined in §312.32 "IND safety reporting."

Additional information can be found in the "Final rule" guidance document on 21 CFR Parts 312 and 320 "Investigational New Drug Safety Reporting Requirements for Human Drug and Biological Products and Safety Reporting Requirements for Bioavailability and Bioequivalence Studies in Humans."

EU

The regulations that currently govern interventional clinical trials in the EU are complex.

EU Directive 2001/20/EC is the underlying basis, but legally binding is its adoption into local legislation in the individual EU member states (e.g., German "Arzneimittelgesetz"), which in parts vary from country to country.

Details on the collection and reporting of safety information can be found in EU "CT-3" guidance document "Detailed guidance on the collection, verification and presentation of adverse event/reaction reports arising from clinical trials on medicinal products for human use" (2011/C 172/01) (European Commission 2011).

EU Clinical Trial Regulation 536/2014 was adopted and entered into force in 2014 and will replace in the second half of 2019 the existing EU Clinical Trial Directive 2001/20/EC and all national legislation that was put in place to implement this Directive. But there will be an interim period of 3 years, where clinical trials can still be run according to the national legislation.

Japan

The Pharmaceuticals and Medical Devices Agency (PMDA) is the supervising authority for the conduct of interventional clinical trials in Japan.

Clinical trials in Japan are carried out in accordance with the Japanese Ordinance on Standards for Conduct of Clinical Trials (GCP) (Enforcement Regulations of the Law on Securing Quality, Efficacy and Safety of Pharmaceuticals, Medical Devices, Regenerative and Cellular Therapy Products, Gene Therapy Products, and Cosmetics 1961; Pharmaceutical Administration and Regulations in Japan 2017; J-GCP Ordinance of the Ministry of Health and Welfare 1997).

Key Definitions (ICH E2A)

Adverse Event or Adverse Experience (AE)

Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment.

Adverse Drug Reaction (ADR)

All noxious and unintended responses to a medicinal product related to any dose should be considered adverse drug reactions. The phrase "responses to a medicinal product" means that a causal relationship between a medicinal product and an adverse event is at least a reasonable possibility, i.e., there are facts (evidence) or arguments to suggest a causal relationship.

Unexpected Adverse Drug Reaction

An unexpected adverse reaction is one, the nature or severity of which is not consistent with the applicable investigator's brochure.

Serious Adverse Event (SAE)

Any untoward medical occurrence that results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect.

Serious Unexpected Suspected Adverse Reactions (SUSARs)

Any adverse event that is serious, is unexpected, and is suspected to be causally related to the investigational drug.

Planning of a Clinical Trial

The Investigator's Brochure (IB) is the Reference

This is the key document for any clinical trial; it forms the core of the submission package when applying for approval to conduct the clinical trial

and as a living document will evolve during the conduct of a clinical trial program, as more information becomes available over time.

Each clinical trial is supported by the IB, containing all information on the study drug available at the time and providing a broad overview of the known safety and tolerability of the investigational medicinal product (IMP). The majority of the IB will be taken up by detailed descriptions of studies and research results available and displaying both the growing knowledge of the targeted efficacy and the gained safety information to inform the investigators, ethics committees (ECs)/institutional review boards (IRBs), and authorities approving the trial. The safety information will be further summarized in a concise section of the IB of a highly technical nature, known as the Reference Safety Information (RSI).

The RSI presents a list of any events identified in previous clinical trials as expected serious adverse drug reactions (SADR). Of note, this means that events which are **not** listed in the RSI are "unexpected" and therefore potentially reportable: the RSI guides the selection of individual safety reports that have to be reported to authorities, other investigators, and ethics committees during the conduct of the trial – because they are new information that has not previously been observed. The requirements for selecting appropriate events for the RSI is guided by detailed rules, particularly clearly laid out within the European regulations.

The Informed Consent Form to Include All Relevant Safety Information

While the IB is a technical, scientific document that is used for reviewing authorities, ethics committees, and investigators, the subject/patient in the study also needs to receive transparent and comprehensive information about the clinical trial, the demands of the protocol, and what is known about the study drug. The relevant safety information available – and what is not yet known about the drug – needs to be summarized in a concise, readable fashion that allows a layperson to understand the possible benefits but importantly the possible risks of participation in the clinical trial.

Routine Safety Data Collection: Adverse Events/Serious Adverse Events

As part of the planning/designing of a clinical trial, the protocol will have to define very clearly what data will be collected in the context of the trial. For interventional clinical trials, the standard set of data to be collected in (almost) all instances are adverse events and serious adverse events, regardless of causality – furthermore, the investigator and study sponsor are required to make an assessment as to whether or not an event observed in a study subject is possibly related to the study drug or not. Where the investigator – or the receiving sponsor of the clinical trial – considers an event possibly related to the study drug, the adverse event is considered a possible adverse drug reaction (ADR).

The study protocol defines not only what safety data will be collected, but it will also define timelines by which the investigator has to communicate with the sponsor and send certain reports in a more expedited manner than simply collecting the information in the case report form (CRF) of the trial subject. Typically events that have to be sent faster to the sponsor of the clinical trial are those that may require onward reporting in a quick turnaround to authorities, ethics committees, and other investigators – these will be the SAEs but may also include certain events of special interest and some so-called special situation reports (SSRs) like pregnancies, or medication errors (see below).

Outcome Events/Unblinding of Data

The adverse event information collected in certain clinical trials may serve a purpose beyond the general understanding of the safety of study subjects, but specific events may also constitute part of the data collected for answering the efficacy objective of the clinical trial. Such trials are named clinical outcome trials and study the impact of an intervention (in most cases the administration of a drug) on

defined clinical outcomes. For example, cardiovascular outcome trials investigate strokes, heart attacks, and death – all of such events would clearly be collected as adverse event safety data but at the same time are the key "endpoint events" for answering the efficacy question whether a certain medical intervention may prevent or reduce such cardiovascular outcomes.

In such clinical outcome trials, the planning of collection, data handling, assessment, and adjudication of defined "outcome events" is a key point and requires a complex system setup to standardize events as much as possible. Beyond the investigator reporting an event, such events routinely go to a separate adjudication committee, who reviews the available information and applies preset definitions and algorithms to maximize homogeneity of diagnosis.

In case of clinical outcome trials, the outcome events often constitute serious unexpected events In case of a suspected causal association these events would meet the definition of a SUSAR and require unblinding and reporting to Authorities and Ethics Committees. It is therefore important in a clinical outcome trial protocol, that the protocol clearly identifies the specific outcome events. The study sponsor will have to obtain agreement with the supervising authorities and ethics committees for special conditions for the reporting or unblinding of outcome events. Global guidelines allow for clearly defined, case-by-case agreements on reporting conditions for outcome events to avoid routine and systematic unblinding of such events, which may threaten the data integrity of the clinical trial.

Definition of Expected Events/ Adverse Events of Special Interest

A primary part of the planning for a clinical trial and indeed a whole clinical development program is to determine (and update regularly) the RSI in the IB but also consider whether there are any adverse events of special interest (AESI). These AESIs may be nonserious in nature, but perhaps the preclinical program suggested that there may be a possible issue with a certain type of event (e.g., skin reactions or diarrhea) or they are part of the symptoms of the treated disease but may be of particular importance to determine whether the study drug is a contributing factor or modifies the event in any way and requires an intensive additional diagnostic pathway to ensure full understanding (e.g., hepatic or neurologic events of a certain nature). Adverse events of special interests are then identified in the IB and the specific study protocol and require the investigator to perform more detailed data collection, possibly perform a set of predefined mandatory additional diagnostic tests and also expedited communication to the study sponsor – often aligned with the timelines for reporting SAEs. Therefore AESIs will undergo special and fast scrutiny by both investigators and sponsors, and the information on such events will be maximized – this may over the course of a development program disprove a preclinical concern or elucidate specific treatment pathways that may be helpful for later use of the drug in the post-approval era (such as effective treatment of drug-induced diarrhea).

Standardized Data Collection for Later Pooling

When planning a drug development program, the overall aim is to develop clinical data to the purpose of marketing authorization. From the very outset, the different clinical trials to be conducted need to be designed in a fashion that will allow them to be used not just in isolation, but the data should be standardized so that pooling of data from multiple studies becomes possible. In order to support analysis and understanding of emerging safety profiles, the data from different studies should be combinable as well as separating different subgroups and populations from across studies.

Data and Safety Monitoring Boards (DSMB)

When planning a clinical trial or a whole program, the setting up of a DSMB is a complex and important task to ensure effective oversight over the clinical study or program. It requires the identification of the appropriate membership and

development of communication routes and mechanisms to supply data, charters, and rules of communication among the members as well as with the sponsor, meeting frequency, review principles, and standards of assessment. The DSMB is set up independently from the conduct or sponsor of a clinical trial with the aim to perform ongoing unblinded data monitoring while the study is ongoing. DSMBs are often set up to support a number of studies in the same program and provide ongoing guidance to the study sponsor as to whether the trial may continue as it is, requires modification, or should be terminated. The sponsor remains blind and will not receive the detailed reviews of the DSMB, but only the final recommendation. The full DSMB materials, minutes, analyses and documentation will be added to the Trial Master File after end of study.

Endpoint Adjudication Committee

Where an endpoint of a clinical trial is not an objective measure (such as a blood pressure or a particular laboratory value) it is important to ensure standardization of the measure and remove subjectivity and variance as much as possible. Particularly for clinical outcome trials, where the endpoints may be composites of multiple adverse events – the standardization of diagnosis of each of the contributing events is critical and should be agreed upon beyond the individual investigator. An endpoint adjudication committee is a way to ensure that the defined endpoint events are diagnosed to a common standard and context independent, based on preset data elements, clinical tests, and diagnostics that allow a central diagnosis to be made.

Collection and Reporting of Adverse Events/Serious Adverse Events

Collection of Adverse Events

Adverse Events and Serious Adverse Events are actively collected by the investigator and sponsor starting from the moment a participant signs the informed consent and until leaving the clinical trial (protocol defined end of data collection).

All AEs experienced by a patient at a certain point in time form an adverse event case report, an ICSR (individual case safety report). The ICSR forms the smallest unit of reporting – it may contain more than one event in the patient.

The investigator uses an AE/SAE form, paper or electronic, to document the AE(s) the study participant has experienced. The AE/SAE form can be integrated into the electronic Case Report Form (eCRF) of the trial or be a separate, loose form.

All adverse events have to be evaluated by the investigator and sponsor concerning seriousness, causal relationship, and expectedness. The assessments given by the investigator should not be downgraded by the sponsor.

Usually all nonserious AEs are entered into the clinical trial database and all serious AEs into the clinical trial and safety database. Consistency of the two databases has to be ensured by either the way of collecting AEs (simultaneously and electronically into both databases) or later reconciliation between the databases.

Case reports containing a serious unexpected suspected adverse reaction (SUSAR) event are usually unblinded, unless they are defined as exempted outcome events by the clinical trial protocol. The investigational drug given (verum, comparator, or placebo) is then documented in the case report in the safety database.

Adverse events are coded using the Medical Dictionary for Regulatory Affairs (MedDRA).

Special situation reports without an associated adverse event (e.g., pregnancy, overdose, medication error, etc.) might not qualify for individual case reporting but nevertheless need to be recorded by the sponsor for continuous and cumulative safety analysis and presentation and discussion in periodic aggregate safety reports such as the Development Safety Update Report.

Reporting of Serious Adverse Events

The obligations on the collection and reporting of safety information of a clinical trial are directly

conferred with the approval to perform a clinical trial and lie with the study sponsor (the applicant for the clinical trial authorization).

The safety information collected in the clinical trial is reported by the investigators to the study sponsor – in the (e)CRF or on special forms and there is a subset of data that requires prompt (expedited) reporting to the sponsor – these are mainly serious events. Case reports containing a serious unexpected adverse reaction then have to be reported further by the sponsor.

When a case is initially reported to a sponsor, the report may be incomplete – for reasons where there is still a lack of knowledge, as the situation of the patient is still evolving or because a form has not been effectively completed. Only cases that contain a minimum set of data should be reported onward – these are considered "valid" cases for the purposes of reporting.

For the purpose of reporting cases of suspected adverse reactions, the minimum data elements for a valid case are:

- An identifiable reporter
- An identifiable patient
- An adverse reaction
- A suspect medicinal product (see Annex IV, ICH-E2D Guideline (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2003))

If any of these elements are missing, the case is considered "nonvalid" for expedited reporting purposes, and while at that time therefore not reportable, it is incumbent on the sponsor of the clinical trial to follow up the case intensively and urgently to ensure that at least the minimal criteria are all met. In the setting of a clinical trial, there should not really be any nonvalid cases, as the structured and tight control over the treatment and data collection on all patients should ensure that all four minimal criteria are always available even at the first instance.

Depending on national legislation, the individual case reports have to be as a minimum reported in an expedited manner to national competent authorities. There are strict timelines defined for the reporting, usually 7 calendar days after first awareness, if the case report is classified as fatal or life-threatening, and 15 days for all other serious case reports. Based on the applicable legislation, these case reports might have to be reported, if the adverse reaction not only is causally associated to verum, but to a comparator drug or even placebo (although this is rare, if there is a suspicion of the ADR being related to an excipient).

The identification of potentially reportable cases in a timely manner as they are received by the sponsor of the clinical trial is a key activity in the ongoing support of a clinical trial. Particularly where a clinical trial is conducted in a double-blind fashion, the regulatory requirements of the EU require unblinding of the treatment code for the patient, unless there is a prior documented agreement with the authority who approved the clinical trial that certain events do not require unblinding (such as in the case of endpoint events in a clinical outcome trial). In the context of international clinical trials conducted in several countries worldwide, the sponsor will usually apply a common standard to reporting to authorities – so that all authorities receive the same information and therefore may report cases in an unblinded manner to authorities who do not require it.

Figure 1 presents an algorithmic decision tree to determine whether a case requires unblinding and potentially reporting. This is taking general principles only and would need to be adjusted and supported by detailed reporting requirements for each country, ethics committee and site, as any of these may have additional requirements for other events to be reported than SUSARs.

The challenge with unblinding of SUSARs for the purposes of expedited reporting is to ensure that only a few select individuals in the sponsor responsible for reporting the event (typically a part of the drug safety personnel) receive the information on the treatment allocation of a patient and that this information is not shared with personnel otherwise involved in the conduct of the clinical trial or investigators.

In addition to reporting of SUSARs to supervising national competent authorities, individual countries may require reporting of individual case

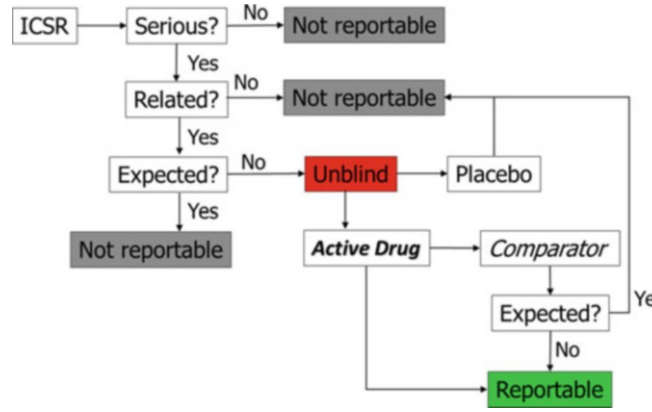


Fig. 1 Assessing reportability in blinded clinical trials (by ARvanTroostenburg)

reports to ethics committees and to participating investigators. Timelines and report format vary from 7/15 days individual case reports to, e.g., quarterly batch reports with cases included in form of a line listing. Also, while reports to ethics committees will be of the same standard as the reports going to the national competent authority – i.e., in an unblinded fashion, reporting to investigators is done blinded, so as not to affect the integrity of the blinded study design and bias the investigators. Only in the case of a major safety concern arising from received safety information is consideration given to unblinding investigators at the same time as all other personnel – this is in general only when major actions for the safety of the patients in the clinical trial may have to be considered (see further below for Urgent Safety Measures).

If required or possible, expedited reporting can be done electronically using the ICH defined E2B standard.

Reporting requirements for serious unexpected adverse reactions continue even after the end of the clinical trial. Therefore collection, processing (only in the safety database), and reporting of respective case reports received after the end of the trial continue.

Events relevant for patient safety may occur during a clinical trial which do not fall within the definition of a serious unexpected adverse reaction and thus are not subject to the reporting requirements described above, e.g. a major safety finding from a newly completed animal study, such as carcinogenicity. Sponsors are also obliged to inform the national competent authority, ethics committee, and investigators of findings that could adversely affect the safety of subjects, impact the conduct of the trial, and might materially alter the current benefit-risk assessment for an investigational drug.

Monitoring of Patient's Safety and Actions

Continuous Monitoring of Patient's Safety

The sponsor is responsible for the ongoing safety evaluation of the investigational drug. Ongoing safety evaluation consists of various layers of safety monitoring activities.

Evaluation and assessment of individual case reports are performed during case processing. An important aspect is the causality assessment. To decide about a potential causal association between drug treatment and adverse event a variety of aspects can be taken into account, e.g., timely relationship, pharmacological plausibility, de-challenge and rechallenge, concurrent diseases, and concomitant medication.

Cumulative interim safety analyses are performed at specified time points. It may include a comparison of adverse event rates for verum against comparator/placebo or a comparison of adverse event rates against predefined expected

rates based on epidemiological data, or identification of trigger events to detect cases of interest, e.g. Hy's Law cases and drug-induced liver injury.

In case of blinded clinical trials, the support of an external Data and Safety Monitoring Board for unblinded analysis is needed to keep staff of the sponsor involved in the conduct of the clinical trial blind.

Another tool for a cumulative safety evaluation is the Development Safety Update Report.

Development Safety Update Report

The ICH E2F Guideline (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2010) defines format and content of the Development Safety Update Report (DSUR). It provides periodic (typically annually) analysis of clinical trial safety for an investigational drug. The DSUR summarizes identified and potential risks, describes new safety issues that arose during the period of the report, determines if reporting period safety information is in accordance with prior product safety knowledge, and provides an update on the clinical development program. The main focus is on interventional clinical trials, ongoing or completed during the reporting interval. The investigator's brochure is the reference document for the DSUR.

Included in the DSUR is a line listing of serious adverse reactions arising in the reporting interval and a cumulative tabulation of all serious adverse events organized by MedDRA System Organ Class, from the start of development to date, for verum, comparator, placebo, and still blinded cases.

Actions and Measures

New relevant safety information, such as newly identified safety issues, changed product safety knowledge, or DSMB recommendations, may warrant actions to be taken in a clinical trial or even across a whole clinical program.

A basic measure is the update of the investigator's brochure. New product safety knowledge is included and newly identified side effects added to the Reference Safety Information (RSI). Updates of the IB are usually performed on a frequent basis (e.g., yearly) to include and summarize the increasing knowledge about the developmental drug.

Safety issues having a significant impact on the safety of the subjects may require a substantial protocol amendment, e.g. changing exclusion criteria or introducing a new monitoring procedure and need to be submitted to national competent authorities and ethics committees for approval.

Safety issues requiring urgent safety measures to protect subjects against any immediate hazard, such as temporarily halting of the clinical trial, may be taken immediately without prior authorization from the competent authority in form of an urgent amendment. The sponsor must inform the competent authority and the ethics committee concerned as soon as possible. Some countries have strict timelines in place for the reporting of Urgent Safety Measures, and there may be only a few hours from the information being available to the study sponsor to the need to inform competent authorities. Therefore a rapid assessment and decision-making system needs to be in place at the sponsor to be able to respond to major safety issues with all due haste for the protection of patients in the clinical trial – or also possibly a whole development program of many clinical trials.

References and Further Reading

Code of Federal Regulations (CFR) – Title 21 food and drugs – Chapter I food and drug administration (FDA), Department of health and human services – Subchapter D – Drugs for human use – Part 312 – Investigational new drug application

Directive 2001/20/EC of the European Parliament and of the council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use

Enforcement Regulations of the Law on Securing Quality, Efficacy and Safety of Pharmaceuticals, Medical Devices, Regenerative and Cellular Therapy Products, Gene Therapy Products, and Cosmetics (MHW ordinance no. 1, February 1, 1961). Final revision: MHLW ordinance no.82, April 10, 2015 and MHLW ordinance no.92, July 31, 2014 (to be enforced on June 12, 2017) article 273

European Commission – Communication from the Commission – Detailed guidance on the collection, verification and presentation of adverse event/reaction reports arising from clinical trials on medicinal products for human use ('CT-3') (2011/C 172/01) June 2011

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E2A – Clinical safety data management: definitions & standards for expedited reporting October 1994

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E2D – Post-approval safety data management: definitions and standards for expedited reporting November 2003

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E2F – Developmental safety update report (DSUR) August 2010

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E2B (R3) clinical safety data management: data elements for transmission of individual case safety reports July 2013a

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E2B (R3IWG implementation: electronic transmission of individual case safety reports July 2013b

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) – Guideline E6 'good clinical practice' (GCP) – R1 June 1996 and integrated addendum R2 November 2016

J-GCP Ordinance of the Ministry of Health and Welfare No. 28 of March 27, 1997 (As last amended by the ordinance of ministry of health, labour and welfare No. 161 of December 28, 2012) and guidance on the ministerial ordinance on the standards for the conduct of clinical trials of medicinal products (PFSB/ELD notification no. 1228/7 dated 28 December 2012)

Pharmaceutical Administration and Regulations in Japan (2017) (1.3 safety information on adverse reactions and infections during the study)

Regulation (EU) no 536/2014 of the European Parliament and of the council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing directive 2001/20/EC

Part III

Regulations

Regulatory Guidance: ICH, EMA, FDA

59

Gerd Bode

Contents

Abstract

Two revolutionary breakthroughs have favored the international development of pharmaceutical compounds: on one hand the International Conferences on Harmonization (ICH), on the

G. Bode (✉)
Institute for Pharmacology and Toxicology, University Medical Center, University of Goettingen, Goettingen, Germany
e-mail: gerd-bode@t-online.de

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_58

other hand the computer sciences. ICH provides tremendous support for the strategies in general by creating global guidelines, and these recommend the basis for research, for preclinical and clinical development. A great number of solutions for detailed pending issues in the areas of Quality, Safety, and Efficacy are offered. The enormous amount of data is then organized for the Common Technical Dossier by teams of experimental researchers, physicians, computer scientists, and statisticians for submission to Regulatory Agencies for the final evaluations necessary before the market authorization. Thus, the pharmaceutical field is passing through a major period of transformation. Many changes are driven by information technology but also by enormous progress in medical scientific research. Therefore, it has to be recognized that many breakthroughs, made in labs as well as on laptops, have changed and facilitated the pharmaceutical world.

This chapter highlights the progress in the International Conferences on Harmonization during the last decades; the focus is on preclinical Safety (Pharmacology and Regulatory Toxicology) with the objectives to identify risks for drugs in development by in silico, in vitro, and in vivo methods to prevent, treat, and cure diseases. When needed, then also regional guidances, e.g., from US Food and Drug Administration (FDA) or European Medicines Agency (EMA), are referred to. This global, complex procedure reveals: we are now living in the age of big data and success is only possible by an enthusiastic multidisciplinary communication and transparent cooperation.

General Considerations

The International Conferences for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use is now rebaptized into *the International* ***Council*** *for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use* (ICH). ICH continuous in bringing together the regulatory authorities and pharmaceutical industry to discuss scientific and technical aspects of drug development and registration. This was an important decision, because in the 1980s a great number of differing discrepancies came into existence for the drug development in different areas of the world with the result that many studies for quality, safety, and efficacy had to be repeated. Figure 1 illustrates one example from 1986, where more than 12 months studies were requested, even when clinical treatment was only intended for 1 week.

Since its inception in 1990, ICH has gradually evolved to respond to the increasingly global face of drug development. ICH's mission is to achieve greater comparable requirements worldwide to ensure that safe, effective, and high-quality medicines are developed and registered in the most resource-efficient manner.

The International Conferences was jointly supported and organized by the Commission of the European Communities (CEC), the US Food and Drug Administration (FDA), and the Japanese Ministry of Health and Welfare (MHW), together with the pharmaceutical industry, as represented by the International Federation of Pharmaceutical Manufacturers Associations (IFPMA), the European Federation of Pharmaceutical Industry Associations (EFPIA), the US Pharmaceutical Manufacturers Association (PMA), and the Japanese Pharmaceutical Manufacturers Association (PMA) (Table 1).

Many important initiatives have been undertaken between regulatory authorities and industry associations, particularly on a bilateral basis first, to promote harmonization of regulatory requirements between the three regions Japan, the USA, and the European Community. ICH owes much to these initiatives of experts during international symposia.

ICH has been very successful in this harmonization process; its expansion from a three-party movement to an internationally active process is supported worldwide. ICH marks today 25 years of successful transparent cooperation. Its reduction from huge international conferences to a more counseling organization reflects the fact that many of the major issues in quality, safety, and efficacy have been agreed upon and today focuses on special problems.

The organization for the process of creating guidances has been clarified; the ICH Steering

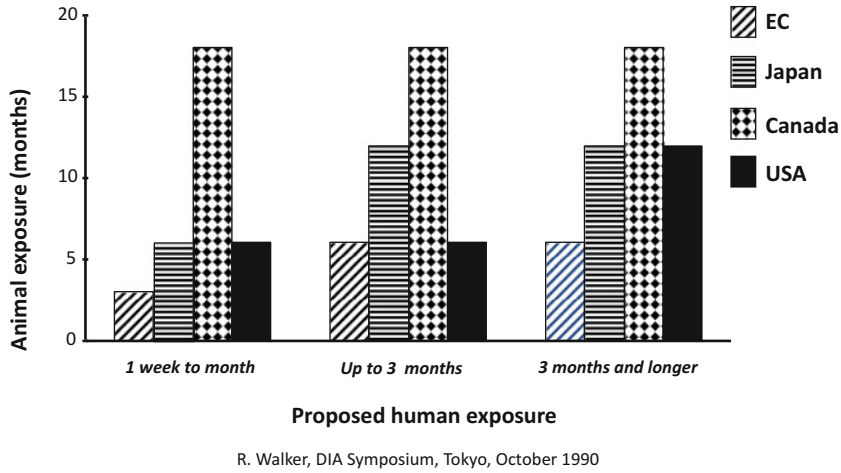


Fig. 1 Durations of animal studies for proposed clinical trials (1986)

Table 1 Organization of International Conferences on Harmonization (ICH)

Organization	
Trade associations	
EFPIA	European Federation of Pharmaceutical Industries and Associations
JPMA	Japanese Pharmaceutical Manufacturers Association
PhRMA	Pharmaceutical Research and Manufacturers Association (USA)
Regulatory authorities	
European Union	Commission and appropriate "parties"
	(CHMP, QWP, BQWP, EWP, SWP, PhVWP)
Japan	Ministry of health and welfare and appropriate expert support
	(MHLW, NIHS, universities)
USA	Food and drug administration and appropriate expert support
	(regulatory communication, efficacy, safety, quality – review staff, expert leads – CDER/CBER chief officers)
Observers	Canada, EFTA, WHO
Interested parties	World self-medication industry – WSMI, international generic pharmaceutical alliance – IGPA
"Umbrella" organization, ICH secretariat	
IFPMA	International Federation of Pharmaceutical Manufacturers Associations

Committee governs and controls the objectives and their progress on every step (Fig. 2). The groups of experts are competent and self-critical, knowing that often special expertise is needed. ICH then allows specialists to be invited as consultants to the discussions to lift conclusions up to the state-of-the art levels. In addition, the ICH Step 2 Consensus papers are published and submitted to comments for a period of 6 months. Hereby all levels of information and recommendation need to be respected for harmonization development to optimize the design of the individual tests and supported the endeavor to avoid redundant studies.

ICH guidelines represent always the highest level of scientific and regulatory recommendations. Regional guidelines like those from the EMA or FDA refine the ICH guidances by adding important details which also have to be taken into consideration. As an example, the ICH guidelines on carcinogenicity (S1 A-C) are supplemented by additional information via the EMA guideline on carcinogenicity on which organs need to be sampled and assessed by microscopic evaluation, or that this work should preferably be done by one board certified pathologist, or that historical data are only acceptable if these are not older than 5 years, due to a possible genetic drift in the number of generations during breeding. Thereafter, on the level of OECD guidelines, all details of the individual studies can be considered. Accordingly, complete regulatory guidance can only be gained when all specific guidelines are respected (Table 2).

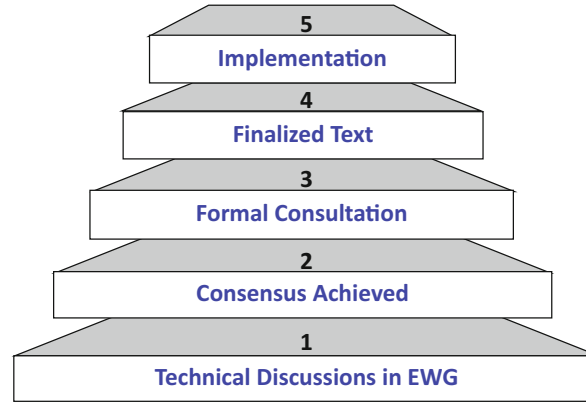


Fig. 2 Five steps of ICH before implementation

Table 2 Different levels of regulatory guidances

Levels of recommendations for regulatory toxicology
ICH
Highest level of information, accepted by regulatory agencies, focus on the need and timing of studies. Example: Carcinogenicity ICH S1 A-C
EMA or FDA
Additional details helpful for strategies and individual tests
Example: EMA (January 2003) – note for guidance on carcinogenic potential
FDA (May 2002): Carcinogenicity studies protocol submissions
OECD
Study details, accepted internationally as standard methods for safety testing, regularly updated with the assistance of thousands of national experts from OECD member countries. Example: Test no. 451 carcinogenicity studies

This chapter focuses on safety. Safety means in the ICH language: Nonclinical safety investigations like toxicology or safety pharmacology studies, which identify undesirable and often unexpected adverse effects induced by pharmaceutical or chemical compounds. In vitro or in vivo investigations describe the type and degree of toxicity and assess the risks first for nonclinical models. The results support the subsequent management of the risks in humans, when participating in clinical trials or later as patients after market authorization, and attempt to analyze the mechanisms behind the alterations; they extrapolate preclinical hazards to humans and finally help to communicate these risks to populations concerned with exposure of that particular substance.

Test strategies for toxicological investigations have been refined and the extrapolation of preclinical results to humans improved.

Soon after the finalization of important guidelines it became visible, that unnecessary studies were avoided, and as a result a number of animals could be saved for experiments as shown in 1997 (Table 3).

The assessments of all properties should be conducted using high-quality scientific standards with data collection records readily available and in compliance with good laboratory practices (GLP) regulations stressing the requirement to document all results and provide safe conditions for archiving.

The safety of drugs is of public interest; therefore, Ethical Committees take care of the safety of volunteers and patients as long as clinical investigations are running. Regulatory agencies have become partners for Industry and offer scientific advice, especially when the development reaches critical steps as, e.g., the decisions for administering the new drug for the first time to humans or later at the end of Phase II when an enormous extension in the number of patient enrolments takes place or finally for the MAA (Market Authorization Application).

With the creation of the International Conferences/Council of Harmonisation (ICH), many of international initiatives were channeled into a global process, which continuous with the running revisions and updates of all considerations. The World Health Organization (WHO) is highly active in distributing existing knowledge, also in countries

Table 3 Animal usage before and after the influence of ICH

Advantages			
Animal use for toxicological studies			
		Before ICH1	After ICH4
Single-dose toxicity (2 routes)	2 rodent	200–300	50–100
	1 non-rodent	16–32	0
Repeated-dose sub-chronic	1st duration (e.g., 1 month)		
	Rodent	80	160[a]
	Non-rodent	24	32
	2nd duration (e.g., 3 months)		
	Rodent	160	0
	Non-rodent	32	0
	Recovery		
	Rodent	200	0
	Non-rodent	40	40
Repeated-dose chronic	1st duration (e.g., 6 or 9 months)		
	Rodent	160	160
	Non-rodent	32	32
	2nd duration (e.g., 12 months)		
	Rodent	160	160
	Non-rodent	32	32
Reproduction	Segment I rat		
	Japanese style	192	192
	US/EU style	96	0
	Segment II rat		
	Japanese style	96 + 48	0
	US/EU style	96	96
	Segment II rabbit	60	60
	Segment III rat	96	96
Carcinogenicity	1st species (e.g., rat)	400–500	400–500
	2nd species (e.g., mouse)	500	500
	Medium- or short-term study	0	160
Total		2720–2936	1478–1583

Lumley and van Cauteren (1997)
[a]Excludes offspring and dose-range finding studies

of development. There are a number of regional societies supporting this process (Table 4). Global Cooperation Groups (GCG) address increasing interest by non-ICH parties in ICH guidelines and operations. These groups facilitate dissemination of information on ICH activities, guidelines, and their use. The underlying principle is that ICH will not seek to impose its views, but the GCG will serve as resource for information. Four brochures have been published on ICH and GCG, available at ICH website www.ich.org.

The ICH focuses on guidances for quality, preclinical safety (= toxicology), and efficacy (clinical

Table 4 Regional initiatives

Regional harmonization initiatives
APEC
Asia-Pacific Economic Cooperation
ASEAN
Association of the Southeast Asian Nations
GCC
Gulf Cooperation Council
PANDRH
Pan American Network for Drug Regulatory Harmonization
SADC
Southern African Development Community

effective substances) of pharmaceutical compounds. This section illustrates all ICH guidelines important for the safe use of drugs worldwide.

Under the heading of S are all guidelines collected which contribute to the preclinical **safety** of pharmaceutical compounds. These can be in vitro experiments or classical toxicity studies in animals identifying unexpected adverse reactions or safety pharmacology investigations focusing on damages of physiological functions.

Besides the general GLP requirements, all safety studies should always be supported by kinetic information; there is a need to know at which exposure levels which undesirable effects appear and if these reactions have any clinical importance or if thresholds for reactivity can be denominated for human conditions and therefore are acceptable if used with care.

Activities for ICH Conferences

A Steering Committee was appointed with members from EU, FDA, MHW, EFPIA, JPMA, PMA, and IFPMA, with observers from the WHO and from the regulatory authorities of Canada and Switzerland (for EFTA).

The Steering Committee sets up three joint industry/regulatory Expert Working Groups to deal with the technical aspects of the three main subject areas – quality, safety, and efficacy – which were discussed in three parallel workshops during the conferences. Each of the Expert Working Groups has members, representing EU, FDA, MHW, EFPIA, PMA, and JPMA. With advice from these technical Working Groups, the Steering Committee was responsible for the selection and prioritization of the topics discussed at the Workshops at the ICH conferences (Table 1).

The International Conference on Harmonization differs from many other harmonization initiatives in that it has a recognized status and is backed by a commitment on the part of both industry and regulators, to facilitate greater harmonization of technical requirements in the three regions and today worldwide. Commitment to these objectives, set out in the terms of reference, was reaffirmed by the Steering Committee in a statement issued first following the meeting held in Tokyo, October 1990.

Objectives of the ICH

The objectives of ICH have been set out from the beginning and continue to be respected in all subsequent activities:

1. To provide a forum for a constructive dialogue between regulatory authorities and the pharmaceutical industry on the real and perceived differences in the technical requirements for product registration in the CEC, USA, and Japan
2. To identify areas where modifications in technical requirements or greater mutual acceptance of research and development procedures could lead to a more economical use of human, animal, and material resources, without compromising safety
3. To make recommendations on practical ways to achieve greater harmonization in the interpretation and application of technical guidelines and requirements for registration

There is considerable success of the international discussions and constructive solutions documented in the guidelines.

Overview of Guidelines for Quality

Quality: There are more than 20 guidelines focusing on issues concerning the quality of compounds. These deal with issues like stability, analytical validation, detailed solutions how to handle impurities, further aspects of pharmacopeias, quality of biotechnology-derived, specifications, manufacturing practices, pharmaceutical development, quality risk assessments, pharmaceutical quality systems, development and manufacture of drug substances, and life cycle

management. In addition there are a number of issues with translateral importance, many of these can be found under the headings of multidisciplinary guidelines and sections.

Overview of Guidelines for Efficacy

Efficacy: There are more than 30 guidelines dealing with problems of efficacy during clinical trials or after marketing authorization. Among them are as follows:

Population Exposure: The Extent of Population Exposure to Assess Clinical Safety
Good Clinical Safety Data Management: Definitions and Standards for Expedited Reporting
Clinical Safety Data Management: Data Elements for Transmission of Individual Case Safety Reports
Clinical Safety Data Management: Periodic Safety Update Reports for Marketed Drugs
Structure and Content of Clinical Study Reports
Dose Response Information to support Drug Registration
Ethnic Factors in the Acceptability of Foreign Clinical Data
Good Clinical Practice
Studies in support of Special Populations: Geriatrics
General Considerations for Clinical Trials
Statistical Principles for Clinical Trials
Choice of Control Group for Clinical Trials
Clinical Investigation of Medicinal Products in the Pediatric Population
Principles for Clinical Evaluation of New Antihypertensive Drugs
Clinical QT; Definitions in Pharmacogenetics, Pharmacogenomics
Quantification of Genomic Biomarkers
Multiregional Clinical Trials, Genomic Sampling and Safety Data Collection

And in addition, like in quality and safety, there are a number of multidisciplinary guidelines, clarifying the interrelationships of efficacy with quality and safety.

The essential objectives, scopes, and principles of the clinical guidelines are summarized in the following.

ICH Topic E1 the Extent of Population Exposure to Assess Clinical Safety for Drugs Intended for Long-Term Treatment of Non-life-Threatening Conditions (Step 5 in 1995)
E1 guideline was finalized under Step 4 in October 1994. This document gives recommendations on the numbers of patients and duration of exposure for the safety evaluation of drugs intended for the long-term treatment of non-life-threatening conditions.

ICH Topics E2 A to F

See ▶ Chap. 58, "General Principles of Pharmacovigilance in Clinical Development"

ICH Topic E3 Structure and Content of Clinical Study Reports (Step 5 in 1996)
The overall safety evaluation of the test drug(s)/investigational product(s) should be reviewed, with particular attention to events resulting in changes of dose or need for concomitant medication, serious adverse events, events resulting in withdrawal, and deaths. Any patients or patient groups at increased risk should be identified and particular attention paid to potentially vulnerable patients who may be present in small numbers, e.g., children, pregnant women, frail elderly, people with marked abnormalities of drug metabolism or excretion, etc. The implication of the safety evaluation for the possible uses of the drug should be described.

The discussion and conclusions should clearly identify any new or unexpected findings, comment on their significance, and discuss any potential problems such as inconsistencies between related measures. The clinical relevance and importance of the results should also be discussed in the light of other existing data. Any specific benefits or special precautions required for individual subjects or at-risk groups and any implications for the conduct of future studies should be identified.

ICH Topic E4 Dose Response Information to Support Drug Registration (Step 5 in 1994)

Dose-response data are desirable for almost all new chemical entities entering the market. These data should be derived from study designs that are sound and scientifically based; a variety of different designs can give valid information. The studies should be well-controlled, using accepted approaches to minimize bias. In addition to carrying out formal dose-response studies, sponsors should examine the entire database for possible dose-response information.

Dose-response data should be explored also for possible differences in subsets based on demographic characteristics, such as age, gender, or race. Knowledge whether there are pharmacokinetic differences among these groups, e.g., due to metabolic differences, differences in body habitus, or composition will be helpful.

Informative dose-response data, like information on responses in special populations, on long-term use, on potential drug-drug and drug-disease interactions, is expected but might, in the face of a major therapeutic benefit or urgent need, or very low levels of observed toxicity, become a deferred requirement.

ICH Topic E5 (R1) Ethnic Factors in the Acceptability of Foreign Clinical Data (Step 5 in 1998)

This guidance describes how a sponsor developing a medicine for a new region can deal with the possibility that ethnic factors could influence the effects (safety and efficacy) of medicines and the risk/benefit assessment in different populations. Results from the foreign clinical trials could comprise most, or in some cases, all of the clinical data package for approval in the new region, so long as they are carried out according to the requirements of the new region.

Acceptance in the new region of such foreign clinical data may be achieved by generating "bridging" data in order to extrapolate the safety and efficacy data from the population in the foreign region(s) to the population in the new region.

ICH Topic E6(R2) Guideline for Good Clinical Practice (Step 5 in 2017)

Clinical trials should be conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki and that are consistent with GCP and the applicable regulatory requirement(s).

Before a trial is initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks. The rights, safety, and well-being of the trial subjects are the most important considerations and should prevail over interests of science and society.

The available nonclinical and clinical information of an investigational product should be adequate to support the proposed clinical trial.

Clinical trials should be scientifically sound, and described in a clear, detailed protocol, which has received acceptance prior institutional review board (IRB)/independent ethics committee (IEC) approval/favorable opinion.

The medical care given to, and medical decisions made on behalf of, subjects should always be the responsibility of a qualified physician or of a qualified dentist.

Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s).

Freely given informed consent should be obtained from every subject prior to clinical trial participation. Guideline for good clinical practice E6(R2) EMA/CHMP/ICH/135/1995 Page 15/70 2.10. All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, and verification.

Addendum: The confidentiality of records that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

Investigational products should be manufactured, handled, and stored in accordance with applicable good manufacturing practice

(GMP). They should be used in accordance with the approved protocol.

ICH Topic E7 Studies in Support of Special Populations: Geriatrics Questions and Answers (Step 5 in 2010)

This guideline provides answers to a number of questions. As an example:

Why do we need an adequate representation of geriatric patients in the clinical database?

Geriatric patients can respond differently from younger patients to drug therapy in a number of ways, and such differences can be greater in patients 75 years and older:

(a) The geriatric population has age-related physiological changes that can affect the pharmacokinetics of the drug, and the pharmacodynamic response to the drug, both of which can influence the drug-response and the dose-response relationship.
(b) Geriatric patients are more prone to adverse effects since they often have comorbidities and are taking concomitant therapies that could interact with the investigational drug. The adverse effects can be more severe, or less tolerated, and have more serious consequences than in the non-geriatric population. With the increasing size of the geriatric population (including patients 75 and older) and in view of the recent advances in pharmacokinetics and pharmacodynamics since the ICH E7 guideline was established in 1993, the importance of geriatric data (from the entire spectrum of the geriatric patient population) in a drug evaluation program has increased. Not all potential differences in pharmacokinetics, pharmacodynamics, disease-drug and drug-drug interactions, and clinical response that can occur in the geriatric population can be predicted from non-geriatric populations, as the geriatric patients are far more likely to have multiple illnesses and to be receiving multiple drugs.

Therefore, to assess the benefit/risk balance of a drug that will be used in the geriatric population, these patients should be appropriately represented in clinical trials.

ICH Topic E8 General Considerations for Clinical Trials (Step 5 in 1998)

Objectives: In the three ICH regions, the evolution of drug development strategies and evaluation processes has led to the establishment of regional guidances on general considerations for clinical trials and the process of clinical development of pharmaceuticals for human use. This harmonized guideline is derived from those regional documents as well as from ICH guidelines.

The ICH document "General Considerations for Clinical Trials" is intended to:

(a) Describe internationally accepted principles and practices in the conduct of both individual clinical trials and overall development strategy for new medicinal products.
(b) Facilitate the evaluation and acceptance of foreign clinical trial data by promoting common understanding of general principles, general approaches, and the definition of relevant terms.
(c) Present an overview of the ICH clinical safety and efficacy documents and facilitate the user's access to guidance pertinent to clinical trials within these documents.
(d) Provide a separate glossary of terms used in the ICH clinical safety and efficacy related documents that pertain to clinical trials and indicate which documents contain them.

The term "drug" should be considered synonymous with "investigational (medicinal) product," "medicinal product," and "pharmaceutical product" including vaccines and other biological products.

The principles established in this guideline may also be applied to other clinical investigations (e.g., radiotherapy, psychotherapy, surgery, medical devices, and alternative therapies).

ICH Topic E9 Statistical Principles for Clinical Trials (Step 5 in 1998)

Scope and direction: The focus of this guidance is on statistical principles. It does not address the use

of specific statistical procedures or methods. Specific procedural steps to ensure that principles are implemented properly are the responsibility of the sponsor.

This guidance should be of interest to individuals from a broad range of scientific disciplines. However, it is assumed that the actual responsibility for all statistical work associated with clinical trials will lie with an appropriately qualified and experienced statistician, as indicated in ICH E6.

The role and responsibility of the trial statistician, in collaboration with other clinical trial professionals, is to ensure that statistical principles are applied appropriately in clinical trials supporting drug development.

For each clinical trial contributing to a marketing application, all important details of its design and conduct and the principal features of its proposed statistical analysis should be clearly specified in a protocol written before the trial begins. The protocol and subsequent amendments should be approved by the responsible personnel, including the trial statistician. The trial statistician should ensure that the protocol and any amendments cover all relevant statistical issues clearly and accurately, using technical terminology as appropriate.

The principles outlined in this guidance are primarily relevant to clinical trials conducted in the later phases of development, many of which are confirmatory trials of efficacy. In addition to efficacy, confirmatory trials may have as their primary variable a safety variable (e.g., an adverse event, a clinical laboratory variable, or an electrocardiographic measure) and a pharmacodynamic or a pharmacokinetic variable (as in a confirmatory bioequivalence trial).

Although the early phases of drug development consist mainly of clinical trials that are exploratory in nature, statistical principles are also relevant to these clinical trials.

The main principle is to minimizing bias and maximizing precision. It is important to identify potential sources of bias as completely as possible to draw valid conclusions from clinical trials. Bias can occur in subtle or unknown ways and its effect is not measurable directly.

ICH Topic E10 Choice of Control Group in Clinical Trials (Step 5 in 2001)

Control groups have one major purpose: to allow discrimination of patient outcomes (e.g., changes in symptoms, signs, or other morbidity) caused by the test treatment from outcomes caused by other factors, such as the natural progression of the disease, observer or patient expectations, or other treatment. The control group experience tells us what would have happened to patients if they had not received the test treatment or if they had received a different treatment known to be effective.

The choice of control group is therefore always a critical decision in designing a clinical trial. That choice affects the inferences that can be drawn from the trial, the ethical acceptability of the trial, the degree to which bias in conducting and analyzing the study can be minimized, the types of subjects that can be recruited and the pace of recruitment, the kind of endpoints that can be studied, the public and scientific credibility of the results, the acceptability of the results by regulatory authorities, and many other features of the study, its conduct, and its interpretation.

The general principles described in this guideline are relevant to any controlled trial, but the choice of control group is of particularly critical importance to clinical trials carried out during drug development to demonstrate efficacy. The choice of the control group should be considered in the context of available standard therapies, the adequacy of the evidence to support the chosen design, and ethical considerations.

ICH Topic E11 Clinical Investigation of Medicinal Products in the Pediatric Population (Step in 2001)

Objectives of the guidance: The number of medicinal products currently labeled for pediatric use is limited. It is the goal of this guidance to encourage and facilitate timely pediatric medicinal product development internationally. The guidance provides an outline of critical issues in pediatric drug development and approaches to the safe, efficient, and ethical study of medicinal products in the pediatric population.

Scope of the guidance: Specific clinical study issues addressed include:

1. Considerations when initiating a pediatric program for a medicinal product
2. Timing of initiation of pediatric studies during medicinal product development
3. Types of studies (pharmacokinetic, pharmacokinetic/pharmacodynamic (PK/PD), efficacy, safety)
4. Age categories
5. Ethics of pediatric clinical investigation

General principles: Pediatric patients should be given medicines that have been appropriately evaluated for their use. Safe and effective pharmacotherapy in pediatric patients requires the timely development of information on the proper use of medicinal products in pediatric patients of various ages and, often, the development of pediatric formulations of those products. Advances in formulation chemistry and in pediatric study design will help facilitate the development of medicinal products for pediatric use.

Drug development programs should usually include the pediatric patient population when a product is being developed for a disease or condition in adults and if it is anticipated that the product will be used in the pediatric population.

Obtaining knowledge of the effects of medicinal products in pediatric patients is an important goal. However, this should be done without compromising the well-being of pediatric patients participating in clinical studies. This responsibility is shared by companies, regulatory authorities, health professionals, and society as a whole.

ICH Topic E12 Principles for Clinical Evaluation of New Antihypertensive Drugs (Step 5 in 2000)

This document provides general principles for the clinical evaluation of new antihypertensive drugs. It describes accepted principles in the three ICH regions versus some region-specific differences. These differences may be harmonized in future but may require discussions with regional regulatory authorities.

General principles for the assessment of efficacy: The primary goal is the effect of the drug on systolic **and** diastolic blood pressures. In the past the primary endpoint of most studies was diastolic blood pressure. Although all drugs to date have reduced both systolic and diastolic blood pressures, the recognition of isolated or predominant systolic hypertension as a significant and remediable risk factor demands explicit evaluation of the effect of a drug on systolic blood pressure.

Many clinical trials of many interventions (including low- and high-dose diuretics, reserpine, and beta-blockers, usually as part of combination therapy) have shown consistent beneficial effects on long-term mortality and morbidity, most clearly on stroke and less consistently on cardiovascular events.

Whether some drugs or combinations have better effects than others on overall outcomes or on particular outcomes is not yet known. Formal mortality and morbidity outcome studies are not ordinarily required for approval of antihypertensive drugs, and the kind of active control mortality and morbidity studies that would be convincing is not well defined. Results of a large number of ongoing outcome studies could affect this policy and modify requirements. It should be noted that, even if an antihypertensive effect has been proven, a significant concern about a detrimental effect on mortality and/or cardiovascular morbidity might lead to a need for outcome studies.

ICH Topic E14 the Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs (Step 5 in 2005)

An undesirable property of some non-antiarrhythmic drugs is their ability to delay cardiac repolarization, an effect that can be measured as prolongation of the QT interval on the surface electrocardiogram (ECG). The QT interval represents the duration of ventricular depolarization and subsequent repolarization and is measured from the beginning of the QRS complex to the end of the T wave.

A delay in cardiac repolarization creates an electrophysiological environment that favors the

development of cardiac arrhythmias, most clearly torsade de pointes (TdP) but possibly other ventricular tachyarrhythmias as well.

TdP is a polymorphic ventricular tachyarrhythmia that appears on the ECG as continuous twisting of the vector of the QRS complex around the isoelectric baseline. A feature of TdP is pronounced prolongation of the QT interval in the supraventricular beat preceding the arrhythmia. TdP can degenerate into ventricular fibrillation, leading to sudden death.

While the degree of QT prolongation is recognized as an imperfect biomarker for proarrhythmic risk, in general there is a qualitative relationship between QT prolongation and the risk of TdP, especially for drugs that cause substantial prolongation of the QT interval.

Because of its inverse relationship to heart rate, the measured QT interval is routinely corrected by means of various formulae to a less heart rate-dependent value known as the QTc interval. It is not clear, however, whether arrhythmia development is more closely related to an increase in the absolute QT interval or QTc. Most drugs that have caused TdP clearly increase both the absolute QT and the QTc (hereafter called QT/QTc).

Documented cases of TdP (fatal and nonfatal) associated with the use of a drug have resulted in the withdrawal from the market of several drugs and relegation of other drugs to second-line status. Because prolongation of the QT/QTc interval is the ECG finding associated with the increased susceptibility to these arrhythmias, an adequate premarketing investigation of the safety of a new pharmaceutical agent should include rigorous characterization of its effects on the QT/QTc interval.

Objectives: This document provides recommendations to sponsors concerning the design, conduct, analysis, and interpretation of clinical studies to assess the potential of a drug to delay cardiac repolarization. This assessment should include testing the effects of new agents on the QT/QTc interval as well as the collection of cardiovascular adverse events.

The assessment of the effects of drugs on cardiac repolarization is the subject of active investigation. When additional data (nonclinical and clinical) are accumulated in the future, this document may be reevaluated and revised.

Scope: The recommendations of this document are generally applicable to new drugs having systemic bioavailability but may not apply to products with highly localized distribution and those administered topically and not absorbed. The focus is on agents being developed for uses other than the control of arrhythmias, as antiarrhythmic drugs can prolong the QT/QTc interval as a part of their mechanism of clinical efficacy. While this document is concerned primarily with the development of novel agents, the recommendations might also be applicable to approved drugs when a new dose or route of administration is being developed that results in significantly higher exposure (i.e., Cmax or AUC).

Additional ECG data might also be considered appropriate if a new indication or patient population was being pursued.

The evaluation of the effect of a drug on the QT interval would also be considered important if the drug or members of its chemical or pharmacological class have been associated with QT/QTc interval prolongation, TdP, or sudden cardiac death during post-marketing surveillance.

This guideline should be read together with the preclinical guideline ICH S7 B.

The Expert Groups of E14 and S7 B have and will continue to cooperate to create strategies which will provide greater safety for patients.

The ICH guidelines E15, E16, and E18

ICH Topic E15 Definitions for genomic biomarkers, pharmacogenomics, pharmacogenetics, genomic data, and sample coding categories (Step 5 in 2008)

ICH Topic E16 Genomic Biomarkers Related to Drug Response: Context, Structure and Format of Qualification Submissions (Step 3 in 2009)

ICH guideline E18 on genomic sampling and management of genomic data (Step 5 in 2018)

ICH Guideline E17 on General Principles for Planning and Design of Multiregional Clinical Trials (Step 5 Coming into Effect in June 2018)

Objectives of the guideline: With the increasing globalization of drug development, it has become

important that data from multiregional clinical trials (MRCTs) can be accepted by regulatory authorities across regions and countries as the primary source of evidence, to support marketing approval of drugs (medicinal products).

The purpose of this guideline is to describe general principles for the planning and design of MRCTs with the aim of increasing the acceptability of MRCTs in global regulatory submissions.

Background: Data from MRCTs are often submitted to multiple regulatory authorities without a previously harmonized regulatory view on the development program. There are currently no ICH guidelines that deal specifically with the planning and design of MRCTs, although the ICH E5 guideline covers issues relating to the bridging of results from one region to another. MRCTs conducted according to the present guideline will allow investigation of treatment effects including safety evaluations in the overall population as well as investigations of the potential impact of intrinsic and extrinsic factors, thus offering an earlier access to new drugs worldwide.

Scope of the guideline: MRCT is defined in the present guideline as a clinical trial conducted in more than one region under a single protocol. Such data will be submitted to multiple regulatory authorities for drug approval (including approval of additional indications, new formulations, and new dosing regimens) and for studies conducted to satisfy post-marketing requirements.

ICH E19 Optimization of Safety Data Collection (Concept Paper of ICH in August 2017)

This new guideline is proposed to provide harmonized guidance on when it would be appropriate to use a targeted approach to safety data collection in some late-stage premarketing or post-marketing studies and how such an approach would be implemented.

Recognizing that protection of patient welfare during drug development is critically important, unnecessary data collection may be burdensome to patients and serve as a disincentive to participation in clinical research. By tailoring safety data collection in some circumstances, the burden to patients would be reduced, a larger number of informative clinical studies could be carried out with greater efficiency, studies could be conducted with greater global participation, and the public health would be better served. The proposed guideline would be consistent with risk-based approaches and quality-by-design principles.

Detailed Overview for Guidelines for Safety

In the following a survey is given in regard to the guidelines on preclinical safety (Toxicology or Safety Pharmacology).

Carcinogenicity S 1 A-C

There are a number of endpoints which should not be tested in humans; these are mainly the assessment of genotoxicity, teratogenicity, and cancerogenicity. Especially the cancerogenic risk can usually not be investigated in humans; it is ethically forbidden and such a risk by new drugs for patients is unacceptable. The potential to induce tumors, therefore, can only be evaluated by nonclinical studies.

The nonclinical testing for carcinogenic potential today employs short-, mid-, and long-term studies in rodents. Such studies are considered to have a relatively high power of predictivity for the carcinogenic risk in humans.

General Regulatory Background

Treatment with compounds associated with carcinogenic potential is unacceptable for banal indications; for severe indications like life-threatening cancer diseases, the treatment with carcinogenic compounds often does not increase the overall risk of the underlying disease the patient suffers from.

Tumors are usually the result of a multihit, multistep long-term process which needs in humans many years. These long durations are reflected in animal testing: carcinogenicity studies are often lifelong studies; the duration in these models is mostly 24 months in rats and today also in mice (former duration was 18 months in mice). Carcinogenicity studies are therefore the most expensive (approximately 1.5–4 million dollars or euros)

Table 5 Usual strategies for cancerogenicity studies

Carcinogenicity testing: the usual way
Usually in rodents (rats, mice)
Usually starts with 3 month dose-range-finding study followed by the main study
Usually over a treatment period of 2 years
Usually 3 treatment groups + control
Usually needed for MAA/NDA(marketing authorization application, new drug application)

Table 6 When are carcinogenicity studies needed?

Need for carcinogenicity studies? (ICH S1A)
Clinical use for >6 months(continuous or intermittent)
Known carcinogenic potential for product class
Structure-activity relationship suggesting carcinogenic risk
Preneoplastic lesions in repeated dose toxicity studies
Long-term tissue retention resulting in local tissue reactions or other pathophysiological responses
Equivocal genotoxicity tests

preclinical studies. Accordingly, they should be well designed and conducted in such a way that they clearly indicate any risk involved. The basis of the strategies is always literature research, what is known from the chemical class of the new drug, which pharmacodynamic effects can be expected, detail by detail is collected for a weight-of-evidence approach before the experimental part starts with dose-range finding tests (Tables 5 and 6).

Guideline on the Need for Carcinogenicity Studies of Pharmaceuticals (S1A) (2)

This guideline was adopted in 1997 and implemented in all regions.

The objectives of carcinogenicity studies are:

- To identify a tumorigenic potential in animals
- To assess the relevance of these identified risks for humans

Any cause for concern derived from laboratory investigations, animal toxicology studies, and data in humans may lead to a need for carcinogenicity studies.

Table 7 When are bioassays not necessary?

No need for carcinogenicity studies? (ICH S1A)
Short-term indications
For example, anesthetics, diagnostics
Unequivocally genotoxic compounds
Assumption of trans-species carcinogens
Low life expectancy of (tumor)patients (2–3 years)
Prolongation of survival may trigger carcinogenicity testing
Topical use drugs without systemic exposure
Unless there is cause for concern

Carcinogenicity studies should be performed for any pharmaceutical whose expected clinical use is continuous for at least 6 months. Certain classes of compounds may not be used continuously over a minimum of 6 months but may be expected to be administered repeatedly in an intermittent manner. For pharmaceuticals used frequently in an intermittent manner during the treatment of chronic or recurrent conditions, carcinogenicity studies are generally also needed. Examples of such conditions include allergic rhinitis, depression, and anxiety.

Pharmaceuticals administered infrequently or for short duration of exposure (e.g., anesthetics and radiolabel imaging agents) do not need carcinogenicity studies (Table 7). The time after consumption is too short to cause any clinically relevant damages leading to neoplasms.

But on the other hand, there may be causes for concern. Examples are:

- Previous demonstration of carcinogenic potential in the product class that is considered relevant to humans
- Structure-activity relationship suggesting carcinogenic risk
- Evidence of preneoplastic lesions in repeated dose toxicity studies
- Long-term tissue retention of parent compound or metabolite(s) resulting in local tissue reactions or other pathophysiological responses

Unequivocally genotoxic compounds need not to be subjected to long-term carcinogenicity studies. It is assumed that these products, damaging

the DNA, will cause neoplasias. The S1A guideline offers in addition an alternative: if a drug is intended to be administered chronically to humans, a chronic toxicity study (up to 1 year) may be necessary to detect early tumorigenic effects. In practice, this option has not been used since 1997. Main concern against this type of studies by industry is the lack of historical comparison data. The tumor evaluation was always based on 2-year data.

Other conditions which justify not conducting carcinogenicity assays are summarized in (Table 7).

- In instances, where the life expectancy in the indicated population is short (i.e., less than 2–3 years), no long-term carcinogenicity studies may be required. For example, oncolytic agents intended for treatment of advanced systemic disease do not generally need carcinogenicity studies.
- In cases where the therapeutic agent for cancer is generally successful and life is significantly prolonged, there may be requirements to provide knowledge about the tumorigenic risk.
- When such pharmaceuticals are intended for adjuvant therapy in tumor-free patients or for prolonged use in non-cancer indications, carcinogenicity studies are usually needed.
- Pharmaceuticals showing poor systemic exposure from topical routes in humans may not need studies by the oral route to assess the carcinogenic potential to internal organs.
- Carcinogenicity studies are not generally needed for endogenous substances, when given essentially as replacement therapy (i.e., physiological levels), particularly, where there is previous clinical experience with similar products (e.g., animal insulins, pituitary hormones).

In general, the need for the 2-year rodent assay to assess a carcinogenic potential continues to be questioned because of too many positive outcomes. In addition, retrospective analyses of various datasets (PhRMA, FDA, JPMA, and EU) concluded that based on genotoxicity and non-genotoxic mechanisms, detectable in pharmacology and chronic toxicity data (usually present at the end of phase 2 in the development of a new pharmaceutical), the outcome of the 2-year rat carcinogenicity study could be predicted with reasonable assurance (Table 8).

Table 8 Conditions of a negative outcome of cancerogenicity studies

Negative prediction, initiative of PhRMA	
Rat chronic toxicology studies =	
→	Good predictors of negative outcome in 2-year rat carcinogenicity studies:
(a) Analysis of 182 compounds:	
If	No preneoplasia in chronic toxicity studies
	No genotoxicity
	No hormonal perturbation signals
	No immunosuppression
→	No value added from 2-year rat carc. study

Sistare FD et al. (2011), Toxico- Pathol;39(4):716–744

There are two extremes of the spectrum: Negative predictions can be made when predictive carcinogenic signals are absent, and positive predictions are possible when such signals are present. In between a category of compounds still remain for which the outcome cannot be predicted with sufficient certainty and where experimental studies may add value to identify real hazards.

These hypotheses stimulated ICH to test in an ongoing common exercise by Drug Regulatory Agencies and Industry if there are chances to reduce the number of rodent bioassays. Such prospective evaluation is necessary to justify any revision of the present recommendations of the ICH guideline S1 (EMA 2016). Sponsors are strongly encouraged to submit Carcinogenicity Assessment Documents (CADs) to Drug Regulatory Agencies (DRAs) for all investigational pharmaceuticals with ongoing or planned 2-year rat carcinogenicity studies. The CAD would address the overall carcinogenic risk of the investigational drug as predicted by the available knowledge on pharmacology and toxicology and a rationale for why the conduct of long-term studies would or would not add value to that assessment, in the latter case by a request of a "virtual" waiver. DRAs independently review the submitted documents and evaluate the degree of concordance with sponsors.

During this prospective evaluation period, waiver requests will not be granted, but the data

are intended solely for collecting real-world experience.

Submitted CADs will finely be compared to the real outcome of the 2-year carcinogenicity studies and/or any other factors of a weight-of-evidence evaluation. Main objective will be the assessment of accuracy of the predictions, with emphasis on the "virtual" waivers.

The final conclusion was originally intended to be based on 50 examples; by the end of 2017, 46 cases were available; they are considered sufficient to modify any conclusions. Their prospective assessments by industry and agencies are now compared to the final outcome of the real long-term studies.

There seems hope that a revision of the guideline S1 A may be justified. The competent ICH Expert Groups will come to a conclusion in 2019.

Testing for Carcinogenicity of Pharmaceuticals (S1B)

S1B was adopted and implemented in 1997.

Historically, the regulatory requirements for the assessment of the carcinogenic potential were the conduct of long-term carcinogenicity studies in two rodent species, usually the rat and the mouse. It was the mission of ICH to examine whether this practice could be reduced without compromising human safety. The discussion in the Expert Working Group soon revealed that the rule of testing in two species had to be resumed and the American consumer societies demanded continuation with the same standard of safety as before, based on two species testing.

But more flexibility in the strategies was developed. As a new experimental approach to test for carcinogenic potential, a basic scheme was set up to comprise one long-term rodent carcinogenicity study, plus one other study of the type that supplements the long-term carcinogenicity study and provides additional information that is not readily available from the long-term assay:

S1B: Basic Principle for Testing the Carcinogenic Potential (Fig. 3):

- One long-term rodent carcinogenicity study plus
- One short- or medium-term study that supplements the long-term carcinogenicity study and provides additional information not readily available from the long-term assay

The species selected should be appropriate, based on considerations on pharmacology, repeated-dose toxicology data, metabolism (see also guidelines S1C and S3A), toxicokinetics (see also guidelines S1C, S3A, and S3B), and route of administration (e.g., less common routes such as dermal and inhalation). In the absence of clear evidence favoring one species, it is

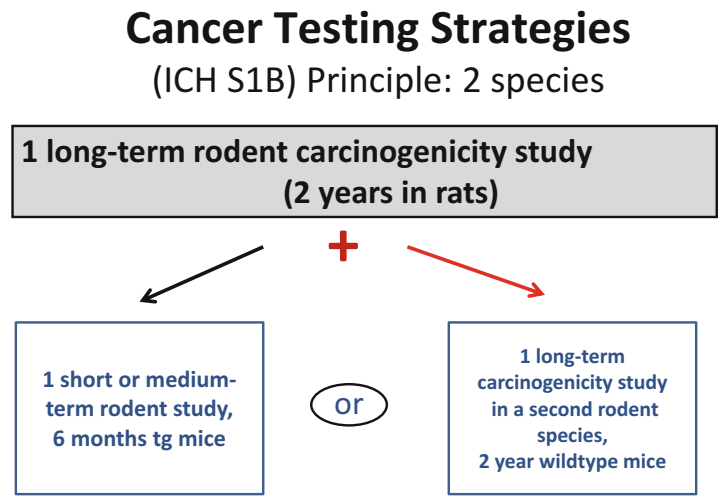


Fig. 3 Options for assessment of carcinogenic potential

recommended that the rat be selected for the long-term study, because data collections by the European, American, and Japanese colleagues resulted in the conclusion that the rat was the more sensitive species and historical data were more meaningful in this species.

Additional tests may either be short- or medium-term in vivo rodent test systems, usually selecting the mouse. According to the guideline, these models of carcinogenesis may use transgenic or neonatal rodents or may include models of initiation and promotion in rodents (but the latter option is today considered to be useful models for hepatic carcinogenesis or adequate mechanistic studies, but not as assays appropriate as general screen for drug-induced potential for carcinogenesis). Also neonatal animals are hardly used any longer.

All these models were assessed with approximately 40 different compounds in a strenuous study by the International Life Science Institute (ILSI/HESI) (Cohen et al. 2001).

The most appropriate models today are transgenic mouse models with either activated oncogenes, like Tg.rasH2 model or Tg.AC skin model (favored especially in the USA) or a transgenic model with inactivated tumor suppressor gene, the p53 "knock out" (=p53+/−) model (Table 9).

Table 9 Transgenic mouse models in practice 2018

Transgenic mouse models
Activated oncogenes =
TgrasH2 model (Japan)
Tg.AC skin model (USA); also gavage
Inactivated tumor suppressor gene =
p53 "knock out" (=p53+/−) model (USA)

The Food and Drug Administration (FDA) supports the selection of these models and considers that $P53^{+/-}$is the right choice, if a compound is clearly or equivocally genotoxic. The Tg.AC model is adequate for dermally applied products, neonatal mice have been used when the drug is clearly or equivocally genotoxic, and the Tg.Hras2 can be recommended for genotoxic or non-genotoxic products.

The EU/EMA concluded in addition that p53 and Tg.rasH2 are equally sensitive to genotoxic compounds (although some false positives or false negatives have been identified). Tg.rasH2 is more sensitive to peroxisome proliferators, and altogether p53 and Tg.rasH2 are acceptable in a regulatory context as alternatives.

The guideline describes in the "Notes" important information about the new models.

Note 1 informs about the Syrian hamster embryo (SHE) cell-transformation assay (Fig. 4) Mauthe et al. 2001 reviewed the SHE cell-transformation assay and described the following:

The Syrian hamster embryo (SHE) cell-transformation assay represents a short-term in vitro assay capable of predicting rodent carcinogenicity of chemicals with a high degree of concordance (LeBoeuf et al. 1996). The SHE assay model identifies the earliest identifiable stage in carcinogenicity, morphological cell transformation. In contrast to other short-term in vitro assays, both genotoxic and epigenetic

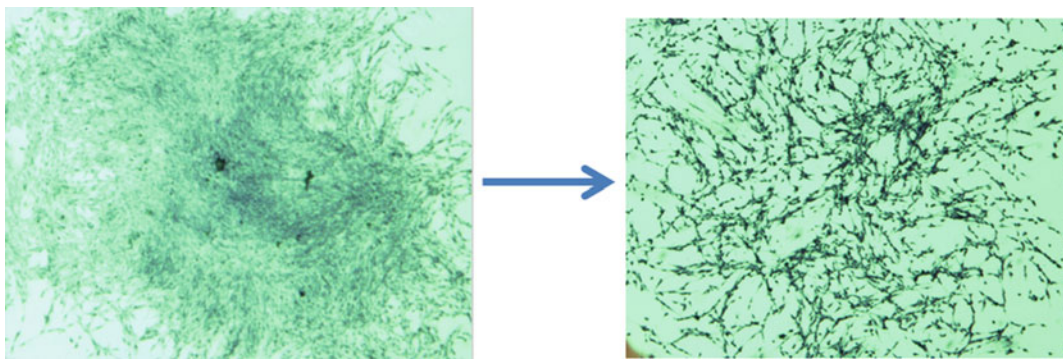


Fig. 4 Negative control and positive result of Syrian hamster embryo test

carcinogens are detected. The SHE assay, originally developed by Berwald and Sachs (1965) and modified as described by LeBoeuf and Kerckaert 1986, was included in the International Life Sciences Institute, Health and Environmental Sciences Institute (ILSI/HESI). Alternative carcinogenicity testing (ACT) collaboration provides additional information on the use of short-term in vitro tests in predicting carcinogenic potential. A total of 19 ILSI compounds have been tested in the SHE assay: 15 were tested for this project, whereas clofibrate, methapyrilene, reserpine, and di(2-ethylhexyl)phalate (DEHP) were tested previously. Of the three non-carcinogenic compounds tested, two were negative in the SHE assay, whereas ampicillin was tested positive. The remaining 16 compounds tested were either known rodent carcinogens and/or human carcinogens. From this group, 15 tested positive in the SHE assay, whereas phenacetin, a genotoxic carcinogen, was tested negative.

Therefore, overall concordance between the SHE assay and rodent bioassay was 89% (17/19), whereas concordance with known or predicted human carcinogens was 37% (7/19).

Based on these data, it is concluded that the SHE cell-transformation assay has utility for predicting the results of the rodent carcinogenesis bioassay but lacks the selectivity to distinguish between rodent and human carcinogens.

Note 2 on conditions to limit testing with one species only: if the findings of a short- or long-term carcinogenicity study and of genotoxicity tests and other data indicate that a pharmaceutical clearly poses a carcinogenic hazard to humans, a second carcinogenicity study would not usually be useful.

Note 3 provides details about the short- or mid-term models.

Evidence of tumorigenic effects of the drug in rodent models should be evaluated in light of the tumor incidence and latency, the pharmacokinetics of the drug in the rodent models as compared to humans, and data from any ancillary or mechanistic studies that are informative with respect to the relevance of the observed effects to clinical conditions.

Table 10 Types of cancerogenicity studies for MAA. Categorization of active substances with carcinogenicity data according to type and number of studies performed

Active substances with carcinogenicity data	Number	%
2 long-term carcinogenicity studies	116	80.5
1 long-term carcinogenicity study in rats +1 transgenic mouse study	8	5.5
1 long-term carcinogenicity study in mice or rats	13	9
1 transgenic mouse model	1	1
No carcinogenicity studies performed	6	4
Total	144	100

Source: Friedrich and Olejniczak (2011)

The results from any tests cited above should be considered as part of the overall "weight-of-evidence" taking into account the scientific status of the test systems.

It needs finally to be stressed that a long-term carcinogenicity study in a second rodent species (e.g., mice) is still considered acceptable.

When looking back at recent applications, these options of life span studies in the mouse have not been replaced by transgenic models. This is the result of the analyses of Friedrich and Olejniczak in 2011 reflecting the situation until 2009 (Table 10). It looks that there is a trend in experimental strategies that transgenic mice are increasingly part of the routine testing paradigm. Further observations should be performed.

Dose Selection for Carcinogenicity Studies of Pharmaceuticals (S1C)

International Conference on harmonization (1997): Guideline on dose Selection for carcinogenicity studies of Pharmaceuticals

This guideline was adopted and implemented in the different regions in 1997.

Traditionally, carcinogenicity studies for chemical agents have relied upon the maximally tolerated dose (MTD) as the standard method for high-dose selection.

The MTD is generally chosen based on data derived from toxicity studies of 3 months'

duration. Testing options for dose-range-finding studies are as follows:

- Usually, 3 months for long-term studies or 1 month for neonatal or transgenic mice with a range of different dose levels, often 5.
- There is a focus on toxicity endpoints, determination of MTD with profiling of AUC (e.g., 1 and 3 months for rats or 1 and 4 weeks for alternatives).
- Toxicokinetics: In many cases it is acceptable by agencies to use instead of transgenic animals their wild type associates.

Ideally, the doses selected for rodent bioassays for non-genotoxic pharmaceuticals should provide an exposure to the agent, and an adequate margin of safety, without any significant severe chronic dysfunction and compatible with good survival of up to 2 years.

The guideline calls for a flexible approach to dose selection. The guideline proposes five different approaches (Tables 11 and 12):

Table 11 Criteria for high-dose selection

Criteria of high-dose selection (ICH S1C)
Maximum-tolerated dose (MTD) (10% of weight loss compared to controls)
25-fold AUC ratio (rodent/human)
Saturation of absorption
Dose-limiting pharmacodynamic effects (e.g., hypotension, inhibition of blood clotting)
Maximum feasible dose, limit dose: 1500 mg/kg

Table 12 Examples for recommendation in the EU guidance for cancerogenicity assays

EMA, CPMP (2003) note for guidance on carcinogenic potential
Historical control data not older than 5 years, same strain and testing facility
Histopathological terms according to well-defined classifications (e.g., (ILSI, StP, IARC, Reni)
Ideally, one *pathologist* with board certification responsible for histological evaluation
Peer-review (slides) required for 10% of all tumors

1. Toxicity-based endpoints (MTD = maximum-tolerated dose or minimum toxic dose)
2. Pharmacokinetic endpoints (25 times the human AUC)
3. Saturation of absorption
4. Pharmacodynamic endpoints
5. Maximum feasible dose

Ad 1) The ICH Expert Working Group on safety has agreed to continue the use of the MTD as an acceptable toxicity-based endpoint for high-dose selection for carcinogenicity studies.

The MTD is defined as the top dose or maximum-tolerated dose that produces a minimum toxic effect over the course of the carcinogenicity study. Factors to consider are alterations in physiological functions, which would alter the animal's normal life span or interfere with interpretation of the study. Such factors include no more than 10% decrease in body weight gain relative to controls and target organ toxicity or significant alterations in clinical pathological parameters.

Ad 2) For relatively well-tolerated drugs, the alternative of a systemic exposure representing 25 times multiple of the human AUC (at the maximum recommended daily dose) may be an appropriate endpoint for dose selection for carcinogenicity studies for non-genotoxic pharmaceuticals, as a pragmatic solution without any clear-cut scientific justification but supported by data indicating that such a compound does not induce severe toxicities at lower exposure levels.

Ad 3) Limitation of absorption: High-dose selection based on saturation of absorption measured by systemic availability of drug-related substances is acceptable. The mid- and low-doses selected for the carcinogenicity study should take into account saturation of metabolic and elimination pathways. If such saturation is identified, then there is no need to raise the dose levels, since exposure will not increase.

Ad 4) Pharmacological properties: Pharmacodynamic endpoints for high-dose selection will be highly compound-specific. The high-dose

selected should produce a pharmacodynamic response in dosed animals of such magnitude as would preclude further dose escalation. However, the dose should not produce disturbances of physiology or homeostasis. Examples include hypotension, inhibition of blood clotting, or insulin-like effects. Too high exposure would exclude survival up to 2 years and jeopardize the study conclusion.

Ad 5) The maximum feasible dose by dietary administration was considered 5% of diet. By many scientists this amount of drug is considered to be too high. Therefore, a new and more reasonable solution was formulated in the guideline ICH/S1C(R), which follows.

Addendum to Dose Selection for Carcinogenicity Studies of Pharmaceuticals

Addition of a limit dose and related notes of pharmaceuticals.

This addendum S1C (R) was adopted in 1998.

In determining the high dose for carcinogenicity studies, it may not be necessary to exceed a dose of 1500 mg/kg/day. This limit dose applies only in cases where there is no evidence of genotoxicity and where the maximum recommended human dose does not exceed 500 mg/day.

This limit dose helps to calculate the amount of compound needed for such long-term assays. Compared to the upper options, it is not applied very often, since in most cases drug-induced toxicity can be identified.

Genotoxicity Guidelines (ICH/S2A, S2B and S2(R1))

A permanent alteration of genes or chromosomes can cause heritable effects leading to malformations and dysfunctions in the next generation or inducing tumors in the individual patient.

The former genotoxicity guidelines: Genotoxicity Guidelines (ICH/S2A and S2B), Genotoxicity: Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals, and Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals have been replaced in 2011 by the integrated Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use S2(R1) (Current Step 4 version dated 9 November, 2011).

The main reasons to revise the previous guidelines were to merge them into one concise guideline, to refocus the sensitivity and specificity of the in vitro assays (e.g., via restriction of upper concentrations to be tested), to include the option to replace these in vitro assays by using mammalian cells, and to define a suitable in vivo test battery option according to improved scientific standards and guidelines from other sources (e.g., OECD guidelines).

The revised guideline continues to support the goals of testing to deselect mutagenic/genotoxic drug candidates as early as possible in the development of new molecular therapeutic entities to contribute to the safety of humans.

In terms of coverage of genetic endpoints, valid information on gene mutations, structure chromosome aberrations (clastogenicity), and numerical chromosome aberrations (aneugenicity) is required.

In genetic toxicology, no single test is capable of detecting all relevant genotoxic agents; therefore, a battery of tests is considered appropriate. It traditionally starts with a test for gene mutations in bacteria (Ames). Ames published this test as early as 1973, and it remains in the screening battery also for its easiness of conduct and fast/reliable turnaround of data. It is considered to be robust and predictive for mutagenic carcinogens (Kirkland et al. 2006).

As a second test, the ICH S2R guideline offers, besides the in vitro chromosomal aberration test and the mouse lymphoma tk assay, an in vitro micronucleus test (the latter test was not available with a standard protocol before).

This assay is supported by many years of validation and considered to be an adequate alternative to the traditional chromosome test (Corvi et al. 2008).

Regarding the dose and exposure levels having been used, awareness increased that this procedure did often substantially exaggerate clinical exposure conditions. Accordingly, concentrations of 1 mM (instead of 10 mM) were considered sufficient for nontoxic drug candidates in

mammalian cells in vitro. And additionally, these perspectives were further strengthened by limiting the levels of cytotoxicity of not more than 50%.

ICH generally pursues the important goal of a reduction in animal testing in all experimental strategies. While traditionally an area of importance of in vitro models, this principle has been followed also in S2 R. The assessment of genotoxicity should, where possible, be integrated into repeat dose toxicity studies, e.g., into 4-week rodent studies, where the bone marrow micronucleus test could help to optimize the usage of animals (Fig. 5).

Modifications of the standard battery may be necessary for some classes, e.g., antibiotics which are toxic to bacteria or, e.g., for compounds like topoisomerase inhibitors which interfere with the mammalian cell replication system.

A selection of additional assays is being proposed; further modifications may be acceptable via discussion in the ICH maintenance process. Alternative strategies may consider assays like the in vivo comet assay (single cell gel electrophoresis measuring DNA strand breaks) or gene mutation tests with transgenic animals or in vivo DNA adduct studies.

Support for the interpretation of positive test results and considerations on conditions leading to false-positive data can be found in many excellent publications of this special field of science. It is sharing this special emphasis with the carcinogenicity testing guidelines in terms of the general notion that effects, once incurred, are not reversible.

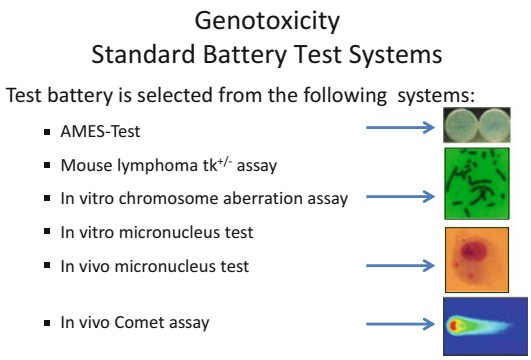


Fig. 5 Selection of genotoxicity assays

Toxicokinetics/Pharmacokinetics (ICH/S3A and S3B)

ICH/S3A

The objectives of toxicokinetics are primarily to describe the systemic exposure achieved in animals and its relationship to dose level and the time course of the toxicity study and, further, to relate exposure levels to toxicological findings, to assess the relevance of these findings to clinical safety, and to support the design as the choice of species and treatment regimen in nonclinical studies.

If animals or humans are exposed to pharmaceutical compounds, they will either elicit a pharmacodynamic effect, e.g., show a suppression of blood pressure, or reveal in analytical blood samples exposure levels of the compound. This kinetic information of the parent compound and its metabolites is an important contribution for the extrapolation of safety data from animal studies to humans. Species differ considerably in regard to their kinetic conditions as C_{max}, T_{max}, area under the curve (AUC), $t_{1/2}$, and ADME (absorption, distribution, metabolism, and excretion).

Therefore, it is important to know what a drug does with the body (what pharmacology and toxicology is induced?); it is also crucial to know what the body does with the drug.

The following toxicokinetic studies are usually distinguished: concomitant toxicokinetics, which are normally integrated in the toxicity studies and other supportive studies which mimic the conditions of the toxicity study.

The focus before IND (Table 13) is on T_{max}, C_{max}, and AUCs, while the complexity of pharmacokinetic characterization (like oral bioavailability, plasma half-life, volume of distribution, mean residence time, absorption, solubility, and concentration) is built up during clinical trials and on the basis of comparable human data.

Kinetic data should be considered in repeat dose toxicity studies, in genotoxicity studies when there are discrepancies between in vitro and in vivo assays and sufficient exposure can be characterized in the indicator tissue confirming that lack of exposure was not the reason for a negative reaction. Such data are useful for the evaluation of reproductive tests and in

Table 13 Kinetic data and timing

Kinetic information before first in man
Area under the curve (AUC)
Cmax and Tmax
In vitro metabolism (P 450 cytochrome)
Plasma protein binding data for animals and humans
Kinetic information during clinical trials
ADME (animal and humans) (absorption, distribution, metabolism, excretion)
In vivo metabolism animal versus human
Bioavailability
Half-life
Accumulation versus inhibition
Safety margins

cancerogenicity studies where often high exposures cause the appearance of adverse effects.

The compound can be bound to plasma proteins, erythrocytes, or other cells or tissues. Therefore, a distinction between "unbound drug" and "free fraction" is relevant. Distribution studies help to optimize the design of preclinical studies. Demonstration of accumulation can, for example, explain toxicity at the site of increased compound accumulation. The affinities of different organs to some drugs can vary considerably.

The International Conferences deal with increasing knowledge; scientific data more and more dominate the regulatory area. Translateral data exchange increases. Adverse effects need to be interpretated in the light of dose levels or more precisely of exposure levels.

When in 1995 ICH S3 was created, the rule was accepted that all safety data should be supported by kinetic analyses. But today this guideline is more than 20 years old; it needs urgently a revision. AUC and cmax alone are not sufficient any more at the start of development. In contrast to the guideline text, cancerogenicity studies need data also after longer periods than 6 months, e.g., after 12 and 24 months, sometimes including more metabolic refinement. Consider in that context that some drugs induce after longer-term exposure, e.g., hepatotoxicity or renal lesions, leading to increases of blood concentrations of parent compounds or metabolites.

Therefore, already before FIM (before first human clinical trials), evaluation of **in vitro** metabolic data (cytochrome p450) is needed (see ICH M3 R). Plasma ***protein binding*** data for animals and humans are essential before exposing large numbers of humans (Phase III). Data for **ADME** (administration, distribution, metabolism, and excretion) must be compared in animals and humans. This facilitates in the end the decision, if the preclinical safety assessment had used the *Most human-like* animal species for optimal predictions for both the desirable and undesirable effects for patient conditions.

These objectives are supported by new progresses in techniques; there are new sampling techniques, e.g., microsampling, blood spot analyses, tissue samples, or in vitro concentrations; there are more sophisticated and refined analytical procedures, and there is an increased interrelationship of kinetics and pharmacological/toxicological qualifications, facilitated by an increasing number of identified biomarkers, e.g., genomics.

ICH/S3B

Single-dose studies provide usually sufficient information about tissue distribution, but there may be cases where assessments after repeated dosing may provide better information. Such studies are necessary when:

1. Single-dose distribution studies suggest that the half-life of the test compound and/or metabolites in organs or tissues significantly exceeds the half-life of the elimination phase in plasma.
2. Steady-state levels of a compound/metabolite in the circulation, determined in repeated dose pharmacokinetic or toxicokinetic studies, are markedly higher than those predicted from single-dose kinetic studies.
3. When histopathological changes were observed that were not predicted from short-term toxicity studies.

This information is provided in the ICH guideline S3B: "Pharmacokinetics: Guidance for

repeated dose distribution studies (CPMP/ICH/395/95)."

Other relevant kinetic questions are the investigations of the potential of compounds to penetrate the barriers of placenta, blood-brain, or excretion into milk.

Knowledge about metabolites and their activity is further an important criterion for the assessment of species-specific effects and differences, e.g., the search for the most human-like test model as the best predictor for human reactions focuses on such differences. Metabolism can lead to pharmacologically active metabolites; such knowledge is desirable early in development. In vitro metabolism studies normally precede in vivo preclinical safety assessments.

For safety reasons, it is important to identify, and perhaps eliminate, drugs from further development if they are subject to polymorphic metabolism or extensive metabolism by key human enzymes. Knowledge about the cytochrome P450 (CYP450) superfamily of drug metabolizing enzymes is of particular interest.

ICH/S4A (20)

This is one of the guidelines where harmonization continued to be difficult for a long time. This lack of harmonization to recommend clear advice for duration led to the decision for pharmaceutical companies to perform partially duplicative studies for both 6 and 12 months duration. As a consequence the regulatory authorities in the EU, Japan, and USA wondered whether a more unified duration for chronic toxicity testing could be identified.

There was clear agreement that **rodent** studies were only necessary for a maximum duration of 6 months. When a 6-month study in rodents is conducted, then continuous application of pharmaceutical compounds to the indicated population can be assumed. Under such conditions, a long-term carcinogenicity study will usually be performed. This allows long-term exposure and lifelong observation and compensates any lack of long repeated observations.

For **non-rodents**, lifelong exposure is not required, but which duration can be considered as sufficient? The EU proposed as a start a common maximum duration of 6-month studies for rodents and non-rodents.

A unified analysis of data from 6- to 12-month dog studies was performed by the ICH members. In 16 cases there were findings observed by 12 months, but not by 6 months. It was concluded that these would or could have been detected in a study of 9 months duration. An agreement on the clinical relevance of these findings could not be reached.

The guideline S4A recommends the following: For non-rodents, 12-month studies are usually not necessary, and in the EU 6-month studies are acceptable based on the Directive 75/318/EEC. Accordingly, as a compromise, the Expert Working Group agreed that 9-month studies could be recommended in general.

In the US Federal Register, this guideline was published with an FDA Note on ICH/S4A as follows: *9-month studies are acceptable for most development plans; shorter ones like 6-month studies may be acceptable for some drugs, while longer durations,* e.g., *12 months, may be more appropriate for others.*

This advice is characterized by the FDA as their current line of thinking. In practice, FDA agreed with the sponsors in 50% for 9-month studies and in 32% for 6-month studies, while 12-month studies were only requested for compounds with novel mechanisms, or when only sparse clinical data were available, as in indications like HIV. Aids patients should be provided early on with the new treatment, and the lack of complex clinical data is compensated by longer non-rodent studies. Sponsors are advised by the FDA to get in contact with the agency when the maximum duration needs to be determined for non-rodents.

Table 14 summarizes the duration of repeat dose studies as agreed in ICH S4.

S4 focuses only on dog studies. There is no recommendation on what to do with primate studies. S6 offers some help for tests with biotechnology-derived drugs; the final decision is often only

Table 14 Duration of repeat dose toxicity studies

Duration of repeat dose toxicity studies
During clinical trials: Ratio 1:1 (4-week clinical trial requests 4-week toxicity studies)
For marketing authorization: Ratio 1:2/3 (4-week treatment of patients requires 3-month toxicity data)
Maximum duration of rodent repeat studies (6-month studies, followed by 2-year cancer test)
Maximum duration of non-rodent repeat studies (9-month tests in general)

possible by asking for a scientific advice from the agencies.

Further, a new species gains relevance for repeat dose assays, which is the **Goettingen minipig**, a result of breeding without genetic manipulation by the Department of Animal Sciences, Georg-August University Goettingen, Germany. There are many situations where the minipig is excellently predictive for human conditions, as examples cardiovascular or gastrointestinal issues. In the Rethink project (Forster et al. 2010), an impact assessment on minipigs as models for the toxicity of new medicines and chemicals was performed. In this context, Bode et al. (2010) discuss intensively the utility of the minipig as an animal model in regulatory toxicology. Further details were edited in 2012 by McAnulty, Dayan, Ganderup, and Hastings (CRC Press).

When new systems or methods are introduced into regulatory toxicological assessments, then a sequence of preclinical studies will be conducted to validate these new methods. The European Center for the Validation of Alternative Methods (ECVAM) has a long tradition to support the validation of such methods which reduce, refine, or replace the use of animals for safety testing and efficacy/potency testing of chemicals, biologicals, and vaccines. The introduction of new animal models is more difficult out of reasons of animal protection. Here the support from regulatory agencies by sharing their experience from often huge databases provides the stimulus. A. Jacobs and JW. van der Laan (2012) provide an overview on the regulatory acceptability of minipigs as a species of choice for the safety evaluation of human pharmaceuticals and medical devices and stress that this species can be a valuable alternative to dogs or nonhuman primates.

Reproductive Studies (ICH/S5 A+B)

ICH/S5A+B (21 and 22)

The special toxicology discipline "Reproductive and Developmental Toxicity" focuses on adverse effects on male and female fertility, birth defects (developmental toxicity, malformations, teratogenicity), and nonphysiological changes that appear shortly before, during, and after birth and during the weaning period. The relevant ICH guidelines describing the requirements for testing and detailing experimental studies are the ICH guideline S5A and S5B. These were among the first guidelines finalized in 1995, supporting the successful ICH process.

What was missing initially was the integration into the combinatorial test programs, the issues addressing male fertility, namely, for reason of protecting especially young male volunteers in clinical Phase I, when humans are being exposed for the first time to developmental medicinal compounds.

Finally, the decision was taken to incorporate the considerations of such male fertility testing into the main text of the guideline, which took place when the previous addendum, actually dated from 9 November, 2000, was incorporated into the common guideline S5 (R) in November 2005 by a successful revision process.

In humans, malformations and changes in development are relatively rare (around 6%) and often caused by accidental genetic errors; some are induced by external factors, e.g., chemical drugs. If a compound is labeled as a developmental toxicant, then the occurrence of structural or functional abnormalities in offspring is significantly increased at a dose level which does not induce severe maternal toxicity. If malformations are developed, then exposure levels of the individual fetuses are important, but also at which time of the development exposure was high. The inclusion of kinetic data helps to support the evaluation.

In general, there is a distinction between the following time periods (Figs. 6 and 7):

Segment I: male and female fertility studies

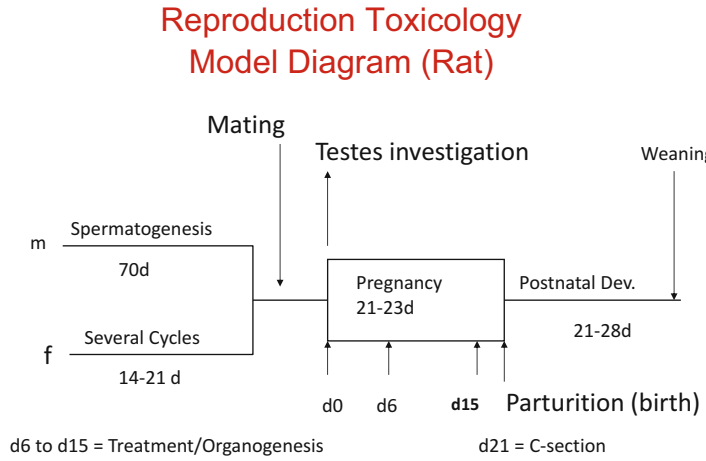


Fig. 6 Timing and design of reproductive toxicity studies

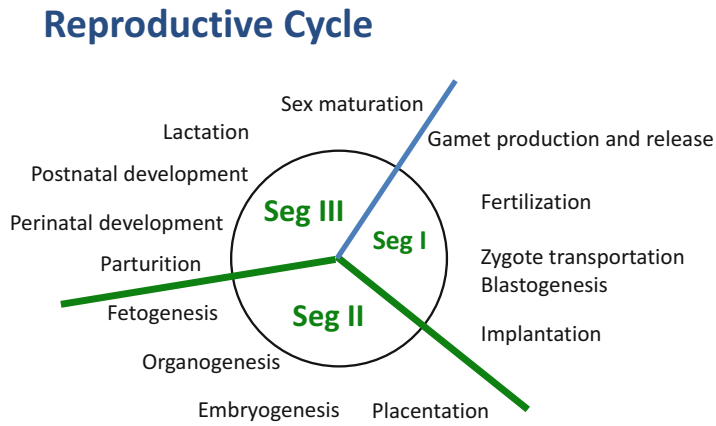


Fig. 7 Segment I, II, and III of reproductive cycles

Segment II: the embryo-fetal development (embryo during major organ development, i.e., organogenesis, the fetus in the post-embryonic period)

Segment III: the prenatal and postnatal development (neonate or postnatal offspring)

These segments can be tested separately or in a combined manner. All stages of development from conception to maturity and the detection of acute and delayed effects of exposure through one complete life cycle should be examined.

The standard species are rodents; rats as the preferred rodent species for all study types. The rabbit is the second non-rodent species for the embryo-fetal toxicity studies. The introduction of the rabbit as a sensitive and predictive model is due to the thalidomide (Contergan) catastrophe in the 1960s, when 10.000 of babies were born with a range of severe and debilitating malformations after their mothers had been exposed to this teratogen to treat their morning sickness (Vargesson 2015).

In some rare cases, mice, monkeys, or Goettingen minipigs (Bode et al. 2010) are used too, if special conditions – usually kinetic data – justify such species.

The guideline S5 (R) should be read in conjunction with the multidisciplinary ICH guideline M3, which was first introduced in 1997 and then revised several times until its current version from June 11, 2009. In this M3 (R), the different conditions are illustrated which have to be fulfilled before men, women of childbearing potential, pregnant women, or children are included into clinical trials.

It is interesting that men and women not of childbearing potential can be included into clinical trials without any experimental fertility studies because it is considered that the histopathological assessments of the gonads recovered from repeat dose toxicity studies provide sufficient insight into any possible disturbances of, e.g., the spermatogenesis. Our ICH collaborating Japanese colleagues contributed to this conclusion convincingly, literature examples are Kishi et al. (1995), Takayama et al. (1995), or Sakai et al. (2000).

In future, it is to expect that in vitro methods may facilitate early discovery of reproductive risks, whole body cultures, or stem cell investigations that help to validate these procedures. Another aspect will continue to expand scientific approaches but combined with political impact to give women of childbearing potential the same early access to new medicines.

ICH/S6: Preclinical Safety Evaluation of Biotechnology-Derived Products

ICH guideline S6 outlines the development of products which have been created by using biotechnology techniques. These compounds have been synthesized in living systems (cell cultures, virus, bacteria, and transgenic animals). These drugs are special also in regard to their size; in general they are >30 kDa, up to 800 kDa, which limits their distribution within the body or penetration into cells and organelles. This guidance is applicable for compounds such as recombinant DNA proteins, vaccines, peptides, plasma-derived products, endogenous proteins extracted from human tissues, oligonucleotide drugs, etc., while heparin, vitamins, and cellular blood components, for example, are not covered.

The production process of these compounds is complex; impurities may appear, e.g., host cell-derived products, which are difficult to qualify. Degradation or proteolysis may complicate their purification, to guarantee their stability and avoid aggregation; these are challenges for the formulation. All these characteristics explain why "the process is the product."

These bio-products are generally administered parenterally (subcutaneous, intramuscular, or intravenous application) and then submitted to catabolism and degradation leading to small peptides and nontoxic amino acids. This explains why studies of metabolism and excretion are generally not required.

An important characteristic is further their kinetics; their half-life is usually days to weeks with a limited distribution to plasma. When efficacy is long, then not surprisingly, there are long recovery periods to be expected.

In general, for biotech products high flexibility for the developmental scheme is recommended and should be reconsidered on a case-by-case basis.

Usually, as in other toxicity studies, two species should be used, but when the biological activity is well understood or when in short-term toxicity studies the effects were similar in both species, then longer-term studies could be run only with one species. In any case, a justification that tests are conducted with a relevant species (predictive for humans) is needed. This means using a species in which the test material is pharmacologically active due to the expression of the receptor or an epitope (in case of monoclonal antibodies).

When no relevant species can be identified, one should consider the use of homologous proteins or transgenic animals.

With regard to safety pharmacology, testing of the vital function such as cardiovascular, respiratory, and CNS functions is recommended, but S6 also mentions renal function, which, from today's perspective, would not be necessary before first administration of a compound in humans if there is no specific concern.

Some information on absorption, disposition, and clearance in the animal models are desirable before clinical trials; systemic exposure of the compound should be monitored as well as the appearance of antibodies and their ability to neutralize the intended effect. On the other hand, appearance of antibodies in animals when treated with human proteins is not predictive for clinical conditions.

Nevertheless, **immune** responses could alter the pharmacokinetic or pharmacodynamic effects. Anaphylactic responses tested in the guinea pig, at any rate, are not predictive for humans and therefore not necessary. The same holds true for the

standard testing batteries for immunotoxicity; these are not recommended.

Studies with duration between 2 weeks and 3 months are often sufficient, and it is only with chronic use in humans that 6-month studies should be considered.

Flexibility may allow further reduction of the traditional testing program for the **reproductive endpoints**. Such studies may not be necessary if a new compound, related to well-known compounds, shows similar effects. In addition, there are profound differences in maternofetal transfer of immunoglobulins between species with extensive gestational transfer of maternal immunoglobulins in primates (including humans) via the chorioallantoic placenta as well as in rabbits and guinea pigs via the inverted yolk sac splanchnopleure. In contrast, other neonatal rodents (rats and mice) receive passive immunity predominantly postnatally. This transfer is mediated principally via FcRn receptors.

Therapeutic monoclonal antibodies (mAbs) are most commonly of the IgG1 subclass, which is transported most efficiently to the fetus. In all animal species used for testing developmental toxicity, fetal exposure to IgG is very low during organogenesis, but this increases during the latter half of gestation (when there is only growth of the organs) such that the neonate is born with an IgG1 concentration similar to the mother (but not rats and mice). Review of mAb developmental toxicity studies of licensed products reveals that cynomolgus monkey is the species used in the majority of the cases (10 out of 15). Pregnancy outcome data from women gestationally exposed to mAb is limited. In general, the findings are consistent with the expected low exposure during organogenesis. Guinea pigs and rabbits are potential candidates as "alternatives" to the use of nonhuman primates as the maternofetal transfer in the last part of gestation is at a level similar in humans. Based on the pattern of placental transfer of IgG in humans, study designs that allow detection of both the indirect effects in early gestation plus the effects of direct fetal exposure in mid and late gestation are recommended for developmental toxicity of mAbs (Pentsuk and van der Laan, 2009).

The risk of **genotoxicity** also reveals an exception: The standard studies are not appropriate; DNA damage is not expected, but testing on a case-by-case basis of impurities or promoter studies might be helpful.

Long-term **carcinogenicity** studies are usually not appropriate, but when there is cause for concern, studies with a single rodent species are sufficient. Cause for concern may have arisen in general toxicity studies or when a stimulation of the growth of normal as well as malignant cells can be assumed.

For biotechnology-derived compounds, always a high flexibility has been asked for. The main principle was the case-by-case strategies, developers, and agencies contributing their knowledge to improve the predictive value for extrapolating to humans and patients. The guideline was first published in 1997, but the accumulation of collective observations and experience caused permanent discussions and justified revisions until the recent S6 R in June 2011.

Within the biotechnology-derived compounds, the monoclonal drugs play a major role. Here clinical experience has shown that more care for patients is needed especially when the first dose is administered in humans, mostly volunteers.

The use of the minimum anticipated biological effect level (MABEL) for selection of first human dose in clinical trials with monoclonal antibodies (Fig. 8) is offered. Muller et al. (2009) report the selection and strategies.

The authors plead that the dose selection for first-in-human (FIH) clinical trials with monoclonal antibodies (mAbs) is based on specifically designed preclinical pharmacology and toxicology studies, mechanistic ex vivo/in vitro investigations with human and animal cells, and pharmacokinetic/pharmacodynamic (PK/PD) modeling approaches and requires a thorough understanding of the biology of the target and the relative binding and pharmacological activity of the mAb in animals and humans. The authors highlight the scientific and regulatory challenges and the intensive care for enrolled humans when estimating the minimal anticipated biological effect level (MABEL).

The EMA supports these prudent strategies with their recommendations in the *Guideline on strategies to identify and mitigate risks for first-in-*

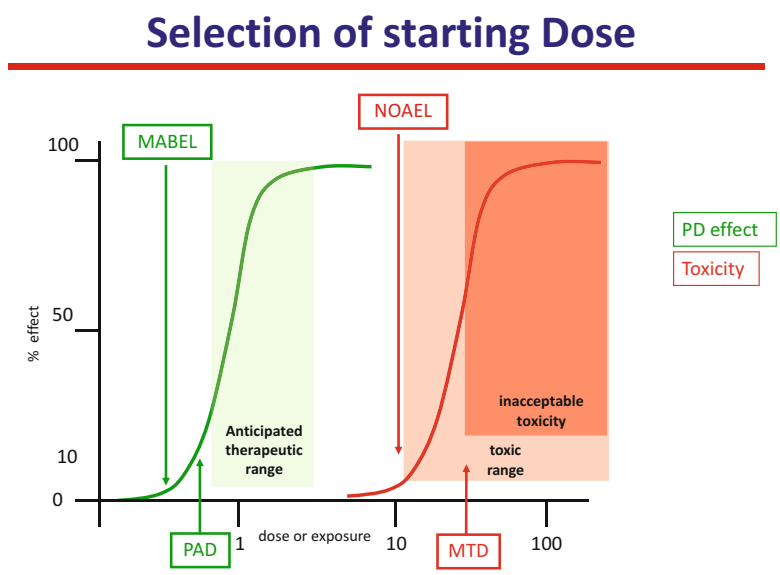


Fig. 8 Illustration of Mabel versus minimum toxic dose

human and early clinical trials with investigational medicinal products (February 2018).

Safety Pharmacology (ICH/S7 A+B)

S7A was implemented in 2001 and the S7B reached Step 4 in June 2005 and has been implemented in 2006.

S7A informs in general about the requirements necessary for testing the vital functions usually in single-dose studies in safety pharmacology.

S7A differentiates between three types of studies: core battery, follow-up, and supplemental studies. The core battery of tests/S7A consists of an investigation of the effects of a test substance on vital functions: central nervous system, cardiovascular system, respiratory system, and other systems as appropriate. The exclusion of a system or function should be justified.

Safety pharmacology studies carried out as necessary are:

- Follow-up studies for core battery (they provide a greater depth of understanding than, or additional knowledge to, that provided by the core battery [e.g., mechanistic studies])
- Supplemental studies: they evaluate effects of the test substance on systems not addressed by the core battery when there is cause for concern not addressed elsewhere (e.g., in toxicology).

Table 15 Timing of safety pharmacology studies

Before first administration to humans:
Core battery tests
Possibly follow-up/supplemental studies if concern
During clinical development:
Additional studies as required to clarify observed or suspected undesirable effects in animals or humans
Before approval
Effects on all organ systems, either covered by preclinical or clinical investigations
GLP conditions (S7A)
Core battery tests should be conducted according to GLP
Follow-up + supplemental tests according to GLP as far as possible
Primary and secondary pharmacodynamic studies need not be conducted according to GLP

S7A expresses very clearly when (Table 15) such studies should be available and what conditions should be considered in regard to good laboratory procedures.

Special focus is given to the cardiovascular system.

For the core battery of the cardiovascular system according to S7A blood pressure, heart rate, and electrocardiogram should be assessed,

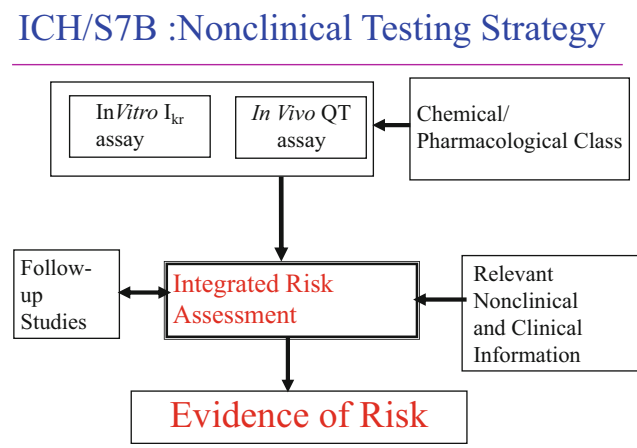


Fig. 9 Strategies of cardiovascular safety

but also in vivo, in vitro, and/or ex vivo evaluations, including methods for repolarization and conductance abnormalities, should be considered.

This text was finalized at a time when the details of S7B were not yet outlined. During recent years there has been an increase of regulatory concern. The awareness that non-cardioactive drugs (used for sometimes non-life-threatening diseases) can cause QT prolongation, and serious dysrhythmias such as torsades de pointes (TdP) was intensified.

A greater number of compounds became known to be associated with QT prolongation and the potential to cause torsades de pointes (severe arrhythmias).

Accordingly, the ICH Expert Working Group for S7B was created to work on this specific concern.

ICH S7B: Nonclinical Studies for Assessing Risk of Repolarization: Associated Ventricular Tachyarrhythmia for Human Pharmaceuticals

The background of S7B is summarized as follows:

- The QT interval (time from the beginning of the QRS complex to the end of the T wave) of the electrocardiogram (ECG) is a measure of the duration of ventricular depolarization and repolarization.
- QT interval prolongation can be congenital or acquired (e.g., pharmaceutical-induced).
- When the QT interval is prolonged, there is an increased risk of ventricular tachyarrhythmia, including torsade de pointes (TdP), particularly when combined with other risk factors (e.g., hypokalemia, structural heart disease, bradycardia), see Fig. 1.

For this important issue, a complex strategy was set up in the ICH guideline S7B (Fig. 9).

The basis for integrated risk assessment is in vitro and in vivo assays, supported by any knowledge about the chemical/pharmaceutical class. This first risk assessment may later be modified when results from follow-up studies or relevant nonclinical or clinical information becomes available. The evidence of risk summarizes the preclinical evaluation of the proarrhythmic potential as essential information for clinicians.

Parallel to the development of S7B, a clinical guideline (ICH/E14) was drafted and reached Step 4, also in 2005. During the discussion between these two expert groups, the question was raised again and again, if toxicologists could exclude any risk for QT prolongation for humans in their testing strategies. The answer is: they cannot. Preclinical researchers and clinical developers identify hazards and assess the definitive risk (the probability that an adverse effect may be induced), but a prediction for the individual patient is not possible. There is extrapolation from animal to man, from volunteers to patients under defined conditions, and from these to patients under daily living and health conditions, but all these data do not have the power to exclude

risks for future patient generations as can be seen by the fact that post-marketing authorization withdrawals can happen due to severe side effects.

The Food and Drug Administration (FDA) has during the ICH process compared preclinical data with clinical results and identified few cases where QT prolongation was observed under clinical conditions, while the preclinical tests were negative. This discrepancy is the basis for the diplomatic text in regard to the need for availability of S7B QT studies:

Timing of S7B Nonclinical Studies and Integrated Risk Assessment in Relation to Clinical Development/Step 4, June 2005, Brussels:

- Conduct of S7B nonclinical studies assessing the risk for delayed ventricular repolarization and QT interval prolongation prior to first administration in humans should be considered.
- These results, as part of an integrated risk assessment, can support the planning and interpretation of subsequent clinical studies.

The term "should be considered" allows flexibility either to do the studies before first time in humans or at a later stage of development. In practice, these studies are most often available before IND, because one wants to cope with this issue in time and wants to provide best safety to volunteers and patients.

But in conclusion it has to be stated that:

S7B proposes a series of nonclinical tests which are believed to predict the likelihood that a compound will prolong cardiac repolarization in vivo, in animals and in humans.

These data currently seem to have limited impact on the clinical development proposals contained in the draft E14 guideline now. The ICH E14 Expert Working Group was not convinced that the preclinical in vitro and in vivo study results would predict the clinical cardiovascular tolerance.

Under the auspices of the International Life Science Institute (ILSI)-Health and Environmental Sciences Institute (HESI), a consortium involving representatives from pharmaceutical companies, regulatory agencies, and opinion leaders from the scientific and medical research communities has been initiated (Trepakova et al. 2009).

The objectives are:

1. To assess the concordance between signals in nonclinical repolarization assays and clinical QT interval prolongation
2. To investigate the mechanisms for any discrepancy identified between nonclinical and clinical results and to determine viable and successful alternative approaches to identify these compounds
3. To assess the proarrhythmic potential of such compounds

At present, the consortium is conducting a retrospective analysis of nonclinical and clinical data from both FDA and contributing companies' databases and supplementing with a literature review.

The overall objectives of these efforts are to establish a quantitative integrated risk assessment for each compound, to define criteria for concordance, and to apply them to the database in order to identify non-concordant compounds.

Immunotoxicology Studies (ICH/S8)

Drug-induced immune dysfunction can induce increased susceptibility to infections, hypersensitivity reactions (immunological sensitization due to a drug and/or its metabolites), autoimmunity (immune reactions to self-antigens), or finally the development of tumors, but the present status of S8 is restricted to unintended immunosuppression and immunoenhancement, excluding allergenicity or drug-specific autoimmunity. The missing elements may be resolved in future revisions.

The guideline applies to new pharmaceuticals intended for use in humans, as well as to marketed drug products proposed for different indications or other variations on the current product label. The guideline does not apply to biotechnology-derived pharmaceutical products covered by ICH S6.

Immunosuppression or enhancement can be associated with two distinct groups:

Table 16 Immunotoxicity testing assay

Methods to evaluate immunotoxicity
I. Standard repeat dose toxicity tests first (signals for immunosuppression or stimulation)
If concern then
II. Additional immunotoxicity studies
T-cell-dependent antibody response (TDAR)
Immunophenotyping (lymphocyte subsets)
Natural killer cell activity assays
Host resistance studies
Macrophage/neutrophil function
Assays to measure cell-mediated immunity

1. Drugs intended to modulate immune function for therapeutic purposes (e.g., to prevent organ transplant rejection), here adverse immunosuppression can be considered as exaggerated pharmacodynamics.
2. Drugs not intended to affect immune function but cause immunotoxicity due to, e.g., necrosis or apoptosis of immune cells or interaction with cellular receptors shared by both target tissues and nontarget immune system cells.

Methods include (Table 16) standard toxicity studies (STS) and additional immunotoxicity studies conducted as appropriate. Whether additional immunotoxicity studies are necessary should be determined by a weight-of- evidence review of cause(s) for concern.

Illustrations of the thinking behind the guidances of the European proposals first as a forerunner of the final ICH S8 are publications like van der Laan et al. (1997) or in 2000 Putman et al.: Assessment of immunotoxic potential of human pharmaceuticals; Drug Info J 36: 417–427.

Table 16 illustrates the options and Fig. 10 the decision tree for the tests procedure.

The assessment of immunotoxicity should include the following:

- Statistical and biological significance of the changes
- Severity of the effects dose/exposure relationship
- Safety factor above the expected clinical dose
- Treatment duration
- Number of species and endpoints affected
- Changes that may occur secondarily to other factors (e.g., stress)
- Possible cellular targets and/or mechanism of action
- Doses which produce these changes in relation to doses which produce other toxicities and reversibility of effect(s).

Additional immunotoxicity testing should be considered (Table 17):

- If the pharmacological properties of a test compound indicate it has the potential to affect immune function (e.g., anti-inflammatory drugs).
- If the majority of the patient population for whom the drug is intended is immune-compromised by a disease state or concurrent therapy
- If a compound is structurally similar to compounds with known immunosuppressive properties
- If the compound and/or its metabolites are retained at high concentrations in cells of the immune system
- If clinical findings suggestive of immunotoxicity in patients exposed to the drug occur

More in detail: If the weight-of-evidence review indicates that additional immunotoxicity studies are needed, there are a number of assays which can be used (Fig. 10). It is recommended that an immune function study be conducted, such as a T-cell-dependent antibody response (TDAR). If specific cell types are affected in STS not involving cells participating in a TDAR, assays that measure function of that specific cell type might be conducted. Immunophenotyping of leukocyte populations, a nonfunctional assay, may be conducted to identify the specific cell populations affected and may provide useful clinical biomarkers.

Generally accepted are 28 consecutive daily doses in rodents. Adaptations of immunotoxicity assays have been described using non-rodent species. The species, strain, dose, duration, and route

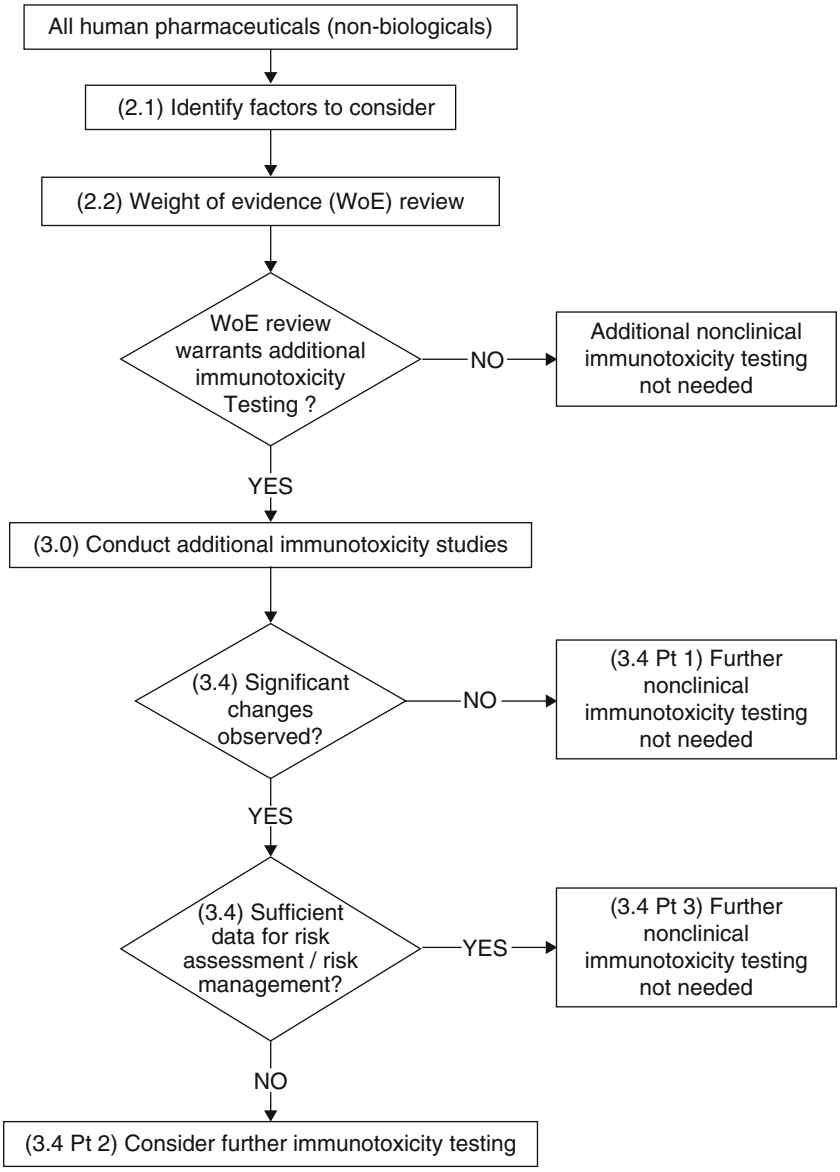


Fig. 10 **Decision tree for immunotoxicological evaluation**

of administration used in immune function assays should be consistent with the nonclinical toxicology study in which an adverse immune effect was observed.

Usually both sexes should be used in these studies, excluding nonhuman primates. The high dose should be above the no observed adverse effect level (NOAEL) but below a level inducing changes secondary to stress. Multiple dose levels are recommended in order to determine dose-response relationships and the dose at which no immunotoxicity is observed.

If no risk of immunotoxicity can be detected, then no further testing is needed.

If a risk of immunotoxicity is identified, additional testing may offer precision for the risk-benefit decision. Should the risk of immunotoxicity be considered as acceptable and/or can be addressed in a risk management plan (see ICH E2E), then no further testing in animals might be called for.

Table 17 Concern for immunotoxicological reactions

Causes for immunotoxicological concern
Findings from standard toxicity studies (STS)
Pharmacological properties for the intended patient population
Structural similarities to known immunomodulators
Disposition of the drug
Hematological changes, e.g., leukocytosis, granulocytopenia, or lymphopenia
Alterations in immune system organ weights and/or histology (e.g., changes in the thymus, spleen, lymph nodes, and/or bone marrow)
Changes in serum globulins (without plausible explanation)
Increased incidence of infections
Increased occurrence of tumors (possibly a sign of immunosuppression in absence of other plausible causes such as genotoxicity, hormonal effects, or liver enzyme induction

Any additional immunotoxicity studies should be completed before exposure of a large population of patients, usually Phase III.

If the target patient population is immunocompromised, immunotoxicity testing can be initiated at an earlier time point in the development of the drug.

The **TDAR** is administered to rats a week before the end of a 4-week repeat dose study recognized T-cell-dependent antigens (e.g., sheep red blood cells, SRBC or keyhole limpet hemocyanin, KLH), leading to a robust antibody response. Use antigens for immunization without adjuvants, (Alum acceptable for nonhuman primate studies).

Antibody measurement is done via ELISA or other immunoassay methods (samples can be collected serially during the study).

Immunophenotyping (lymphocyte subsets) differentiates T- and B-lymphocytes and subtypes. Their maturation and functions are in bone marrow for B-lymphocytes and NK cells and in thymus for T-lymphocytes with subsequent differentiation in cytotoxic (Tc) or T-helper (Th) cells. The cytotoxic T cells are cytotoxic to virus-infected cells; they form CD 8 positive cells as memory cells. The T-helper cells fortify defense mechanisms, secrete IL-2, IFN, and TNF, and form CD4 pos. cells.

Natural killer cells mature as lymphoid precursors in bone marrow. They show fast reactions during defense mechanisms and are therefore the first line of defense, e.g., for virus or tumors. They are cytotoxic when glycoproteins are recognized on cell membranes, but with identification of MHC-class 1, molecules on the surface of healthy cells show immediate activation of killing inhibitory receptors.

All NK cell assays are ex vivo assays in which tissues (e.g., spleen) or blood are obtained from animals that have been treated with the test compound. Cell preparations are coincubated with target cells that have been labeled with ^{51}Cr. New methods that involve nonradioactive labels can be used if adequately validated.

Host resistance studies involve challenging groups of mice or rats treated with different doses of test compound with varying concentrations of a pathogen (bacteria, fungal, viral, parasitic) or tumor cells. Infectivity of the pathogens or tumor burden observed in vehicle versus test compound treated animals is used to determine if the test compound is able to alter host resistance.

Host resistance assays involve innate immune mechanisms, for which specific immune function assays have not been developed. Careful observation of the direct or indirect (nonimmune mediated) effects of the test compound on the growth and pathogenicity of the organism or tumor cell is needed, e.g., compounds that inhibit the proliferation of certain tumor cells can increase host resistance.

Macrophages living for several months have a strong affinity to inflammatory locations and are powerful in phagocytosis, cytotoxicity, and secretion of mediators.

Localized macrophages are, e.g., alveolar macrophages, Kupffer cells in the liver, sessile macrophages in the spleen, and osteoclasts in bones. They phagocyte bacteria, fungus, parasites, damaged cells, old erythrocytes, immune complexes, they are cytotoxic to tumor cells, virus-infected cells, transplanted cells, and secrete enzymes, cytokines, prostaglandins, etc. Accordingly, their part in immunoregulation is the antigen production, cytokines, and unspecific immunosuppression.

Assays to measure cell-mediated immunity have not been as well established as those used for

the antibody response. They are in vivo assays where antigens are used for sensitization. The endpoint is the ability of drugs to modulate the response to challenge. Examples are delayed-type hypersensitivity (DTH) reactions with protein immunization and challenge reported for mice and rats.

This new **ICH guideline S8** replaces all guidances from the EU, USA, and Japan. It represents a very pragmatic approach and uses studies, e.g., standard toxicity studies, which are conducted anyhow. There is great confidence in the prediction of these assays for any potential of new compounds to induce immune suppression or immune stimulation. This guideline helps to reduce the number of animals and requires additional studies only in special cases for concern.

ICH S9: Nonclinical Evaluation of Anticancer Pharmaceuticals

This guideline was endorsed by ICH Steering Committee in May 2007. It had been observed that separate regional oncology guidances were in development.

All experts realized that malignant tumors are life-threatening, death rate from diseases is high, and existing therapies have limited effectiveness.

This explains the need to provide new effective anticancer pharmaceuticals to patients more expeditiously. The question was if this group of indications may allow higher flexibility in the design and timing of nonclinical studies for anticancer pharmaceuticals.

A working group was established delivering answers to questions like which nonclinical studies to support the development of anticancer pharmaceuticals in patients with *advanced disease and limited therapeutic options are really needed?* Does one sufficiently understand the toxicological profile of a pharmaceutical, e.g., identification of the target organs, exposure-response relationship, and reversibility in regard to the fact that with anticancer pharmaceuticals, the clinical dose level is often close to or at an adverse effect level?

How could one nevertheless protect patients from unnecessary adverse effects and at the same time avoid unnecessary use of animals in accordance with the 3 R principles (reduce/refine/replace)?

As all indications, the discipline of **pharmacology/pharmacodynamics** needs to deliver the nonclinical proof of principle, what is the preliminary characterization, what are the underlying mechanism(s) of action, how does the drug exert its antitumor activity, which schedule dependence is needed, and, if necessary, which combinations can be justified?

The discipline of **safety pharmacology** is requested to identify any acute risks of the vital organ functions. Are the cardiovascular, respiratory, and central nervous systems not disturbed? This answer must be given before the initiation of clinical studies and based either on stand-alone safety pharmacology studies or study part included into general toxicology studies. How could specific concerns be managed? In the absence of a specific risk, such studies will not be called for to support clinical trials or for marketing.

Kinetics: limited pharmacokinetic parameters (e.g., Cmax in plasma/serum, AUC, and half-life) in the animal species used for nonclinical studies can facilitate dose selection, schedule, and escalation during Phase I studies. Further information (ADME) can be generated in parallel with clinical studies.

General toxicity: the identification of a NOEL/NOAEL is not essential. Small molecules should be usually conducted in both rodent and non-rodent, with the exception for genotoxic drugs targeting rapidly dividing cells; here one rodent species might be considered sufficient. For biopharmaceuticals, see ICH S6. Important is the issue of reversibility: severe toxicity at approximate clinical exposure and recovery cannot be predicted by scientific assessment, but the demonstration of complete recovery is not considered essential.

Reproductive toxicity: Embryo-fetal toxicity studies of anticancer pharmaceuticals should be available for marketing with the exception if they target rapidly dividing cells in general toxicity studies or belong to a class which has been well characterized in causing developmental toxicity.

If a pharmaceutical is positive for embryo-fetal lethality or is teratogenic, a confirmatory study in a second species is usually not warranted.

For biopharmaceuticals, the assessment might be done by evaluating the toxicity during the period of organogenesis or study designs as described by ICH S6. Alternative approaches might be considered appropriate if scientifically justified. The alternative approaches might include a literature assessment, assessment of placental transfer, the direct or indirect effects of the biopharmaceutical, or other factors.

A fertility study is generally not warranted to support the treatment of patients with advanced cancer. Information available from general toxicology studies on reproductive organs should be incorporated into the assessment of reproductive toxicology.

A peri- and postnatal toxicology study is generally not warranted to support the treatment of patients with advance cancer.

Genotoxicity: Studies should be performed to support marketing. The principles outlined in ICH S6 should be followed for biopharmaceuticals. If the in vitro assays are positive, an in vivo assay might not be warranted.

Carcinogenicity: Studies are usually not warranted based on the underlying disease.

Immunotoxicology: General toxicology studies are considered sufficient to evaluate the immunotoxicological potential.

Photosafety studies: Initial assessment of phototoxic potential should be conducted prior to Phase I. Photochemical properties of the drug and information on other members in the class should be assessed. If assessment of these data indicates a potential risk, appropriate protective measures should be taken during outpatient trials.

If the photosafety risk cannot be adequately evaluated based on nonclinical data or clinical experience, a photosafety assessment consistent with the principles described in ICH M3 should be provided prior to marketing.

Setting the starting dose: Pharmacologically active dose should be reasonably safe to use.

Start dose should be scientifically justified using all available nonclinical data (e.g., pharmacokinetics, pharmacodynamics, toxicity). Interspecies scaling usually based on normalization to body surface area. Interspecies scaling based on body weight, AUC, or other exposure parameters might be appropriate.

The FDA is offering good advice and help for calculations in their Guidance for Industry: *Estimating the maximum safe starting dose in initial clinical trials for Therapeutics in adult healthy volunteers* (CDER July 2005). Here conversions on animal doses to human equivalent doses based on body surface area are tabled.

For biopharmaceuticals with agonistic properties, minimally anticipated biologic effect level (MABEL) should be considered

During clinical trials: Highest clinical dose is not limited by nonclinical data. In Phase I, treatment dose depends on the patient's response. No new toxicology study is called for to support continued treatment beyond duration of the toxicology studies. Examples of duration and schedule of toxicology studies to support initial clinical trials are provided.

In cases where the available toxicology information does not support a change in clinical schedules, an additional toxicology study in a single species is usually sufficient. The highest clinical dose is not limited by nonclinical data.

Beyond Phase I, the pharmaceutical development could continue with the delivery of results from repeat dose studies of 3 months prior the initiation of phase III studies. This maximal duration would for most pharmaceuticals be considered sufficient to support marketing.

A special option for this indication is offered by a Phase 0 clinical trial:

This is a first-in-human clinical trial conducted under an exploratory IND that has no therapeutic or diagnostic intent and involves very limited human exposure.

The results of such a Phase 0 trial can provide essential pharmacodynamic, pharmacokinetic, and/or imaging data at the initial stage of the clinical trials process to inform and expedite the subsequent development of promising new agents.

The results of the first "Phase 0" clinical trial in oncology of a therapeutic agent under the Exploratory Investigational New Drug Guidance of the US FDA have recently been reported and considered to be a successful and expeditious new paradigm for early therapeutics development in oncology, but according to Kummar et al. (2009), several additional phase 0 trials will need to be completed under the Exploratory IND Guidance, before phase 0 trials will be considered to have an established role in the anticancer drug development process. Nevertheless, the US Food and Drug Administration's new regulatory policy has provided an important and timely opportunity to expeditiously conduct and complete novel, proof-of-principle clinical trials of molecularly targeted therapeutic and imaging agents. The potential for a major impact of phase 0 trials and the exploratory IND on developing new anticancer drugs provides a strong stimulus for the broader uptake and enhanced application of carefully conceived, pharmacodynamically driven early-phase clinical trials in oncology.

In **Summary** (Table 18), one can conclude that this new guideline ICH S9 facilitates and speeds up the development of new oncology compounds, since there is no need for 6-/9-month studies, no need for fertility and peri- and postnatal studies, only one embryo-fetal study if a positive reaction is observed, no in vivo micronucleus test if in vitro genotoxicity assays are positive, safety pharmacology assessments that could be conducted within the general toxicology studies, no need for non-rodent studies for initiation of clinical trials with cytotoxic drugs, and no or limited studies to be conducted in late-stage development. All restrictions help to conserve resources and reduce animal use.

Table 18 Facilitation of development of oncologics

Development of oncologic compounds		
1.	Pharmacology	Primary PD studies (in vivo and/or in vitro)
2.	Safety pharmacology	Safety pharmacology core battery studies
3.	Kinetics	Kinetics: AUC, Cmax, ADME
4.	Human metabolites	Nonclinical test of human metabolite(s)
5.	Toxicity 4 weeks	Repeat dose toxicity studies 4 weeks
6.	Toxicity 3 months	Repeat dose toxicity studies 3 months
7.	Genotoxicity	Genotoxicity
8.	Teratogenicity	Embryo-fetal studies
9.	Phototoxicity	Phototoxicity

ICH Guideline S10

The ICH S10 guideline on photosafety evaluation reached Step 4 of the ICH Process in November 2013 and has entered the implementation period (Step 5) between 2014 and 2016 in the different areas. Forerunners of this global guideline were the EMA guideline (European Agency for the Evaluation of Medicinal Products), note for guidance on photosafety testing. Committee for Proprietary Medicinal Products from June 27, 2002. CPMP /SWP/ 398 /01 and the Guidance on photosafety testing, from May 2003. US Department of Health and Human Services Food and Drug Administration (CDER) (Fig. 11).

The multidisciplinary guideline ICH M3(R2) recommends that an initial assessment of phototoxicity potential be conducted before exposure of large numbers of subjects (Phase 3). This safety evaluation is especially important for compounds with known phototoxic risk, like antibiotics or oncology products. However, neither ICH M3(R2) nor ICH S9 provide specific information regarding testing strategies. Here ICH S10 guideline outlines when photosafety testing is warranted and which possible assessment strategies can be recommended. This guideline should hereby reduce the likelihood that substantial differences in recommendations for photosafety assessment will exist among regions.

S10 generally applies to new active pharmaceutical ingredients (APIs), new excipients, clinical formulations for dermal application (including dermal patches), and photodynamic therapy products but excludes peptides, proteins, antibody drug conjugates, or oligonucleotides.

The photosafety assessment of a drug considers for the evaluation the photochemical

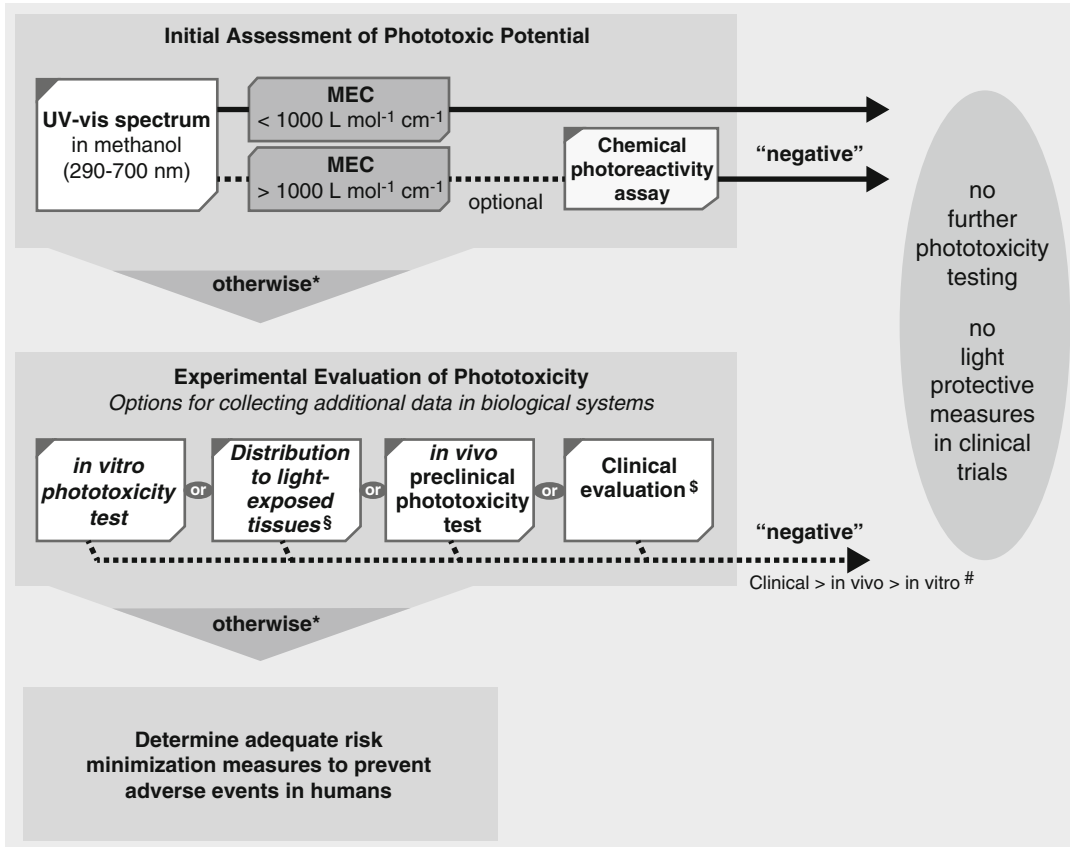


Fig. 11 Assessment of phototoxic evaluation

characteristics, results from nonclinical studies, and human safety information. Four different effects have been discussed in connection with photosafety testing: phototoxicity, photoallergy, photogenotoxicity, and photocarcinogenicity. But testing for photogenotoxicity (Note 2) and photocarcinogenicity (Note 6 of ICH M3 (R2)) is not currently considered useful for human pharmaceuticals. Accordingly, the focus is on phototoxicity and photoallergy effects as defined below:

- Phototoxicity (photoirritation): An acute light-induced tissue response to a photoreactive chemical
- Photoallergy: An immunologically mediated reaction to a chemical, initiated by the formation of photoproducts (e.g., protein adducts) following a photochemical reaction

Photosensitization is a general term occasionally used to describe all light-induced tissue reactions. However, in order to clearly distinguish between photoallergy and phototoxicity, the term photosensitization is not used in this guideline.

For a chemical to demonstrate phototoxicity and/or photoallergy, the following characteristics are critical:

- Absorbs light within the range of natural sunlight (290–700 nm).
- Generates a reactive species following absorption of UV-visible light.
- Distributes sufficiently to light-exposed tissues (e.g., skin, eye).

If one or more of these conditions is not met, a compound will usually not present a concern for direct phototoxicity.

Photochemical properties: The first question when assessing the photoreactive potential is whether a compound absorbs photons at any wavelength between 290 and 700 nm. A compound that does not have a molar extinction coefficient (MEC) greater than 1000 L mol-1 cm-1 at any wavelength between 290 and 700 nm (Bauer et al. 2014) is not considered to be sufficiently photoreactive and therefore able to induce direct phototoxicity. Excitation of molecules by light can lead to generate reactive oxygen species (ROS), including superoxide anion and singlet oxygen via energy transfer mechanisms. Thus, ROS generation following irradiation with UV-visible light can be an indicator of phototoxicity potential.

Photostability testing (EMA/CHMP/ICH/536328/2013 2016a) can also suggest the potential for photoreactivity, but photostability testing alone should not be used to determine whether further photosafety evaluation is warranted.

Tissue distribution: The concentration of a photoreactive chemical in tissue at the time of light exposure is a very important pharmacokinetic parameter in determining whether a phototoxic reaction will occur. This concentration depends on a variety of factors, such as plasma concentration, perfusion of the tissue, partitioning from vascular to interstitial and cellular compartments, and binding, retention, and accumulation of the chemical in the tissue. The duration of exposure depends upon clearance rates as reflected by half lives in plasma and tissue. Further, the longer the concentration of a compound is maintained at a level above that critical for a photochemical reaction, the longer a person is at risk for phototoxicity.

Although a tissue concentration threshold below which the risk for phototoxic reactions would be negligible is scientifically plausible, there are currently no data to delineate such generic thresholds for all compounds. Risks are considered low when a drug shows an overall very low exposure or has a very short plasma half-life or tissue residence. On the other hand, compound binding to tissue components (e.g., melanin, keratin) can explain tissue retention and/or accumulation, but such binding alone does not present a photosafety concern. A single-dose tissue distribution study, with animals assessed at multiple time points after dosing, will generally provide an adequate assessment of relative tissue to plasma concentration ratios, tissue residence time, and the potential for retention and accumulation.

Nonclinical photosafety tests: Generally, it is most important that nonclinical photosafety assays show high sensitivity resulting in a low frequency of false negatives (i.e., a high negative predictive value). This is because negative assay results usually do not warrant further photosafety evaluation.

Natural sunlight represents the broadest range of light exposure that humans might be exposed to regularly. However, sunlight per se is not well defined. Suitability of a sunlight simulator light source should be well defined.

In nonclinical phototoxicity assays, however, the amount of UVB should not limit the overall irradiation and might be attenuated (partially filtered) so that relevant UVA doses can be tested without reducing assay sensitivity. Penetration of UVB light into human skin is mainly limited to the epidermis, while UVA can reach capillary blood. Therefore, clinical relevance of photochemical activation by UVB is considered less important than activation by UVA for systemic drugs. However, UVB irradiation is relevant for topical formulations applied to light-exposed tissues.

S10 discusses pros and cons of a number of different test approaches:

- Photoreactivity tests using chemical assays
  High sensitivity for predicting direct in vivo phototoxicants, but low specificity, generating a high percentage of false-positive result.
- Phototoxicity tests using in vitro assays
  The most widely used in vitro assay for photo toxicity is the three T3 neutral red uptake phototoxicity test (3 T3 NRU-PT).
- Photosafety tests using in vivo assays and systemic administration
  Phototoxicity testing for systemically administered compounds has been conducted in a variety of species, including guinea pig, mouse, and rat. No standardized study design has been established.

For 3T3 NRU-PT, the Organisation for Economic Co-operation and Development (OECD) guideline (EMA/CHMP/ICH/536328/2013 2016b) is available. This is currently considered the most appropriate in vitro screen for soluble compounds.

The sensitivity of the 3T3 NRU-PT is high, and if a compound is negative in this assay, it would have a very low probability of being phototoxic in humans. However, a positive result in the 3T3 NRU-PT should not be regarded as indicative of a likely clinical phototoxic risk, but rather a flag for follow-up assessment.

There are no in vitro models that specifically assess ocular phototoxicity, regardless of the route of administration. While negative results in the 3T3 NRU-PT or a reconstructed human skin assay might suggest a low risk, the predictive value of these assays for ocular phototoxicity is unknown.

The most sensitive early signs of compound-induced phototoxicity are usually erythema followed by edema at a normally sub-erythemogenic irradiation dose. The type of response might vary with the compound. Any identified phototoxicity reaction should be evaluated regarding dose and time dependency, and, if possible, the no-observed-adverse-effect level (NOAEL) should be established. The hazard identification might be further supported by additional endpoints (e.g., early inflammatory markers in skin or lymph node reactions indicative of acute irritation).

If a phototoxicity study is conducted in animals for a systemic drug that absorbs light above 400 nm, phototoxicity of the retina should be assessed using a detailed histopathological evaluation. For compounds that only absorb light below 400 nm, retinal assessment is usually not warranted because such wavelengths do not reach the retina of the adult human eye due to limited penetration of the cornea, lens, and vitreous body.

Testing for photoallergy is not recommended for compounds that are administered systemically. Photoallergy reactions in humans following systemic administration are rare, and there are no established nonclinical photoallergy assays for systemically administered compounds.

Photosafety Tests Using In Vivo Assays and Dermal Administration

For dermal drug products in general, the clinical formulation should be tested. The intended clinical conditions of administration should be used to the extent possible. Irradiation of the exposed area should take place at a specified time after application, and the interval between application and irradiation should be justified based on the specific properties of the formulation to be tested.

For dermal drug products, contact photoallergy has often been assessed in a nonclinical study along with acute phototoxicity (photoirritation). However, no formal validation of such assays has been performed. While the acute photoirritation observed in these studies is considered relevant to humans, the predictivity of these studies for human photoallergy is unknown. For regulatory purposes, such nonclinical photoallergy testing is generally not recommended.

Assessment strategies: The choice of the photosafety assessment strategy is up to the drug developer. ICH M3(R2) suggests that an initial assessment of the phototoxicity potential based on photosafety evaluation of pharmaceuticals, photochemical properties, and pharmacological/chemical class be undertaken before outpatient studies.

Characterization of the UV-visible absorption spectrum is recommended as the initial assessment because it can obviate any further photosafety evaluation.

In addition, the distribution to the skin and eye can be evaluated to inform further on the human risk and the recommendations for further testing. Then, if appropriate, an experimental evaluation of phototoxicity potential (in vitro or in vivo, or clinical) should be undertaken before exposure of large numbers of subjects (Phase 3). Figure 11 provides an outline of possible phototoxicity assessment strategies. The figure is based on the strategies outlined in this section of this document. The strategies are flexible. Depending on the particular situation, some portions of the assessment are optional and might not be conducted.

Recommendations for Pharmaceuticals Given via Systemic Routes

Assessment of Phototoxicity Potential

If the substance does not have a MEC greater than 1000 L mol-1 cm-1 (between 290 and 700 nm), no photosafety testing is recommended and no direct phototoxicity is anticipated in humans.

For compounds with MEC values of 1000 L mol-1 cm-1 or higher, if the drug developer chooses to conduct a test for photoreactivity, a negative result could support a decision that no further photosafety assessment is warranted.

Experimental Evaluation of Phototoxicity

The in vitro approach with the 3T3 NRU-PT is currently the most widely used assay and in many cases could be considered as an initial test for phototoxicity. The high sensitivity of the 3T3 NRU-PT results in good negative predictivity, and negative results are generally accepted as sufficient evidence that a substance is not phototoxic. In such cases no further testing is recommended and no direct phototoxicity is anticipated in humans.

If an in vitro phototoxicity assay gives a positive result, a phototoxicity study in animals could be conducted to assess whether the potential phototoxicity identified in vitro correlates with a response in vivo. Alternatively, drug distribution data could, on a case-by-case basis, support a position that the risk of phototoxicity in vivo is very low and that no further photosafety assessment is warranted. As another option, the photosafety risk could be assessed in the clinical setting or managed by the use of light-protective measures. A negative result in an appropriately conducted phototoxicity study either in animals or humans supersedes a positive in vitro result. In such cases no further testing is recommended and no direct phototoxicity is anticipated in humans.

In all cases a robust clinical phototoxicity assessment indicating no concern supersedes any positive nonclinical results.

Recommendations for Pharmaceuticals Given via Dermal Routes

Assessment of Phototoxicity Potential

If the active substance and excipients do not have MEC values greater than 1000 L mol-1 cm-1 (between 290 and 700 nm), no further photosafety testing is recommended and no phototoxicity is anticipated in humans.

Molar extinction coefficient (MEC) (also called molar absorptivity) reflects the efficiency with which a molecule can absorb a photon at a particular wavelength (typically expressed as L mol-1 cm-1) and is influenced by several factors, such as solvent.

For compounds with MEC values of 1000 L mol-1 cm-1 or higher, negative photoreactivity test results (e.g., a ROS assay) (reactive oxygen species, including superoxide anion and singlet oxygen) can support a decision that no further photosafety assessment is warranted.

Experimental Evaluation of Phototoxicity and Photoallergy

The 3T3 NRU-PT can be used to assess individually the phototoxicity potential of the API and any new excipient(s), provided that appropriate testing conditions can be achieved (e.g., test concentrations not limited by poor solubility, relevant UVB dose can be applied).

In cases where no phototoxic component has been identified in vitro, the overall phototoxicity potential of the clinical formulation can be regarded as low.

Reconstructed human skin models can be used to assess the phototoxicity potential of clinical formulations. Under adequate test conditions, a negative result in a reconstructed human skin assay indicates that the direct phototoxicity potential of the formulation can be regarded as low. In this case, generally no further phototoxicity testing is recommended.

As the predictivity of nonclinical photoallergy tests is unknown, this would typically be a clinical assessment using the to-be-marketed formulation and conducted during Phase 3.

Photosafety evaluation of the clinical formulation delivered via dermal patches can follow the above-described principles for clinical dermal formulations. For transdermal patches, the principles for both dermal and systemic drugs should be applied. In addition, the intended clinical use (e.g., skin area recommended for use, duration of application) and the properties of the patch matrix (e.g., being opaque to UV and visible light) should be considered for the overall risk assessment.

Testing for **photogenotoxicity** is not recommended as a part of the standard photosafety testing program. These tests are substantially oversensitive and even incidences of pseudo-photoclastogenicity have been reported (Lynch et al. 2008).

Furthermore, the interpretation of photogenotoxicity data regarding its meaning for clinically relevant enhancement of UV-mediated skin cancer is unclear.

A survey of pharmaceutical companies indicated that the **3T3 NRU-PT**, as described in Organisation for Economic Co-operation and Development, Test Guideline (OECD TG) 432, generates a high percentage of positive results (approximately 50%), the majority of which do not correlate with phototoxicity responses in animals or humans (Lynch and Wilcox 2011).

In the USA, for products applied dermally, a dedicated clinical trial for phototoxicity (photoirritation) on the to-be-marketed formulation (API plus all excipients) can be warranted in support of product approval.

S11 Nonclinical Pediatric Safety

S11 Nonclinical Safety Testing in Support of Development of Pediatric Medicines

This topic was endorsed by the ICH Steering Committee in November 2014 and is in 2018 still in ICH Step 1 (the consensus building process of ICH).

The S11 guideline will recommend standards for the conditions under which nonclinical juvenile animal testing is considered informative and necessary to support pediatric clinical trials.

Clarification is necessary in regard to the need, study design, and/or timing of juvenile animal studies to support pediatric indications, since considerable differences exist within in the present regulatory guidelines from the ICH, EMA, FDA, and MHLW. The regional guidelines recommend a case-by-case approach for determining the need for a juvenile animal study after consideration of the available data (FDA 2006; EMA 2008; MHLW, 2012).

The guideline may propose a modification of a repeat dose general toxicity study or a pre- and postnatal developmental toxicity study.

The ICH M3(R2) guideline, which focuses on the need and, when warranted, the timing of juvenile animal studies, states, "The conduct of any juvenile animal toxicity studies should be considered only when previous animal data and human safety data, including effects from other drugs of the pharmacological class, are judged to be insufficient to support pediatric studies."

If a study is warranted, one relevant species, preferably rodent, is generally considered adequate. The expectation is that the inconsistencies in interpretation and application of the present different guidelines can be harmonized.

Multidisciplinary Guidelines

These ICH recommendations or guidelines are the crosscutting topics; they do not fit uniquely into one of the quality, safety, and efficacy categories but usually request cooperations between several scientific disciplines. M3 as an example represents the contribution of safety and efficacy. Or M4 refers to all three categories. Additionally multidisciplinary guidelines includes the ICH medical terminology (MedDRA) and the development of Electronic Standards for the Transfer of Regulatory Information (ESTRI). The following list summons up all types of these guidelines; in detail M3 and M4 are handled.

M1 (Medical Dictionary for Regulatory Activities) (Concept Paper Available, Step 1)

The development of a Medical Dictionary for Regulatory Activities was approved by the ICH Steering Committee in 1997 and the terminology launched in 1999. Further development of this topic followed over the years, all information about MedDRA and the points to consider documents developed for every MedDRA version are available on the MedDRA page under the work products. The concept paper was approved in 2016.

M2: Electronic Standards for the Transfer of Regulatory Information (ESTRI) (Concept Paper Available, Step 1)

This Expert Working Group (EWG) was established by the ICH Steering Committee in 1994. The objective was to facilitate international electronic communication by evaluating and recommending, open and nonproprietary – to the extent possible – Electronic Standards for the Transfer of Regulatory Information (ESTRI) that will meet the requirements of the pharmaceutical companies and regulatory authorities.

The ICH Steering Committee modified in 2010 the mandate of the M2 EWG. Important changes included agreement that work related to the Electronic Common Technical Document (eCTD) be undertaken by a newly established M8 EWG and that the M2 EWG would no longer be directly involved in the development of technical solutions in relation to topics such as E2B(R3) and M5 but would instead provide the framework for the efficient and effective development of the solutions by groups dedicated to these topics. Under its new mandate, the M2 EWG continues to be responsible for the evaluation and recommendation of standards.

M5 Data Elements and Standards for Drug Dictionaries (See E2 B)

In 2003, the ICH Steering Committee endorsed a concept paper for topic M5, and EWG develops since the requirements for the standardization of medicinal product identifiers and related terminology. In particular, a need was identified to harmonize product information that would facilitate the electronic exchange of Individual Case Safety Reports (ICSRs) within and across ICH regions using the ICH E2B format in post-marketing pharmacovigilance.

In 2005, an M5 consensus draft guideline was published for public consultation at Step 2 of the ICH process. To support the electronic exchange of the M5 data elements proposed by the M5 EWG, technical messaging specifications were needed. Electronic messaging development for utilization by the ICH Parties had been the domain of the ICH M2 EWG, but in June 2006 the ICH Steering Committee took a key decision to develop electronic specifications in collaboration with Standards Development Organizations (SDOs). This would enable wider interoperability across regulatory and healthcare communities.

M6 Virus and Gene Therapy Vector Shedding and Transmission (Step 1)

The ICH Steering Committee endorsed this topic in 2009.

Work by the Gene Therapy Discussion Group (GTDG) followed, a document "General Principles to Address Virus and Vector Shedding" became available. It was recognized that more extensive information to improve harmonization among the ICH regions was needed. In 2011 this topic was ceased following Steering Committee discussion that concluded due to the current state of science and related resource allocation would not allow this to be supported as a topic for harmonization.

M7(R1)Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk (Step 5)

This guideline was finalized as a step 5 document in May 2017; it offers guidance on analysis of structure activity relationships (SAR) for genotoxicity. Furthermore, it is intended to resolve questions such as whether impurities with similar alerts that potentially have similar mechanism of action should not be combined in calculating a Threshold of Toxicological Concern (TTC) and whether the TTC may differ based on differences in the approved duration of use.

To complement this ICH M7 guideline, an addendum was finalized in 2017 to summarize known mutagenic impurities commonly found or used in drug synthesis. The intent of this addendum is to provide useful information regarding the acceptable limits of known mutagenic impurities/carcinogenic and supporting monographs.

In a maintenance process, work followed to define acceptable limits (acceptable intakes (AIs) or permitted daily exposures (PDEs)) for new DNA reactive (mutagenic) impurities and revising acceptable limits for impurities already listed in the addendum as new data becomes available. Data and/or proposals pertaining to the revision of the ICH M7(R1) guideline with supporting information can be submitted directly to the ICH Secretariat from either an ICH Member or Observer or other interested ICH stakeholders.

M7(R2) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk (Step 1)

The M7(R2) EWG is currently undertaking a maintenance of the guideline, which will result in the future ICH M7(R2) version. The work is now in Step 1.

M8 Electronic Common Technical Document (eCTD) (Step 4)

ICH M8 EWG was formed in 2010 to take over the development and revision of eCTD v 4.0 Implementation Guide and related documents from the ICH M2 eCTD Subgroup. Discussions and revisions took place and updated versions are expected.

The M8 EWG also provides technical review and impact assessment of issues arising from the use of the ICH M4 CTD guidelines within the context of the eCTD.

M9 Biopharmaceutics Classification System-Based Biowaivers (Step 1)

This topic was endorsed by the ICH Management Committee in October 2016.

This new multidisciplinary guideline is proposed to address Biopharmaceutics Classification System (BCS)-based biowaivers. BCS-based biowaivers may be applicable to BCS Class I and III drugs; however, BCS-based biowaivers for these two classes are not recognized worldwide. This means that pharmaceutical companies have to follow different approaches in the different regions. This guideline will provide recommendations to support the biopharmaceutics classification of medicinal products and will provide recommendations to support the waiver of bioequivalence studies. This will result in the harmonization of current regional guidelines/guidance and support streamlined global drug development.

M10 Bioanalytical Method Validation (Step 1)

This topic was endorsed by the ICH Management Committee in October 2016.

This new multidisciplinary guideline will apply to the validation of bioanalytical methods and study sample analyses in nonclinical and clinical studies. Reliable data derived through validated bioanalytical methods are key for the review of marketing authorization application.

This guideline will provide recommendations on the scientific regulatory requirements for bioanalysis conducted during the development of drugs of both chemical and biological origins. It will also address issues on method validation by considering the characteristics of the analytical methods used in bioanalysis, e.g., chromatographic assay and ligand binding assay. A harmonized bioanalytical method validation guideline will promote the prompt, rational, and effective nonclinical and clinical studies, thereby advancing the mission of the ICH.

ICH Multidisciplinary Guidelines M3 (Timing)

This guideline was in Step 5 in 1997 and has been revised now several times (Table 19).

The newest version is called:

Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2)

Current Step 4 version dated June 11, 2009.

This guideline focuses on duration and timing of preclinical studies but also very importantly under which conditions different patient populations (men, women, children) can be included into clinical investigations. The guideline stresses that drug development is a stepwise process. It starts with short-term studies, and during the following 10 years and more, the longer time tests follow (Fig. 12).

ICH/M3 provides information about which studies – and of what duration – are needed before the different clinical phases of development can be started.

Special attention is realized during the transition from preclinical animal testing and the first administration of a new drug to humans, mostly male volunteers.

Table 19 General principles of ICH guideline M3 (timing)

M3 R: general principles
Development of a pharmaceutical is a stepwise process
Evaluation of animal and human efficacy and safety needed
The goals of the nonclinical safety evaluation include:
Characterization of toxic effects
Identification of target organs
Clarification of dose dependency
Relationship of toxicities to exposure
Potential reversibility
Possibly: mechanism of toxicity

Table 20 summarizes the different endpoints needed to be available and acceptable before this FIM. And Table 21 pays attention to the important data analysis of metabolites.

In addition to this timing-schedule of preclinical studies in dependence of clinical development plans, there are recommendations under which conditions different populations can be included into clinical trials, populations such as men, women of childbearing potential, or pregnant women and finally pediatric populations.

The objectives of M3 are to reduce differences between regions, to facilitate timely conduct of clinical trials, to reduce unnecessary use of animals and other resources, and to promote early availability of new drugs.

The background of this guidance was that regulatory recommendations differed among regions of Europe, the USA, and Japan. Would it be possible to develop a mutually acceptable guidance?

The guideline provides general guidance for drug development. A guideline is not a legal requirement. Approaches should be scientifically and ethically appropriate.

The different endpoints are addressed in details in this note for Guidance. Table 20 summarizes all study requirements before the first starting dose in humans.

The core battery of safety pharmacology is recommended to be conducted prior to first administration in humans. Any follow-up or

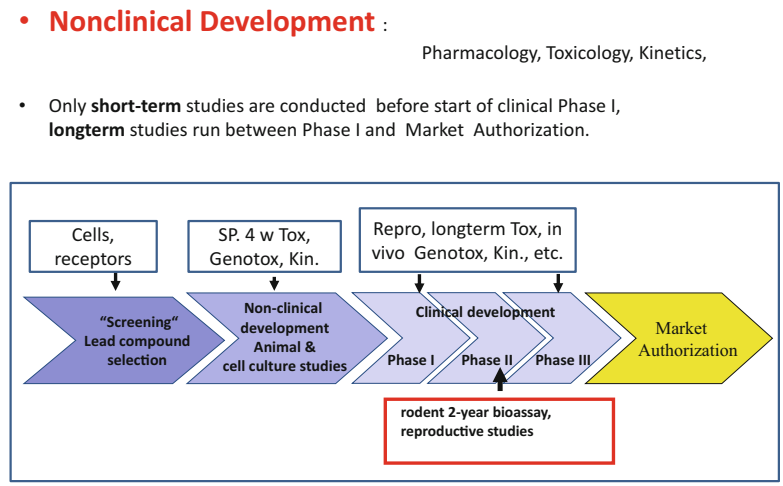


Fig. 12 Schematic design of preclinical development

Table 20 Preclinical studies needed **before** first dose in man

Preclinical studies before FIM	
1.	Primary PD studies (in vivo and/or in vitro)
2.	Safety pharmacology core battery studies
	(CV, CNS, respiration)
3.	Kinetics: In vitro metabolism, plasma protein binding data for animals and humans, systemic exposure data from Tox species (AUC, Cmax)
4.	Repeat dose toxicity studies (minimum duration: 2 weeks in 2 species)
5.	Genotoxicity (only in vitro: gene + chromosomal aberration)
6.	No carcinogenicity
7.	No single-dose studies or data on lethality needed
8.	No reproductive studies
9.	Other studies if concern, e.g., phototoxicity, etc.

Table 21 Investigations of metabolites

Human metabolite(s)
Nonclinical characterization of **human metabolite(s)**, **when** exposures greater than **10%** of total drug-related exposure
Such nonclinical data support phase III clinical trials
Some metabolites not of toxicological concern: (e.g., most glutathione conjugates). No need for testing
Cause for concern may be a unique human metabolite. Consider testing such metabolite in animals

supplemental studies as appropriate. During clinical development, a clarification of observed or suspected adverse effects in animals or during clinical trials may be needed. Before NDA, an assessment of effects on all systems should be provided.

For kinetics, an information on exposure data (AUC) in animals prior human clinical trials is needed, while ADME data are needed at completion of Phase I (human pharmacology) studies.

The duration for repeat dose toxicity studies is related to the duration of clinical trials and their therapeutic indication. In principle, the duration of animal studies is equal to or exceeds the duration of the human clinical trials.

In general, there is a relationship of 1:1 ratio for studies in two mammalian species (one non-rodent). The details can be found in Table 22.

This table was updated by thorough additional Japanese studies in 2000 during the 5th ICH

Table 22 Duration of repeated dose toxicity studies to support phase I and II trials in EU and phase I, II, and III trials in the USA and Japan

Maximum duration of clinical trial	Recommended minimum duration of repeated-dose toxicity studies to support clinical trials	
	Rodents	Non-rodents
Up to 2 weeks	2 weeks	2 weeks
Between 2 weeks and 6 months	Same as clinical trial	Same as clinical trial
> 6 months	6 months	9 months

Conference in San Diego, USA. The Japanese scientists compared the utility of routine 4-week toxicity studies with 2-week studies and concluded that in regard to the prediction of toxicities to the male reproductive organs, 2-week studies were as valid as 4-week studies; therefore, the former regional Japanese requirement to ask for a minimum duration of 4-week studies before starting trials in men was dropped for a global consensus that the minimum duration of non-clinical studies is 2 weeks in rodents and 2 weeks in non-rodents.

On the other hand, regional differences continue in regard to single-dose animal studies supporting single-dose studies in humans.

In the USA, single-dose toxicity studies with extended examinations can support single-dose human trials. This concept encouraged the EU to offer comparable options with the microdosing concept. This principle may be especially valuable for gaining early data for go/no go decisions, when several candidates are being developed in parallel. Table 23 summarizes the different options and their conditions before applying drugs to humans.

Results from two in vitro genotoxicity studies are recommended to be available prior to first administration to humans, while the standard battery should be completed prior to initiation of Phase II studies.

Carcinogenicity studies do not need to be completed in advance of the conduct of clinical trials unless there is cause for concern (ICH: S1A).

For pharmaceuticals to treat certain serious diseases, carcinogenicity testing, if needed, may be concluded post-approval.

Table 23 Options for exploratory assays

Exploratory clinical trials			
Study type	PDy	General Tox	Genotox
Microdose	PD in vitro + in vivo	Single dose, 1 species,	No studies,
		1 route	SAR recommended
Repeat microdose	PD in vitro + in vivo	7 day repeat,	No studies,
		PK, Hem. Chem, path	SAR recommended
Single dose	PD in vitro + in vivo + SP	Extended single dose	Ames
		Day repeat,	
		PK, Hem. Chem, path	
Up to 14 days	PD in vitro + in vivo + SP	2 week study, 2 species	Ames + Chro, abb test
		PK, Hem. Chem, path	
Up to 14 days	PD in vitro + in vivo + SP	2 week study, 2 species, non-rodent as confirmatory, PK, Hem., Chem, path	Ames + Chrom, abb test in vitro or in vivo

The inclusion of different patient populations reveals regional differences, especially for women with child bearing potential. There is a high level of concern for unintentional exposure of an embryo/fetus.

The currently regional differences in the timing of reproduction toxicity studies to support the inclusion of women with childbearing potential are:

1. **Japan:** assessment of female fertility and embryo-fetal development should be completed prior to the inclusion of women of childbearing potential using birth control in any type of clinical trial.
2. **EU:** assessment of embryo-fetal development should be completed prior inclusion of Women of Childbearing Potential, prior Phase III trials female fertility studies are needed.
3. **US:** women of childbearing potential may be included in early, carefully monitored studies without reproduction toxicity studies provided appropriate precautions are taken to minimize risk.
4. **US:** assessment of female fertility and embryo-fetal development should be completed before Phase III trials.

The inclusion of **children** into clinical trials has gained tremendous interest. The following data are requested:

- Safety data from previous adult human exposure:
  - Most relevant information
  - Necessary before pediatric clinical trials
- Prior pediatric trials:
  - Appropriate repeated dose toxicity studies
  - All reproduction toxicity studies
  - Standard battery of genotoxicity tests
- Juvenile animal studies should be considered
- Carcinogenicity testing:
  - Prior to long-term exposure of children, if cause for concern

With the revision of this guideline, a number of important aspects have been incorporated into the overall strategies.

It is stressed that drug development for the preclinical area is always a stepwise process where a number of goals is pursued. Short-term tests come first; investigations of life-threatening conditions dictate their timing. Long-term studies like carcinogenicity or reproductive toxicity studies accompany the clinical Phases II and III.

The guideline provides answers which metabolites need to be characterized.

M3R takes up an old idea of the FDA, to facilitate the kinetic access to humans. FDA had offered already in the late 1990s to use the concept of single-dose toxicity study to allow a single dose in humans. With M3R additional options are given.

Very important are the recommendations of M3R for the inclusion of children into clinical trials (Tables 24, 25, 26, and 27). One needs to realize that children are not small adults; they can be more vulnerable or sometimes more resistant than adults. Their organs need years to mature. The past did not consider the pediatric population to test their tolerance, clinical or toxicological

Table 24 Safety requirements for pediatric population

Clinical trials in **pediatric** populations (1)
Requirements:
Safety data from previous adult human experience, as most relevant information
Review quality and extent of adult human data
Exception: Extensive adult experience might not be available before pediatric exposures (e.g., for pediatric-specific indications)

Table 25 Studies needed for pediatric populations

Clinical trials in pediatric populations
Results from repeat dose toxicity studies of appropriate duration in adult animals
Core safety pharmacology package
Standard battery of genotoxicity tests
Reproduction toxicity studies relevant to the age and gender of the pediatric patient populations with information on direct toxic or developmental risks (e.g., fertility and pre-postnatal developmental studies)
Embryo-fetal developmental studies are not critical to support clinical studies for males or prepubescent females

Table 26 Use of juvenile animals

Pediatric populations + juvenile animals studies
The conduct of any juvenile animal toxicity studies should be considered only when previous animal data and human safety data, including effects from other drugs of the pharmacological class, are judged to be insufficient to support pediatric studies
One relevant species, preferably rodent, is adequate
A non-rodent species appropriate when scientifically justified
Juvenile animal toxicity studies are not considered important for short-term PK studies (e.g., 1–3 doses) in pediatric populations

aspects. Most of the drugs were in former times never tested for children. This has lead now to a paradigm change: We try not to protect children against clinical trials but to protect the pediatric population by conducting specifically designed trials. And the agencies expect now that all activities in regard to the pediatric population should be planned in advance by submitting to the agencies a **pediatric investigation plan (PIP)** as soon as adult PK data available, which means, i.e., end of Phase 1.

Table 27 Pediatric population: need or no need of preclinical studies

Clinical trials in pediatric populations
A **chronic repeat** dose study: Initiated in the appropriate age and species to address this developmental concern (e.g., 12-month duration dog or 6 month in rodent)
A 12-month study covers full development period in the dog
Carcinogenicity studies not generally recommended to support the conduct of pediatric clinical trials
Consider carcinogenicity testing before long-term exposure in pediatric clinical trials, when there is:
Evidence of genotoxicity in multiple tests
Concern for pro-carcinogenic risk based on mechanistic considerations
Findings from general toxicity studies

An important recommendation of M3 guidance refers to the question how to test drugs which are often administered in combination. The strategies are dictated by the possibility of an increase of concern when combined. Testing is needed when little is known for the different compounds, e.g., when compounds are still in early development, but very little is done when the pharmacological and toxicological characteristics are well investigated. Tables 28 and 29 give examples of these two extremes.

Common Technical Document (ICH/M4)

M4 is another very important multidisciplinary guideline, which combines information for the three disciplines within the ICH Process: for quality, safety, and efficacy. The following section focuses predominantly on the preclinical safety issues (Fig. 13). The total document is divided into five modules: Module 1 contains regional specific aspects, it provides for the European Union, e.g., the European Community specific data. This module therefore is not harmonized but region-specific. Module 2 provides the summaries for quality, for safety, and efficacy. The quality part uses as a headline "Quality Overall Summary," for safety and efficacy the terms "Non-Clinical or Clinical Overview," The different names signal that the quality part is a clear summary, while the nonclinical and

Table 28 Activities required for combination of two late-stage compounds

Combination of late-stage drugs
Two late-stage entities with **adequate clinical experience, combination toxicity** studies **not** recommended for support clinical studies or marketing, unless significant toxicological concern (e.g., similar target organ toxicity)
No concern when:
Margins of safety high
Monitoring of adverse effects in humans easy
If concern, complete preclinical combination study before start of clinical trials

Table 29 Activities required for combination of two early-stage compounds

Combination of two early-stage entities:
Nonclinical combination toxicity studies recommended
Duration of clinical trials and nonclinical comparable: Ratio 1.1
A 90-day combination toxicity study would support clinical trials and marketing

clinical part should be critical evaluations. Module 3 provides chemical, pharmaceutical, and biological Quality information. Module 4 contains the nonclinical reports and Module 5 provides clinical study reports.

The objectives of M4 are to assist authors in the preparation of nonclinical pharmacology, pharmacokinetics, and toxicology written summaries in an acceptable format. The CTD is not intended to indicate what studies are required but provides an appropriate format for the nonclinical data. The Common Technical Document is nothing other than a placeholder for the different parts of a documentation for the market authorization process.

No guideline can cover all eventualities; common sense and a clear focus on needs of regulatory authority assessor are best guides to constructing an acceptable document. Therefore, modify the format if needed with the aim to provide best possible presentation and facilitate the understanding for the evaluation of the results.

The CTD-S is organized as follows:

F. Nonclinical summary

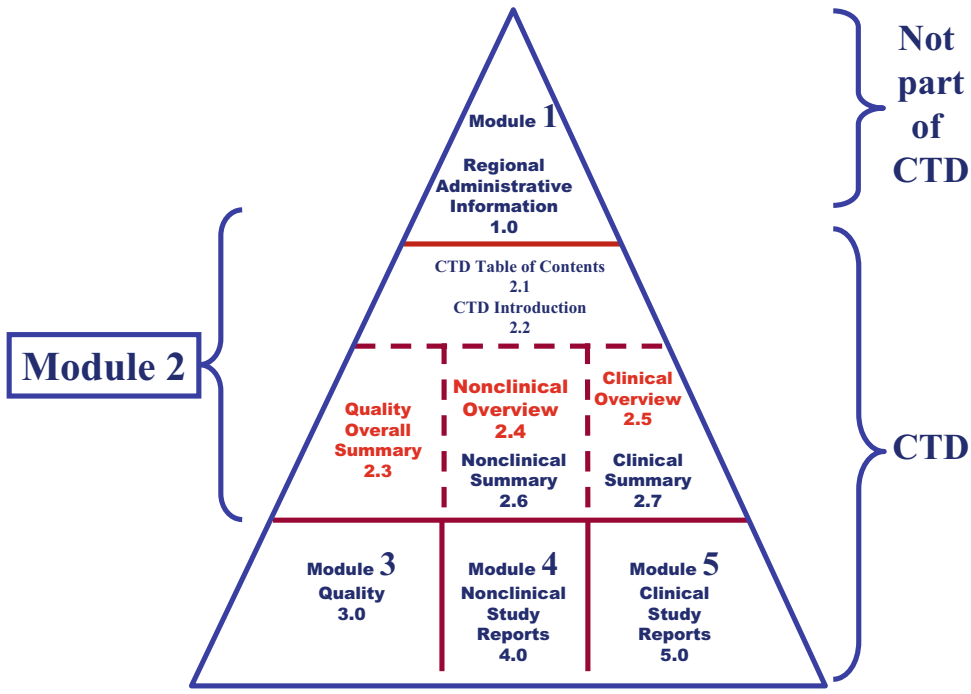


Fig. 13 Structure of common technical document (CTD)

1. **Pharmacology**
   (a) Written summary
   (b) Tabulated summary
2. **Pharmacokinetics**
   (a) Written summary
   (b) Tabulated summary
3. **Toxicology**
   (a) Written summary
   (b) Tabulated summary

This organization is kept up in all parts of the dossier; it is repeated for the overview, the summaries, and the reports.

The detailed organization for pharmacology, kinetics, and toxicology are as follows:

1. **Pharmacology** Written Summary
   - Brief Summary
   - Primary Pharmacodynamics
   - Secondary Pharmacodynamics
   - Safety Pharmacology
   - Pharmacodynamic Drug Interactions
   - Discussion and Conclusions
   - Tables and Figures (either here, or included in text)
2. **Pharmacokinetics** Written Summary
   - Brief Summary
   - Methods of Analysis
   - Absorption
   - Distribution
   - Metabolism
   - Excretion
   - Pharmacokinetic Drug Interactions (Nonhuman)
   - Other Pharmacokinetic Studies
   - Discussion and Conclusions
3. **Toxicology** Written Summary
   - Brief Summary
   - Single-Dose Toxicity
   - Repeat Dose Toxicity
   - Genotoxicity
   - Carcinogenicity
   - Reproduction Toxicity Local Tolerance
   - Other Toxicity Studies
   - Discussion and Conclusions

Examples of detailed advice for sections on discussion and conclusion of pharmacokinetics: Information should be integrated across studies and across species; exposure in the test animals should be related to exposure in humans given the maximum intended doses.

Similar examples for toxicology: in vitro studies should precede in vivo studies. Where multiple studies of the same type need to be summarized within the pharmacokinetics and toxicology sections, studies should be ordered by species, by route, and then by duration (shortest duration first).

The species should be ordered as follows: (1) mouse, (2) rat, (3) hamster, (4) other rodent, (5) rabbit, (6) dog, (7) nonhuman primate, (8) other non-rodent mammal, and (9) non-mammals (see also Table 17.)

It is also recommended to limit the information in the summaries and overview. The overview should contain the essential and critical results on approximately 30 pages. The length of the Nonclinical Written Summaries should in general not exceed 100–150 pages (Table 30).

The brief summaries for pharmacology should be written on 2–3 pages, for pharmacokinetics the same length and for toxicology approximately 6 pages.

The tabulated summaries (examples given in Figs. 14 and 15 and Table 31) help to get a quick insight into the detailed data. A first thorough review is therefore based on reading the nonclinical overview in connection with some of the detailed tabulated summaries.

This CTD has been tested in practice for several years. Although not perfect for every case, it has proven its usefulness. Researchers/Industry know where to place specific information and data and regulators know where to find them. This has facilitated the review process tremendously; only one dossier is necessary for international registration; a lot of resources can be diverted to more important issues.

Table 30 Summary of pages number for CTD

Number of pages recommended in CTD
1. Nonclinical overview 30 pages, use:
Pharmacology brief summary 3 pages
Pharmacokinetics brief summary 3 pages
Toxicology brief summary 6 pages
2. Written summaries 150 pages:
3. Module 4 = reports as long as needed

Fig. 14 Example of tabulated summary **(Pharmacology)**

1 Pharmacology	Overview		Test Article	
Type of Study	Test System	Method of Administration	Study Number	Location Volume, Page
1.1 Primary Pharmacodynamics				
1.2 Secondary Pharmacodynamics				
1.3 Safety Pharmacology				
1.4 Pharmacodynamic Intervention				

1.3 Safety Pharmacology **Test Article:**

Organ Systems Evaluated	Species/ Strain	Method of Admin.	Doses (mg/kg)	Gender and No. per Group	Noteworthy Findings	GLP Compliance	Study Number

Fig. 15 Example of tabulated summary for **safety pharmacology** studies

Table 31 Example of parts of the **toxicology program**

Tabulated summaries: toxicology
Toxicology overview
Overview of toxicokinetics studies
Overview of toxicokinetics data
Batches (and their impurities) used in toxicology studies
Single-dose toxicity
Repeat dose toxicity (non-pivotal and pivotal)

But the CTD is a living document; its weakness is apparent. In order to improve its quality and practicability, suggestions and proposals for improvements are invited from the public. These modifications are being dealt with within the process of "Questions and Answers," the ICH experts publish their conclusions regularly.

Outlook and future of ICH

ICH guidelines are living and dynamic documents which are submitted constantly to revisions and modifications. This is happening to the Common Technical Dossier also. The FDA is proposing a new method which is supported by Japan but also followed with great interest by the European Medicines Agency. Experts work on a Standards for Exchange of Nonclinical Data (SEND). This SEND is an implementation of the CDISC Standard Data Tabulation Model (SDTM) for nonclinical studies, related to animal testing conducted during drug development. Send will facilitate the transfer of raw data of toxicology animal studies. When such experimental studies started after December 18, 2016, to support submission of new drugs to the US Food and Drug Administration, then such data should be submitted to the agency using SEND. Recommendations for implementing SEND, including how to model various nonclinical endpoints, rules to doing so, and examples with sample data are available on the CDISC SEND website.

The Pharmaceuticals and Medical Devices Agency in Japan will enforce its use in the future also, most probably in 2020.

More details are offered by, e.g., Caroline Hroncich (2016) (Steven Denham, MPI Research), the impact of SEND on the pharmaceutical industry:

SEND is a nonclinical version of the Study Data Tabulation Model (SDTM), the standard format for electronic submission of clinical data to FDA. SEND and SDTM are expected to speed up the review process for drug applications by "developing electronic tools to analyze and visualize these submissions, and building data warehouses to rapid query data across drugs, companies, and clinical and nonclinical disciplines."

SEND will become a mandatory part of submissions for relevant investigational new drugs (INDs), new drug applications (NDAs), biologics license applications (BLAs), and abbreviated new drug applications (ANDAs) and will start by December 17, 2017. From now on companies have to provide electronic standards for capturing individual data from nonclinical studies, including both metadata about how the study is conducted and the actual data itself.

The FDA is trying to move away from paper that is creating a physical storage problem. It should improve the review process; not only for review time but for the safety aspects in that there would be a data warehouse at the agency that can be searched across studies and across sponsors.

SEND recommends how one stores truckloads of data and be able to access it in a finite review time and compare it with data from a similar active ingredient. Those studies could well have been submitted by different sponsors. The end result is that they can recommend a more focused approach to safety observations in clinical trials.

For a new submission submitted to FDA/CDER after March 15, 2020, studies need to adhere to the new standard. The other studies in the submission which started before March 15, 2020, are not required to be submitted according to the new standard, although it is encouraged/preferred. Visit the SEND CT page to get the most recent CT and contact edata@fda.hhs.gov for additional advice about such a submission.

The FDA will use the files for the review process, via the Nonclinical Information Management System (NIMS) suite. This suite will facilitate the review of submissions more efficiently than with only PDF or printed submissions that contain the individual animal data.

SEND will only be a requirement in the USA for certain FDA submissions. However, it has operational use, such as transfer between organizations, sponsor warehousing, etc. Accordingly, it is preferable to produce SEND datasets, even if not technically required for submission. However, it is a longer-term goal for the SEND datasets to eventually replace the individual tabulated datasets. A SEND package consists of a number of dataset files (in XPT format, a.k.a., SAS v5 transport format) and a define.xml file (which provides information about what is in the datasets).

The publicly available sample SEND datasets can be seen in, e.g., PhUSE github, https://github.com/phuse-org/phuse-scripts/tree/master/data/send or contact submit@instem.com offering a subject line of "Send me SEND" to get an FDA validated SEND dataset for an example 28 day toxicity study. Finally go to SENDDataSet.org to download an FDA validated multi-organizational SEND dataset.

SEND will offer data mining opportunities, which will be used by the FDA/industry. The benefits can only be realized with time when significant sufficient database of historical data or studies are converted and loaded into repository systems to facilitate such queries.

Developers very certainly should follow FDA recommendation to avoid any delays for the marketing authorization process.

Outlook on ICH

ICH can look back to 25 years of achievements. As the most successful international harmonization initiative in the world, its achievements include an understanding of innovation, a common regulatory platform based on more than 70 guidelines, improvement of scientific guidelines, facilitating communication between different regulatory agencies and between industry and the regulatory authorities, effective use of research and development (R&D) resources, and greater mutual acceptance of R&D data realized.

In addition, the drafting process of guidelines has identified gaps in science; new studies analyzing problems have been conducted to validate the assays being used to confirm safe use of drug in humans, like validity of 2-week studies to assess male fertility or the validity of transgenic mouse models for assessing the carcinogenic potential within the ILSI/HESI evaluation process.

Further guidance documents are in preparation or released for consultation.

Among them gets the guideline on environmental risk assessments for pharmaceuticals more importance. Genotoxic impurities are

handled with great care. Aspects like drug-induced hepatotoxicity, gene transfer, genomics, etc. are waiting to be assessed for regulatory advice.

It is a great success that guidelines in general today are science driven and based on valid assessment in practice. The open-minded acceptance of new technologies, e.g., in vitro systems, molecular databases, in silico data collections, computer sciences, imaging, and critical use of omics as animal and human biomarkers will continue to deliver high-quality products. At the same time, there is awareness that with the increase of high-quality results, the attrition rate does not increase.

The improved cooperation of regulatory agencies and industry has created an atmosphere of mutual confidence and trust. The open dialogue attracts scientific strategic skillfulness and reduces rigidity.

The road of success is the cooperation on a level below legislation while retaining national sovereignty. Many issues have been resolved; several problems wait for a harmonized solution. Enthusiasm continues; dynamic contributions support the further process of improving and facilitating pharmaceutical development under the leadership of ICH with the continuous aim to bring safe drugs faster to the patient.

References and Further Reading

Ames BN et al (1973) Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. PNAS 70:2281–2285

Bauer D, Averett LA, De Smedt A, Kleinman MH, Muster W, Pettersen BA, Robles C (2014) Standardized UV-vis spectra as the foundation for a threshold-based, integrated photosafety evaluation. Regul Toxicol Pharmacol 68(1):70–75

Berwald Y, Sachs L (1965) In vitro transformation of normal cells to tumor cells by carcinogenic hydrocarbons. J Natl Cancer Inst 35:641–661

Bode G, Clausing P, Gervais F, Loegstedt J, Luft J, Nogues V, Sims J (2010) The utility of the minipig as an animal model in regulatory toxicology. Journal Pharm Tox Methods 62:196–226

Contrera JF1, Jacobs AC, DeGeorge JJ (1997) Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. Regul Toxicol Pharmacol 25(2):130–145

Cohen SM, Robinson D, MacDonald J, Health and Environmental Sciences Institute (2001) Alternative models for carcinogenicity testing. Toxicol Sci 64(1):14–19. ilsi.org/publication/alternative-models-for-carcinogenicity-testing/

Contrera JF, Aub B, Barbehenn E, Belair E, Chen C, Evoniuk G, Mainigei K, Mielach F, Sancilio L (1993) A retrospective comparison of the results of 6 and 12 months non-rodent toxicity studies. Adverse Drug React Toxicol Rev 12(1):63–76

Corvi A, Albertinein S, Hartung T, Hoffman S, Maurici D, Pfuhler S, van Bentham J, Vanparys P (2008) ECVASM retrospective validation of in vitro micronucleus test (MNT). Mutagenesis 23(4):271–283

EMA/CHMP/ICH/536328/2013 Rev (2016a) 1 Committee for human medicinal products ICH guideline S1 Regulatory notice on changes to core guideline on rodent carcinogenicity testing of pharmaceuticals, Mar 2016. www.ema.europa.eu/docs/...guideline/.../WC500136405.pdf

EMA/CHMP/ICH/536328/2013 Rev (2016b) 1 CHMP (Committee for human medicinal products) ICH guideline S1 Regulatory notice on changes to core guideline on rodent carcinogenicity testing of pharmaceuticals, status 29 Feb 2016

Estimating the maximum safe starting dose in initial clinical trials for Therapeutics in adult healthy volunteers (CDER July 2005)

Forster R, Bode G, Ellegaard, van der Laan W (2010) The rethink project. Minipigs as models for the toxicity testing of new medicines and chemicals : an impact assessment. Journal Pharm Tox Methods 62:158–159 and 236–242

Friedrich A, Olejniczak K (2011) Evaluation of carcinogenicity studies of medicinal product for human use authorized via the European centralized procedure (1995–2009). Regul Toxicol Pharmacol 60:225–248

From http://www.emea.eu.int/htms/human/ich/safety/ichfin.htm

Guidance for Industry Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) (2005) Pharmacology and toxicology, July 2005

Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use S2(R1) (2011) Step 4 version dated 9 Nov 2011

Hroncich C (2016) The impact of SEND on the pharmaceutical industry. Pharm Technol 2016(1):28–30 (Discussions with Denham S, MCI)

ICH Guideline (2009) M3(R2) on non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals Step 5 from 2009

ICH M 4 R Organisation of the Common Technical Document for the Registration of Pharmaceuticals for Human Use M4 (2016) Current Step 4 version dated 15 June 2016

ICH S 11 (2014) Non-clinical Safety testing in support of development of paediatric Medicines, Concept paper, step 1

Jacobs A, von der Laan JW (2012) Chapter 38. In: McAnulty P, Dayan A, Ganderup NC, Hastings KL (eds) The minipig in biomedical research. CRC Press, Boca Raton, pp 565–571. ISBN 978-1-4398-1118-4

Kirkland D, Aardema M, Mueller L, Hayashi M (2006) Evaluation of the ability of a battery of 3 in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. II. Further analysis of mammalian cell results, relative predictivity and tumor profiles. Mutat Res 608:29–42

Kishi K1, Kanamori S, Maruyama T, Sasaki K, Hara K, Kawai M, Ikeuchi K (1995) Potential parameters of male reproductive toxicity: reproductive performance, histopathology and sperm evaluation in SD rats given nitrazepam. J Toxicol Sci 20(3):329–339

Kummar et al (2009) Phase 0 clinical trial of the poly (ADP-Ribose) polymerase inhibitor ABT-888 in patients with advanced malignancies. J Clin Oncol 27 (16):2705–2711

LeBoeuf RA, Kerckaert GA (1986) The induction of transformed-like morphology and enhanced growth in Syrian hamster embryo cells grown at acid pH. Carcinogenesis 7:1431–1440

LeBoeuf et al (1996) Mutat Res 356:85–127

Lynch AM, Wilcox P (2011) Review of the performance of the 3T3 NRU in vitro phototoxicity assay in the pharmaceutical industry. Exp Toxicol Pathol 63(3):209–214

Lynch AM, Robinson SA, Wilcox P, Smith MD, Kleinman M, Jiang K et al (2008) Cycloheximide and disulfoton are positive in the photoclastogenicity assay but do not absorb UV irradiation: another example of pseudophotoclastogenicity? Mutagenesis 23(2):111–118

Lumley C, van Cauteren H (1997) Harmonisation of international toxicity testing guidelines for contribution in refinement and reduction in animal use. EBRA Bull 1997; pp 4–9

Mauthe RJ1, Gibson DP, Bunch RT, Custer L (2001, 2001) The Syrian hamster embryo (SHE) cell transformation assay: review of the methods and results. Toxicol Pathol (29 Suppl):138–146

McAnulty P, Dayan A, Ganderup NC, Hastings KL (2012) The minipig in biomedical research. CRC Press, Boca Raton. ISBN 978-1-4398-1118-4

Muller PY, Milton M, Lloyd P, Sims J, Brennan FR (2009) The minimum anticipated biological effect level (MABEL) for selection of first human dose in clinical trials with monoclonal antibodies. Curr Opin Biotechnol 20(6):722–729. https://doi.org/10.1016/j.copbio.2009.10.013 Epub 5 Nov 2009

Nonclinical Evaluation for Anticancer Pharmaceuticals S9 (2009) Current Step 4 version dated 29 Oct 2009

OECD Guideline for Testing of Chemicals in vitro 3T3 NR U phototoxicity test (2004)

Pentsuk N1, van der Laan JW (2009) An interspecies comparison of placental antibody transfer: new insights into developmental toxicity testing of monoclonal antibodies. Birth Defects Res B Dev Reprod Toxicol 86(4):328–344

Photosafety Evaluation of Pharmaceuticals S10 (2013) Current Step 4 version dated 13 Nov 2013

Putman E, van Loveren H, Bode G, Dean J, Hastings K, Nakamura K, Verdier F, van der Laan JW (2000) Assessment of immunotoxic potential of human pharmaceuticals. Drug Info J 36:417–427

Sakai T, Takahashi M, Mitsumori K, Yasuhara K, Kawashima K, Mayahara H, Ohno Y (2000) Collaborative work to evaluate toxicity on male reproductive organs by 2- week repeated dose toxicity studies in rats. Overview of the studies. J Toxicol Sci 25:1–21

Sistare FD1, Morton D, Alden C, Christensen J, Keller D, Jonghe SD et al. (2011) An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. Toxicol Pathol 39(4):716–44.

Takayama S, Akaike M, Kawashima K, Takahashi M, Kurokawa Y (1995) A collaborative study in Japan on optimal treatment period and parameters for detection of male fertility disorders in rats induced by medical drugs. J Amer Coll Toxicol 14:266–292

The Clinical Evaluation of QT/QTC Interval Prolongation and Proarrhythmic Potential for Nonantiarrhythmic Drugs E14 (2005) Current Step 4 version dated 12 May 2005

Topic S1C Note for Guidance on Dose Selection for Carcinogenicity Studies of Pharmaceutical

Topic S1C(R) Note for Guidance on Dose Selection for Carcinogenicity Studies of Pharmaceuticals: Addition of a Limited Dose and related Notes

Topic S2A Note for Guidance on Genotoxicity: Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals

Topic S2B Note for Guidance on Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals

Topic S3A Note for Guidance on Toxicokinetics: A Guidance for Assessing Systemic Exposure in Toxicology Studies

Topic S3B Note for Guidance on PharmacoKinetics: Guidance for Repeated Dose Tissue Distribution Studies

Topic S4A Note for guidance on Duration of Chronic Toxicity Testing in Animals (Rodent and non Rodent Toxicity Testing)

Topic S5A Note for Guidance on Reproductive Toxicology: Detection of Toxicity to Reproduction for Medicinal Products

Topic S5B Note for Guidance on Reproductive Toxicology: Toxicity on Male Fertility

Topic S6 Note for Preclinical Safety Evaluation of Biotechnology-Derived Products

Topic S7A Note for Guidance on Safety Pharmacology Studies for Human Pharmaceuticals

Topic S7B Note for Guidance on Non-Clinical Evaluation of the potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals

Topic S8 Immunotoxicity Studies for Human Pharmaceutical

Trepakova ES et al (2009) A HESI consortium approach to assess the human predictive value of non-clinical

repolarization assays. J Pharmacol Toxicol Methods 60(1):45–50. https://doi.org/10.1016/j.vascn.2009.05.002 Epub 9 May 2009

Van Cauteren H, Bentley P, Bode G, Cordier A, Coussment W, Heining P, Sims J (2000) The industry view on longterm toxicology testing in drug development of human pharmaceuticals. Pharmacol Toxicol 86(Suppl I):1–5

van der Laan, van Loveren H, Vos JG, Dean JH, Hastings K (1997) Immunotoxicity of pharmaceuticals: current knowledge, testing strategies, risk evaluation & consequences for human health. Drug Info J 31:1301–1306

Vargesson N (2015) Thalidomide induced teratogenesis: history and mechanisms. Birth Defects Res C Embryo Today 105(2):140–156

www.ema.europa.eu/docs/...guideline/.../WC500136405.pdf. Revision of ICH guideline S 1 A

www.ema.europa.eu/docs/en...guideline/.../WC500002699.pdf. Topic S1A note for guidance on the need for carcinogenicity studies of pharmaceuticals

www.ema.europa.eu/docs/en...guideline/.../WC500003258.pdf. Topic S1B mote for guidance on carcinogenicity: testing for carcinogenicity of pharmaceuticals

Clinical Quality Management System

60

Beat Widler

Contents

Abstract

The ICH (R2) update finally introduces to the world of clinical trials quality and compliance management concepts such as quality by design and quality risk management that have been successfully applied in other industries such as the automotive and airline industries. This chapter reviews how these concepts as well as a high degree of standardization of processes and tools (e.g., trial protocols) can not only drive compliance and quality but also efficiency and cost-effectiveness when engaging in a clinical development program. A lean quality management system needs to support such an approach by focusing on lean processes whereas quality is measured and not "estimated." A proactive mindset with a clear risk management approach is a way to a successful clinical trial conduct.

Quality Is More than Compliance

Even 20 years after the issuing of the ICH GCP Guidelines (ICH E6), the principles of GCP are still sound, and little can be criticized about them. The updates of the Guidelines and especially its

B. Widler (✉)
Widler & Schiemann AG, Zug, Switzerland
e-mail: beat.widler@wsqms.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_42

Addendum [ICH E6(R2)] that became effective in Europe in 2017 have not fundamentally changed the content of this Guideline but introduced technology and approaches that were unknown when the GCP Guideline got issued for the first time. Twenty years ago clinical development was largely a paper-based process, and the Internet and even more so the *Cloud* were technology and tools only known to a few *techies*.

The value of ICH E6 is also evidenced by the fact that ICH GCP successfully made its inroads even into new regions where public welfare priorities and medical practice differ from what is the norm in the original ICH regions: Western Europe, USA, and Japan. When these principles are so robust and still so sound, why is there then an ongoing debate about the inefficiencies of the clinical trial process, the growing disinterest of physicians in clinical trial activities, and an apparently dwindling quality of results generated through clinical trials which results in distrust by patients and the public at large into the clinical trials enterprise? What is causing this disenchantment? Could it be that we – sponsors, health authorities, and also public health services – burdened the sound principles of GCP with bureaucratic requirements that at first seemed to be "good ideas" but as a matter of fact do not contribute to what are the two fundamental dimensions of GCP and quality: protecting patients' safety, right, and integrity as well as ensuring data integrity? Is the ICH GCP revision and especially its *Addendum*, which became effective this year, going to improve the way in which GCP gets implemented? In fact, this *Addendum* stresses the importance of the sponsor's oversight on all stakeholders in a clinical trial, the **personal** responsibilities of the principal investigator, and the roles and responsibilities of a CRO. The *Addendum* also highlights the importance of a robust quality management system and calls for a structured risk assessment of protocols and processes applied to execute a trial protocol that are enablers of risk-based monitoring or what is a more accurate rendition of the concept: evidence-based trial quality management.

When implemented correctly *evidence-based trial quality management* allows:

1. Ensuring the protection of patients' safety, integrity, and rights.
2. As the flip side of the above, ensuring data integrity what is nothing else than the long-term dimension of the first requirement: avoid any false positive or negative conclusions because the data collected in a trial is faulty.

When this principle is recognized and correctly implemented – remember decisions should be data/evidence driven and not be based on opinions, even if decisions are reached by team consensus, as stated by Edwards Deming's *quality without data, is just another opinion* – then a liberating epiphany emerges: errors "that are understood" or "factored into the process," on the basis of the sound principles of a Quality by Design (QbD) methodology and approach, do not matter! In other terms, on this basis an isolated GCP non-compliance, such as an isolated transcription error, omission of a "minor" adverse event can be accepted as "forgivable sins." Conversely, system failures that do result or may result in a risk to the two fundamentals of GCP must be identified, proactively dealt with, and if they have materialized be corrected or at least mitigated swiftly.

Methodologies of Quality Risk Management (QRM) and Quality by Design are instrumental for a systematic quality management approach that allows focusing on the essentials of GCP.

Elements of a Pharma/Health-Care Quality Management System (QMS)

Any QMS should include the following elements:

A. **People**, i.e., those individuals who have a role and responsibility in a
B. **Process** related to the development of a new medicinal product/health-care activity. Processes are typically described in SOPs (standard operating procedures) and systems' (validation) documentation and are owned by the business units involved in these activities. A quality manual is a useful tool to summarize the quality principles and high-level processes of an organization.
C. **Controls**, i.e., those activities implemented by the business or process owner under the oversight of the independent quality assurance

(QA) unit to identify and prevent process deviations and defective products. "Classical" control elements are review and approvals of essential documents, monitoring and co-monitoring of process and its deliverables, auditing, and inspections to verify and confirm compliance. Controls must include evidence (documentation) reviews and approvals that have been executed in a timely manner. In a Quality Risk Management, environment control needs to include KPIs (key performance indicator) and KRIs (key risk indicator) to allow for a continuous monitoring of the performance of systems and products or outputs meeting predetermined specifications.

D. **Documentation**, i.e., a transparent description of the systems and processes used allowing to reconstruct at any time the sequence of events as well as the body of evidence of compliance with stipulated checks and controls.

Quality by Design (QbD) Builds on Robust, Smart, and Well-Documented Processes

The classical approach to a QMS has its shortcomings as more and more the costs (for instance, for monitoring clinical trial centers) and resources needed for the management and controls of a clinical development process are not matched by a commensurate process efficiency and quality of the "product." From a QbD perspective, quality is best defined as *a product, service or process that meets customers' needs*, whereas the customer can be an independent third party such as a buyer, regulator, prescriber, etc. or an "internal" client, e.g., the next in the value chain. To define the process leading to a quality product or service, a concept known from the 6-Sigma methodology – SIPOC – has proven to be very effective. We prefer to refer to SIPOC^2:

- **S**upplier, a **named** individual who delivers a product or service – can also be an instruction – that enables the next individual in the value chain to execute a predefined task. It is good practice not to designate a *team* as the supplier; if a function is designated as the supplier in a process chart, then within the *supplier's* organization, a named individual needs to be identified.
- **I**nput, the (sometimes, semifinished) product or service – can also be a decision such as an approval – that serves as the building element for the next "production" step.
- **P**rocess, predefined, documented, and agreed sequence of activities that each meets predefined, documented, and agreed specifications. Critical process steps are linked to control steps.
- **O**utput, the deliverable – can be a product, service, or decision – of the above process.
- **C**ustomer, a **named** individual (exceptionally a functional entity) who receives the above output. The customer can be a third party or internal client.
- **C**ontrols, this is the C^2 in SIPOC and refers to all quality verification and governance activities as well as any corrective and preventive actions.

The purpose of the SIPOC^2 approach is to break down a complex (generally a multistep) process into discrete elements to drive transparency and accountabilities. Robust controls are defined and implemented at each supplier – customer interface and also specifications for input and output.

The SIPOC^2 concept can also be applied to the design of a protocol and the planning and implementation of a clinical trial. For instance, the protocol "designer" should not only define clear roles and responsibilities for all protocol tasks but as part of the *process* and *control* description also anticipate what could go wrong, build up-front contingency plans, define controls to identify early deviations from the design specifications, and include in the design of a new trial learnings from past, good, or bad experience.

Quality by Design and Quality Risk Management (QRM)

The QbD approach and QRM are intimately related. The smart implementation of a QRM strategy leverages operational or transactional data generated as a by-product of the clinical or pharmacovigilance processes to return inferences about process robustness and compliance with set

specifications. Typically, QRM uses an array of so-called key risk indicators: a KRI is a measurable entity and is always associated with a threshold of acceptance/rejection. A KRI is comparable to a KPI but focuses on the quality rather than the efficiency aspect.

For instance, the audit trail generated for each change to a database entry (i.e., the GCP mandated tracking of the date of a change and the originator of a change) can be trended across all sites of a trial to determine whether an unexpected pattern in these changes emerges. The number of data entry changes is an example of a KRI. For example, investigators in a given country or clinical trial center may "produce" an above average number of changes. QRM requires that the root cause and reason for this "aberration" are investigated and understood in order to take either corrective and preventive action (CAPA) or to accept the fact that there is such an "aberration" when this has a logical and justified reason. As shown in this example, QRM typically uses meta-data as drivers and input for KRIs and thus takes *QC and QA by sampling* to a seamless oversight of mission critical activities, processes, and deliverables.

6-Sigma, FMEA (failure mode and effect analysis), and Kaizen are established methods or models supporting a QbD and QRM approach.

Quality by Design in Other Domains and Industries

ICH had started the discussion about QbD (i.e., ICH Q8, Q9, and Q10 guidelines) in the GMP area, and the revision of ICH GCP and especially its *Addendum* has extended these concepts also to clinical trials. Nevertheless, compared to other industries, the health care and pharmaceutical sector are lagging behind in applying QbD and QRM approaches. Therefore, learning from other industries on how to successfully introduce QbD and QRM processes should become a priority.

The airplane engine manufacturer Pratt & Whitney is a good example for the successful shift from a *trial and error* approach in product development to a disciplined and process driven *engineering standard work* (ESW). In ESW process steps are documented and described in extensive but targeted *Workflow Maps* with a focus on the interdependencies between the successive process steps. This approach establishes *design criteria* with clear deliverables for each ESW step and demands for clear and unequivocal *ownership* of each ESW element. ESW also introduced the *practitioner proficiency assessments* to capture coaching needs as well as coaching capabilities of individuals involved in product development. ESW is a prime example for the application of Edwards Deming's principles of the PDCA cycle (**P**lan–**D**o–**C**heck–**A**ct).

Back in the early 1950s, the US Navy introduced under the leadership of Hyman Rickover a QbD methodology and mind-set that allowed to operate nuclear reactors on their ships without any radiation incident for more than 60 years. Rickover was a stern advocate of a process driven approach. One of his fundamental contributions was reinventing the methodology of testing. In a traditional approach, testing is seen as an enabler for discerning *good* from *bad*. Rickover redefined the purpose of testing as an enabler for discerning an *understood* from a *not understood* process or process step. This coupled with his non-tolerance of a work-around approach – a work-around being seen as evidence of a poorly understood or designed process – a data driven and disciplined (also in terms of detail) problem-solving approach, as well as a low threshold for what is counted as an incident resulted in an unprecedented reliability of a highly complex system such as a nuclear reactor on a submarine.

Quality Means Standardization

The pharmaceutical industry has failed to date to develop and implement common standards for its key activities. As a result of this deficiency and inefficiency, investigators are frustrated, quality of clinical trials is negatively impacted, and inefficiencies in the clinical trial process are the norm.

This "state" is compounded by a silo mentality of this industry by which learnings of pre- or post-competitive nature around clinical trial activities are not shared or leveraged. The consequence is that the same errors are repeated again and again. From this follows that the successful implementation of QbD and QRM approaches will require that standards for routine processes get developed within the company and across companies. As demonstrated in other industries (e.g., airline industry), shared standards drive quality and eventually also efficiency as the need for retraining to different formats for an identical process becomes obsolete. Although TransCelerate – a cross-industry initiative – was established to streamline drug development processes, some of their achievements such as the publication of a template for a clinical protocol, a tool to conduct a risk assessment, etc., fall short of a true standardization: What would truly streamline the clinical trial process was developing and implementing for each "common" indication a shared protocol, which would simplify the review by ethics committees and health authorities, the implementation of a trial by the investigators and their teams, training of all involved stakeholders such as monitors, and also the setting up of the database, the eCRF, by the sponsor.

Misconceptions Around the Building of a QMS

A common error in establishing a QMS is to build it as an afterthought rather than a strategic priority. Especially, start-up companies fall into this trap by focusing on their "science" and consider SOPs and other elements of the QMS as a nuisance rather than an asset. From experience a robust QMS should be in place no later than 6 months prior to starting entry into men trials. On one hand also inspectorates have moved to a *risk-based approach* and may inspect a trial as soon as phase 1 which is a significant change to past practices when inspections were triggered by a submission of a marketing authorization dossier. Moreover, thinking early about the scope and content of the QMS allows streamlining the procurement policy and aligning processes across the various functions of a company. Lack of processes and procedures inevitably lead to waste of time (rework because of errors caused by miscommunication) and money (bad contractual terms or even switching of service providers because of unsatisfactory performance). As a result unhappy or overworked staff members are a sad consequence which often exacerbates the compliance challenge.

Conclusive Remarks

The successful implementation of a modern QMS based on QbD and QRM approaches does not depend so much on the choice of the right methodology or tool kit but primarily requires a change in mind-set that must be initiated by the top management of the involved organizations. QbD needs a long-term commitment and is not a short-term measure to realize cost savings. It will eventually result in *operational excellence* if the process is applied consistently and in a disciplined manner. In this context Deming's profound observations are of significance: "in the 1970s, Dr. Deming's philosophy was summarized by some of his Japanese proponents with the following 'a'-versus-'b' comparison:

(a) When people and organizations focus primarily on quality, quality tends to increase and costs fall over time.
(b) However, when people and organizations focus primarily on costs, costs tend to rise and quality declines over time."[1]

Moreover, it should also be emphasized that the QMS must be owned by senior management and the process owner and that the QMS must be designed and implemented as an in-process and not an epi-process activity. To ensure a disciplined approach, senior management must ensure the true independence of the involved QA and QC functions.

[1]Quote from Wikipedia

References and Further Reading

Deming EW (1966) Some theory of sampling. Dover Publications, Mineola NY, USA

Imai M (1986) Kaizen: the key to Japanese competitive success. McGraw Hill, New York, USA

Lunau S et al (2006) Six sigma + lean toolset. Springer Verlag, Heidelberg, Germany

Spear SJ (2009) The high velocity edge. McGraw Hill, New York, USA

Future of Regulatory Safety Assessments

61

Gerd Bode and Petra Starck-Lantova

Contents

Abstract

Drug development contributes to improve health, duration, and quality of life. Lethal diseases have turned into chronic tolerable conditions, but medical need for many pathological processes continues. Concerns appear that in spite of extensive workload, success of pharmaceutical activity, and included facilitated access to novel drugs, may slow down. The preclinical testing via in vitro and animal experimentation reveals limitations to select the right promising candidates, most likely to be effective in humans and predict undesirable side effects early on.

Therefore, constant efforts are necessary to improve the strategies. Courage needs to be stimulated to leave traditional paths and find new and better ways. This "rethinking" process needs directions to focus on additional options: use of more in silico data, deeper insight via cell cultures or receptor studies, new methods to explore more intensively relevant mechanisms of diseases and pharmacodynamics,

G. Bode (✉)
Institute for Pharmacology and Toxicology, University Medical Center, University of Goettingen, Goettingen, Germany
e-mail: gerd-bode@t-online.de

P. Starck-Lantova
The University of Bonn/German Society for Regulatory Affairs, Bonn, Germany

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0_71

more comparative data from different animal models, which species really deliver signals relevant for patients; for this objective, disease models or implementation of human conditions into transgenic animals may be supportive. More rigorous randomized designs of preclinical studies and their blinded assessment may improve reproducible and therefore validated results.

In times of "big data" regulatory agencies, academic and industry researchers (possibly under political pressure) should feel obliged to stop selective publications (only positive effects) but create access also to options to learn from failures. The use of available knowledge (literature, experience, scientific advice) may limit the risks of reducing attrition rates and help to shorten timelines. Discussions with agencies have already facilitated a number of strategies. Examples are ICH guidelines M3 (allowing early access to new compounds for women of childbearing potential) or S 9 (reducing the preclinical development package for patients suffering from tumors).

The purpose of this chapter is to prompt openness and imagination to use new methods, more science, experience, and communication among researchers to the benefit of patients.

Introduction

The dualism of desirable effects and undesired reactions by chemical and pharmaceutical compounds on biological systems continues to be a fascinating and difficult phenomenon. Research, preclinical and clinical developments carry on to identify the main characteristics. Pharmacology detects new mechanisms for therapeutic purposes and thereby improves quality of life and survival, while toxicology and safety pharmacology submit their results to rigorous evaluations when extrapolation to humans takes place. The success is based on refinement of analytical methods, on paradigm change from morphology to inclusion of physiological functions, on better understanding of diseases and options for correcting such dysfunctions. In silico, in vitro, and in vivo studies support this process, but high-throughput screening does not seem to lead to higher success rates, just the opposite, the attrition rates slow down.

This chapter tries to recommend some options how regulatory safety assessments could try to counterbalance this growing weakness.

Review of History, Methods, Regulatory, and Industrial Environment

Patent expirations, fast-rising competition of developing countries, and an increasingly complicated regulatory environment has reduced the attrition rates. Further, growing cost-constraining healthcare systems reduce the potential of many new drug discoveries to generate revenues sufficient to cover the costs of development, which raised for an individual substance up to US $ 2.6 billion (Kaitin and DiMasi 2011).

The continuous pressure, fueled by public interest for effective and safe medicines, tightens regulatory requirements on the safety assurance before launch. On the other hand, payers enforce drug pricing limits and proofs of therapeutic or economic advantages over existing products. Biotechnological products can serve as an example for this trend. Due to their novelty and their substantial impact on disease progression, biotech products were originally well received in health care systems. Their acceptance, however, quickly became unsustainable when during the last few years, more than ten new products treating cancer were priced at over $100,000 per year per patient, along with a continuing stream of many new orphan drugs costing from $150,000 to $500,000 per year, hepatitis drugs at $80,000 per year, and others. This resulted in the public pressure to advocate for heavy price discounts and formal restrictions (Evens 2016).

As a counterbalance, pharmaceutical companies try to reduce their R&D costs by building strategic alliances with academic institutions, Contract Research Organizations (CROs), patient groups and other developers. Promising drug candidates are often acquired through mergers and acquisitions. This increases complexity,

fragmentation, obstructed communication and coordination among globally active partners and can result in insular solutions and addressing only part of a larger systemic problem. Mergers keep many pharmaceutical companies in a continuous state of reorganization; this does not favor stable research departments allowing to pursuing long-term scientific goals, like drug development, needing mostly more than 10 years of perseverance (Schueler and Buckley 2014; DiMasi et al. 2014).

No doubt, the tragedies of the past (Drug, Food and Cosmetics Act – USA, 1938 – due to mortalities of more than 100 patients after using sulfanilamide elixir with diethylene glycol as a solvent or thalidomide (Contergan) intake leading to phocomelia and other deformities) caused rigid legislations and requirements for animal studies to confirm safety before human treatment with new drugs.

Accordingly, the Kefauver Harris Amendment to the US Drug, Food and Cosmetics Act in 1962 defined the need to prove safety and efficacy of new pharmaceutics (mostly through animal testing) before their exposure to humans and later approval; the predecessor of the EU, the European Economic Community, reacted similarly three years later by introducing the Directive 65/65/EEC. Since then, public attention and expectations on medicines safety is rising and pharmaceutical manufacturers are required to perform safety tests on their new drugs and submit the data to supervisory organs before being allowed to market their products.

The characterization of new products starts today often with high-throughput screening (leading to a ten-fold reduction of the cost of testing compound libraries), with combinatory chemistry (increasing by 800-fold the number of new molecular entities to be potentially synthesized), or new generations of DNA sequencing (Scannel et al. 2012). However, although the turnover of tested substances had significantly increased, this enormous effort did not improve the final goal. Out of every 10,000 new molecular entities discovered, only one receives regulatory approval to be marketed, and the percentage of drugs entering clinical trials resulting in an approved medicine is estimated to less than 11.3% compared to 16.4% success rate ten years ago (tufts CSDD 2015; PhRMA 2015).

This failure rate can partially be explained by the fact that many companies investigate new molecules before they clarify the pathological mechanisms of the diseases they are trying to treat. A more successful strategy would, therefore, attempt to understand the pathophysiology and epidemiology of the disease as early as possible before embarking expensive development programs (pwc 2012).

Some discrepancies can also be stated when looking at the relationship of preclinical and clinical results. The US-American FDA published already 2004 a white paper that critically evaluated preclinical animal models, recognizing them as one of the reasons for the disconnection between increased expenditure in R&D and attrition rate in drug discovery (FDA 2004). It was suggested that the high attrition rates of clinical trials indicate a discrepancy between the promising studies in animal models documenting the efficacy of a drug, and the real effects of the drug in human trial subjects. The quality and translational value of nonclinical research, particularly animal studies, has been therefore questioned.

Two directives referring to animal studies stress different sides of this issue.

In the European Union (EU), the rationale and requirements for animal testing in the development of medicinal products for human use are defined in Directive 2001/**83/EC Annex I**, which states that: "*An integrated and critical assessment of the non-clinical evaluation of the medicinal product in animals/in vitro shall be required. . .*" (Annex I, Part I, Art. 2.4), *"Clinical trials must always be preceded by adequate pharmacological and toxicological tests, carried out on animals* . . ." (Annex I, Part I, Art. 5.2b) with the option that "*Studies in animals can be substituted by validated in vitro tests provided that the test results are of comparable quality and usefulness for the purpose of safety evaluation"* (Annex I, Part I, Art. 4.2.3f). Animal studies have therefore become a standard component of pharmaceutical R&D.

This fact is in contrast to Directive 2010/**63**/**EU** on the protection of animals used for scientific purposes. This Directive has taken full effect on 1 January 2013 and refers directly to the principles of **Three Rs** (Refinement, Reduction, and Replacement). According to this Directive *"the use of animals for scientific or educational purposes should only be considered where a non-animal alternative is unavailable"* (preamble 12) and *"Member States shall ensure that, wherever possible, a scientifically satisfactory method or testing strategy, not entailing the use of live animals, shall be used instead"* (Art. 4.1).

Over the past three decades, the preclinical safety evaluation paradigm has developed in two parallel branches. For new chemical entities, the general approach has provided common ground for evaluation across different product classes; for new biological entities, where classical toxicology was acknowledged to be less relevant, a more product specific approach evolved. This resulted in two guidelines, ICH M3(R2) and ICH S6(R1), both published fist 1997 in the Proceedings of the International Congresses on Harmonization (ICH) and then later revised, which is discernible by the letter R in the title for the recent guidelines.

The guideline ICH M3(R2) delivers practical recommendations for timing (when to conduct which safety studies) and conditions to include different patient population from male adults, over women without or with child bearing potential to pregnant women and children.

The guideline ICH S6(R1), on the other hand, stimulates to reconsider the traditional strategies in preclinical development. As one example, in case of recombinant proteins no studies on carcinogenicity and metabolism are required (because proteins are not metabolized to reactive species) and off-target effects are not expected (because of a high specificity of the recombinant proteins to the target). Another example is the assessment of potential effects on the cardiovascular, respiratory and nervous system (safety pharmacology): instead of stand-alone studies, these functions are recommended to incorporate into the pivotal chronic toxicity studies.

ICH S6(R1)emphasizes especially the selection of a relevant animal species, which are able to predict human reactions. The selection should be usually accomplished by an in-vitro comparison of binding affinity or functional activity of the product in human and animal cells, followed by in-vivo confirmation of the pharmacological activity or cross reactivity in that test species. In case of monoclonal antibodies, relevant species for testing are those that express the desired epitope and demonstrate a similar tissue cross-reactivity profile as for human tissues.

A number of flexible options should be considered: e.g. one relevant species may suffice, e.g., when only one relevant species can be identified, or where the biological activity of the biopharmaceutical is well understood. Or when no relevant species exists, the use of relevant transgenic or gene knock-out animals expressing the human receptor or homologous proteins could be chosen, no doubt, this option will prolong the evaluation process. In humanized mice, the comparability of pharmacodynamics in the animal model and humans is an important conclusion to consider the mouse as a suitable relevant model (van Meer et al. 2015).

The S 6(R1) guideline recognizes also animal models of disease as a relevant option. These models were originally used mainly to better understand the pharmacological action of the product, the pharmacokinetics and dosimetry. In all cases, the use of animal models of disease to support safety should be scientifically justified (ICH S6(R1).

Nevertheless, the guideline caused criticism by some researchers: e.g. the guideline missed the chance to catch up with scientific progress (critical review by Kooijman et al. 2012). Indeed, for example the option to use in-vitro alternative approaches is mentioned only briefly, even in the guideline's revised version S6(R1) (approved 2011): *"Although not discussed in this guidance, consideration should be given to the use of appropriate in-vitro alternative methods for safety evaluation. These methods, if accepted by all ICH regulatory authorities, can be used to replace current standard methods."* (Part II, Chap. 1.1).

It was therefore recommended (e.g., by Kooijman et al. (2012)) to make safety evaluations on a case-by-case basis, driven by product specific aspects such as the cause, mechanisms and reversibility of adverse effects. However, not much of experience and scientific expertise with these products could be gained in the meantime. In addition, the flexible case-by-case approach may lead to diverging interpretations and inconsistency of opinions between regulatory agencies.

The Tegenero case (monoclonal antibody) from 2006, leading unexpectedly to a cytokine storm inducing severe shock symptoms in volunteers, stimulated Agencies to reconsider their recommendations for the First in Man use of new compounds. The EMA guidance for first-in-man studies, "*Guideline on strategies to identify and mitigate risks for first-in-human clinical trials with investigational medicinal products*" (EMA 2017a) came in February 2018 into effect. Requested are a better integration of pharmacokinetic, pharmacodynamic data and toxicological findings into the overall risk assessment; nonclinical data should help to define the estimated therapeutic dose, the maximum dose, and dose steps and intervals. The plead includes stronger stress on using alternative methods and encourages to *use more in vitro studies whenever scientifically relevant and sufficiently validated.*" The "weight of evidence" should include a comparison with humans in regard to target expression, distribution and primary structure, pharmacodynamics, metabolism and other PK aspects, and on- and off-target binding affinities and receptor/ligand occupancy and kinetics. Nevertheless, the guideline warns that even a high degree of homology between the selected animal model and human, or a similar response in human and animal cells in vitro, does not necessarily imply comparable effects in vivo. "*For example, there might be differences in affinity of the new candidate for molecular targets, physiology differences in tissue distribution of the molecular target, cellular consequences of target binding, cellular regulatory mechanisms, metabolic pathways, or compensatory responses to an initial physiological perturbation*." In such situation understanding of the relevance of the animal models and their translational differences may be improved by using in vitro human cell systems or human-derived materials.

Animal Models

Animal models certainly play an important role for the overall assessments. The request is for new pharmaceuticals to include relevant models, relevant for the prediction of reactions in humans.

No animal model can fully reproduce all features of human diseases. And no human models (volunteers or Phase II patients) can predict all reactions possibly seen later under broad exposure of thousands of patients under differing life styles and co-morbidities. But animal models allow to gain early on important signals for any severe effects. But selecting the optimal model is not a trivial task. Despite the S7A ICH Guideline recommends that "*consideration should be given to the selection of relevant animal models or other test systems so that scientifically valid information can be derived*", the selection of animal species follows rather long-established practices and less scientifically justified deliberations.

A short reflexion about animal models may be helpful. Traditionally, over 90% of animals used in drug discovery are mice and rats. In drug development, the two species-testing is the rule. Rodent experiments should be completed with non-rodent species like dogs, non-human primates or minipigs. Non-human primates (NHPs) should only be used when the purpose of the study cannot be achieved by any other species (Article 8.1(b) Directive 2010/63/EU).

Mouse: Easily available; low cost; ease of handling; fast reproduction rate, important for reproductive toxicity studies. Many transgenic models.

Many well-established disease models created in mice, allowing both pharmacology and toxicology investigations. The limited blood volume can be overcome today by new sampling techniques and refined analytical methods, allowing microdosing studies.

Rats: Enormous historical background; important animal model for research in psychology and biomedical science, especially cancer research. Their advantage: good availability; larger body size; easy handling. Gene knockout and embryonic stem cell techniques are relatively more difficult in rats. The role of genes is easy to study: many inbred strains, all members are closely genetically identical.

Syrian hamsters: Less used today; few times for assessing the potential for cancinogenesis, metabolic diseases, non-cancer respiratory diseases, cardiovascular or infectious diseases; less easy to handle because of fighting of males.

R**abbits**: Particularly useful for assessing ocular and dermal irritation; primary non-rodent species for embryo-fetal developmental toxicity studies ever since the tragedy of thalidomide.

Dogs: Most frequently used non-rodent species in preclinical drug development; genetically diverse; a convenient model for many human diseases, e.g. the bone cancer osteosarcoma (Rowell et al. 2012). Dogs naturally develop beta-amyloid plaques (the protein's amino acid sequence in dogs is identical to humans), they show cognitive decline when growing older (Davis and Head 2014). Good model for complex neurocognitive disorders such as Alzheimer's disease. Dogs are expensive; need more sophisticated housing conditions and higher doses of an experimental drug. In addition, there is elevated public scrutiny and reluctance for the public and many evaluators.

Nonhuman primates (NHP): Use allowed when scientifically justified for safety testing. Most frequently used model to study potential adverse effects of monoclonal antibodies (mAbs). High public scrutiny. Target expression and function comparable to humans. Relatively large body size (allows repeated blood sampling), good availability of reagents, assays and methods (often adapted from humans), and generally good availability of animals are advantages of NHPs.

Potential limitations: high costs, limited group size, and often heterogeneous population with occasional background infections (like humans?). Testing often requires adult animals (4–5 years). NHPs inadequate for carcinogenicity testing and inconvenient for reproductive toxicity studies with their low fertility rate, high spontaneous abortion rate and long pre- and postnatal development times (Baumann et al. 2014). When testing mAbs, results may be confounded by anti-drug antibodies formation, leading to a neutralization of pharmacologic effect or a clearance of the mAbs from the circulation (Bussiere et al. 2009).

Chapman et al. (2012) demonstrated that in some therapeutic areas rodents can support biologic development and provide relevant data and could therefore reduce the use of NHP. In controversial discussion is the duration of repeat-dose studies: six months provide sufficient data and nine months or longer did not bring any additional benefit (Clarke et al. 2008). There is also a plead to use only two dose groups instead of standard three, leading equally to relevant data (Chapman et al. 2010).

Parallel to public demands, criticism of unnecessary or even uninformative use of NHPs can be heard from professional circles too. Van Meer et al. (2013) evaluated safety studies in NHPs for mAbs registered in the EU and concluded that NHPs have been used even when there were other pharmacologically-responsive species available and the testing was in some cases not informative. The authors could also show that pharmacology-mediated adverse effects of mAbs are highly predictive from in vitro studies.

Minipigs: Increasingly used for toxicity testing of pharmaceuticals, experience in Europe (Ganderup et al. 2012), also in USA and Japan. Increasing acceptability by regulatory agencies, e.g. FDA. Up today: mostly testing on small molecule-based therapeutics and dermal administration (Ganderup et al. 2012), but increasingly also for repeat-dose administration of biologics (reviewed in Baumann et al. 2014). Zheng et al. (2012) demonstrated on several human mAbs that they show low clearance, long half-life and low volume of distribution in minipigs and therefore good translation to humans. Also, according to Baumann et al. (2014), studies on tissue cross reactivity of biopharmaceuticals as well as safety pharmacology and fertility endpoints in repeat-dose studies can be carried out in minipigs.

Advantage of the minipigs: immune system with structures and functions largely analogous to the human immune system (Bode et al. 2010). Additionally, minipigs show less undue effects than dogs (Weaver et al. 2016). Also the complete genome (of a size as well as the number of genes comparable to humans) is known due to intensive data information based on research of house swines. Minipigs are not genetically transformed but the result of chronic breeding selection. Housing and handling easy with training.

Disadvantage: lack of placental transfer of macromolecules (Bode et al. 2010), which may limit their role in developmental toxicity testing of mAbs. In addition, rapid body weight gain, requiring more flexible testing strategies (shorter studies and use of younger lighter animals) and lack of published experience may be considered disadvantageous (Baumann et al. 2014).

Animal models of human disease: Primarily utilized to gain insight into the potential efficacy and mode of action of novel pharmaceuticals. Their value in understanding safety risks of compounds begins to be recognized (Morgan et al. 2013). Their use as part of a preclinical safety submission/dossier has been driven by the need to test a specific hypothesis and combine efficacy and safety evaluations; these models are even recommended by regulatory authorities (Cavagnaro and Lima 2015). Examples of the use of disease models include: infected animals to test the efficacy of a vaccine, mice inoculated with xenogenic (human) tumors expressing the target antigen, or genetically modified animals that develop spontaneous disease (Bussiere et al. 2009; Te Koppele and Witkamp 2008).

In general, animal models of human disease reflect rather simple mechanistic pathways while human diseases are mostly complicated by multifactorial pathological processes, often poorly understood. From a pathology perspective, the evaluation of animal disease models is challenging as the induced disease results in effects confounding safety assessment (Morgan et al. 2013). Historical data on spontaneous background finding are usually missing. Therefore, discerning whether the clinical and anatomic pathology findings are attributable to incidental age-related or background changes, tested agent, or disease manifestations require additional experience and background data need to be accumulated. Multigenerational studies or increased numbers of control animals may be necessary (Cavagnaro and Lima 2015; Morgan et al. 2013; Bussiere et al. 2009). First observations signal that lifespan of disease models may be limited, therefore, adequacy of such animals in regard to chronic experiments may be an issue (Cavagnaro and Lima 2015).

Further, animal models of disease have intrinsic variability and immutable genetic and species differences. These factors can complicate the interpretation of the data. Investigators should therefore carefully evaluate the results and keep in mind that over- or under-estimating of adverse side effects may be possible. Also, analyze the target behavior in the animals, for low molecular weight chemicals the metabolite profile, for recombinant proteins the pharmacological effect (e.g., activity, clearance, target expression, immune phenotype, and immunogenicity) (van Meer et al. 2015).

Transgenic animals: The most common models are gene-targeted or knock-out (KO) animals; they lack an endogenous gene and therefore fail to express the related protein(s). This property offers the chance to assess drug specificity, investigate mechanisms of toxicity, screen for mutagenic and carcinogenic activities of therapeutic candidates, or study target blockade by novel therapeutic candidates (Bussiere et al. 2009). KO-mice have been valuable to study obesity, heart disease, diabetes, arthritis, substance abuse, anxiety, aging and Parkinson disease (NIH 2009*)*.

The transfer of new genetic information can overexpress a target protein. The "humanized" knock-in animals with a human gene can evaluate the efficacy and toxicity of human biopharmaceuticals that are not pharmacologically active in normal rodents. Transgenic mice generated to carry cloned oncogenes and knockout mice lacking tumor suppressing genes have provided good models for studying risk of human cancer; but in spite of their recommendation by ICH S1A, their use is still limited for the assessing

the carcinogenic potential of new drugs (Friedrich and Olejniczak 2011) .

For testing recombinant proteins and cell therapy products compensatory mechanisms may take on the function of the absent protein(s) or target (e.g., induction of other calcium transporter genes in calbindin-D9k gene KO-mice described by Lee et al. 2007). Additionally, physiological effects of genetic mutations underlying diseases may differ in humans and in mice (Hirano et al. 2007).

The Controversy of the Animal Use in Pharmaceutical R&D

The selection of animal models should be based on their relevance for humans. In practice, the selection of species beyond those cited above is limited. There is a striking paucity of quantitative comparative data for animal models (Schein et al. 1970; Heywood 1981; Greaves et al. 2004; Matthews 2008). This makes any request using validated models and methods difficult. Moreover, literature offers only informative data, where animal models were positively contributing for Market Authorization. Data on failures and lack of success are not selected by researchers and editors for publication: reports about unacceptable adverse effects in animals are unattractive for journals; such data do not raise public interest; there could be also reasons for commercial confidentiality (Matthews 2008). Such data are thus stored in internal databases of pharmaceutical companies or research organizations and forlorn for public research.

This dilemma was addressed by Olson et al. in 2000: they compiled a survey of 150 compounds which revealed to be toxic under clinical test conditions in humans. There was a true positive human toxicity concordance rate of 71% for rodent and non-rodent species, with non-rodents alone being predictive for 63% of human toxicity and rodents alone for 43%. The authors appraised safety testing on (healthy) animals as significantly beneficial. But Matthews (2008) criticizes Olson's analysis for being inconclusive or even misleading because the authors did not attempt to estimate the corresponding specificity (true negative rate) and sensitivity (true positive rate) without which it is impossible to assess the evidential weight provided by the animal models. Van Meer (2013) contributes to this discouraging interpretation and predicts that this poor outcome increases when looking at highly complex and species specific protein drugs, which are usually immunogenic in animals.

Perrin (2014) fortifies this issue by reporting that more than 80% of animal studies on safety and efficiency of potential therapeutics fail to predict the desired success rate in patients. Similar data by Hay et al. (2014) on success rates of 835 drug developers show that the proportion of therapies advancing from Phase 1 to regulatory approval is only around 10%. Bailey et al. (2014) analyzed datasets of 2366 drugs with both animal (rat, mouse and rabbit as preclinical species) and human data. The authors concluded that the absence of toxicity in the animals provided little or virtually no evidential signals for the lack of adverse drug reactions in humans.; a (re-)analysis of data specific for dogs from the same original dataset reinforced recent criticism that dogs are used mainly for historical instead of scientific reasons (Bailey et al. 2013): No evidence appeared, that canine data would predict efficacy and toxicology of medical compounds in clinical trials; they suggested that alternative methods are urgently required (Bailey et al. 2013).

Finally, van Meer et al. (2012), when focusing on post-marketing data, confirmed that animal data were not predictive for detecting serious adverse drug reactions in patients. Because 63% of adverse drug reactions had no counterparts in animals and less than 20% of serious reactions had an actual positive corollary in animal studies the authors conclude that animal safety studies in their current form should not be included in prospective pharmacovigilance studies.

Is there any chance to overcome this dilemma?

Kooijman (2013) explains the persistent use of animal studies in drug development by inertia of the system with animal studies embedded in a set of institutions (i.e., regulations, norms and values) that are taken for granted, normatively endorsed, and backed up by regulatory authorities. This is the motivation why the industry stays reluctant to move away from established conservative models. Although the standard animal models

are as the result of guidelines by definition not binding, and it is possible to provide preclinical data from own scientifically justified experimental programs adapted to a specific product, but this option is in practice not frequently used. This reluctance is the result of fears about possible delays or even failures of the marketing authorization process. Despite of some official support for progressive trend-setting new trends, skepticism prevails towards new strategies and conservatism dominates in industry.

Admittedly the situation of regulators is challenging. On one hand they try to promote innovations and recognize rapid growth in knowledge and technologies (Cavagnaro and Lima 2015), on the other hand they have to protect patients from risks.

Van Meer (2013) attacks both the industry and the regulators by saying:

> *The adoption of the precautionary principle by the regulatory authorities and the relative ease with which this burden of proof is accepted by the pharmaceutical industry – without attempts to improve the current paradigm – has created a stalemate in which animal studies, predictive or not, continue to exist with little room for innovation.*

Therefore, all stakeholders should critically rethink their developmental strategies and should be encouraged to implement new technologies that predict the safety and efficacy of therapeutics better than current animal studies do.

Suggestions for Improvement

To break successfully with long-term traditions is only possible when academic and industrial researchers and developers cooperate openly with regulators and transmit their innovative thoughts into new guidelines and/or good practices. Some options will be illustrated in the following sections.

Consider Entering the Clinic Without Animal Studies

Biosimilars are taken as an example. The option to circumvent in-vivo preclinical studies has been recently recommended in revised guidelines on biosimilar medicinal products (EMA 2012, 2014).

In contrast to generic medicinal products, nonclinical data, including animal studies, have been traditionally requested for marketing authorization of biosimilars. Nonetheless, this praxis has been recently abandoned by European regulators, because animal models turned out not to be sensitive enough to provide sufficient information on pharmaceutical similarity of these products (EMA 2012, 2014). The guidelines on biosimilars acknowledge that "*in-vitro assays may often be more specific and sensitive to detect differences between the biosimilar and the reference product than studies in animal*" and, therefore, "*these assays can be considered as paramount for the non-clinical biosimilar comparability exercise*." Therefore, in-vivo testing should no longer be performed by default and its necessity should be considered on a case-by-case basis in a stepwise approach where the extent and nature of the development program depends on the level of evidence obtained in the previous step(s). This regulatory decision is regarded as revolutionary, it opens new ways for pharmaceutical developments with no new animal testing at all, and it implies that regulators may even discourage developers from performing such studies (van Aerts et al. 2014).

This applies especially to highly specific mAbs, where only NHPs are pharmacologically responsive. Therefore, as toxicological studies in NHPs have notably small group sizes, their conduct has been explicitly not recommended for biosimilars. In situation, when no relevant in-vivo animal model is available, the guidance leaves the option to proceed directly to human studies while applying principles to mitigate any potential risk (EMA 2012, 2014).

The recommended step-wise approach should proceed as follows: after physicochemical and biological characterization of the product, pharmacodynamic comparability should be evaluated in in-vitro assays. Assays using human cells or human receptors can be used to assess binding to the target and the subsequent functional effects. Pharmacokinetic comparability can then be best evaluated directly in clinical studies. When close similarity of the biological and its reference product can be demonstrated, it is highly unlikely that

new safety issues different from the reference product, with the exception of immunogenicity issues, would arise. For immunogenicity, animal studies have no predictive value anyway. Only after performing this biosimilarity exercise, it should be determined whether additional in-vivo non-clinical work is deemed necessary (EMA 2012, 2014; van Aerts et al. 2014).

Reduce the Need of Animal Studies by Gaining Information in Exploratory Clinical Trials

Exploratory clinical trials are an approach described in ICH M3(R) guideline, which recognizes that in some cases early access to human data can provide valuable information on human physiology/pharmacology, on drug candidate behavior, and on therapeutic target relevance to disease. Such data can reduce the need of information gained in animal studies. Central to this approach is the concept that "the best model for man is man." Exploratory clinical trials are conducted in early Phase I (sometimes called Phase 0), have no therapeutic intent, are not intended to explore clinical tolerability, and can be conducted on patients or healthy individuals (ICH M3R). Their advantage is that they may give information on exposure and allow early comparison of kinetic/metabolic data between animal models and humans. They certainly help very early on to prioritize compounds when several candidates are available; these aspects again help to reduce animal usage compared to traditional development.

ICH M3(R) recommends also several approaches based on applying micro- or subtherapeutic doses. **Microdosing** (most often: single microdose of 100 μg) is a method assessing the basic behavior of drugs by applying small doses directly to human volunteers. The doses are well below those expected to produce whole-body effects but high enough to allow the cellular response to be studied. A candidate drug is labeled by radiocarbon isotopes and extremely sensitive analytical methods (mostly positron emission tomography (PET) and accelerator mass spectrometry (AMS)) are used for its biochemical quantitation. AMS is used for determining PK data by taking blood samples over time, processing the samples in the laboratory, and then analyzing their drug content. PET provides primarily PD data through real-time imaging and some limited PK data. The method provides important information about pharmacokinetics and pharmacodynamics, but it does not reveal information about toxicity or toxicology. Those endpoints will be addressed by supportive rigidly reduced conventional study designs.

Eliminating less promising molecules saves costs, resources, animals, and time. It avoids unnecessary exposure of the participants in clinical trials. Because the trials mostly involve a single dose administration (usually 1×100 μg, the alternative is 5×100 μg), the method poses very little risk of human toxic side effects (low dose and short duration of exposure). Very limited number of subjects is usually involved. Further, preclinical safety package required by authorities can be smaller as compared to the traditional Phase I studies, less animals are needed, and also only small quantity of the test drug is required. Other valuable advantages of microdosing studies are that they help to establish a likely pharmacological dose and select the first dose for the subsequent Phase I studies. A limitation of the method is shortage of data that exemplify whether the body's reaction to a particular compound is similar when applied as microdose or in its pharmacologically active dose (Tiwari 2014).

Use Alternative Approaches

Alternative models should be more efficient and provide additional information to supplement the results from traditional animal models. Although animal models are still often considered to be a "gold standard," they have never undergone validation to the same extent as non-animal technologies.

The need for improvement is recognized by Agencies. There is a new *Regulatory guidance on alternative/3Rs testing* approaches in discussion: "Guideline on the principles of regulatory

acceptance of 3Rs (replacement, reduction, refinement) testing approaches" (EMA 2016a) and related reflection papers (still in the form of drafts – EMA 2016b, c). The guideline provides information on the scientific and technical criteria for regulatory acceptance of alternative/3Rs testing approaches and encourages stakeholders and authorities to initiate, support, and accept development and use of such approaches. The reflection papers summarize the main animal tests required for the regulatory testing of medicinal products and presents opportunities for limiting the use of animals.

The guideline recommends the following criteria: availability of test methodology, test protocols with clearly defined and scientifically sound endpoints; relevance of the test for a particular purpose and accuracy/extent to which the test correctly measures the biological effect of interest; robustness of the test (i.e., reproducibility of the test results); a comparison with existing methods; and a description of circumstances under which the 3Rs testing approach is/is not applicable.

The reflection paper on opportunities for implementation of the 3Rs during regulatory testing of medicinal products for human use provides an overview of options to limit or completely skip the use of animal studies in nonclinical evaluation of drug substances. The paper also clearly indicates that the 3R approach is in the state of dynamic development and there will be more options coming. It is, however, already clear that, for example, toxicity evaluation will change. Traditionally, repeat dose toxicology studies follow a standard design and in rodents and nonrodents yield information on general characteristics of the toxicity, the target organs of toxicity, the dose–response (curve) for each toxicity endpoint, responses to toxic metabolites formed in the organism, delayed responses, cumulative effects, the margin between toxic and nontoxic dose, information on reversibility/irreversibility of the effect, and NOAEL (no observed adverse effect level), NOEL (no observed effect level) for toxicity (EMA 2008, 2010; ICH M(R2)). In contrast to this standard approach, the reflection paper (EMA 2016b) concedes the option to perform the tests on one species only ("*on a case by case approach, and if clearly justified*") and to possibly omit a study on reversibility of compound-related effects.

Also changes for safety evaluation paradigms are recommended; for instance, in vivo genotoxicity can be assessed by integrating this endpoint into repeated dose toxicity studies, usually of 4 weeks duration. The reflection paper recommends a standard test battery (in-vitro tests plus in vivo genotoxicity integrated in repeated dose toxicity study) without the isolated single in-vivo study. Likewise, carcinogenicity and reproductive toxicity test requirements are currently under revision with the aim to induce new testing paradigms based on a more comprehensive weight-of-evidence approach and potential to replace in-vivo studies or not doing them at all (Bode and Van der Laan 2016). "Core battery" tests for safety pharmacology could also be integrated in repeated dose toxicity studies. And a variety of tests aiming at manufacture, characterization, and control of the drug substance should be primarily performed in-vitro unless thoroughly justified. Other more specific examples of the recommended 3Rs approaches involve avoiding physiological distribution test of radiopharmaceutical preparations as required by the Ph.Eur., using duck cells rather than live animals when testing plasma derived hepatitis B vaccine, or discouraging from using animals for potency testing of investigational, or biological medicinal products.

Alternatives to animal testing (called 3Rs testing approaches in EMA's guidelines) are being developed for – besides ethical reasons – their time efficiency, less man power required, and cost effectiveness. Two most important approaches involve in-vitro cell culture techniques and in-silico computer simulations. These two approaches are then combined in another method known under the name "organs on a chip." Also microdosing described above can be considered as an alternative method. All these approaches do not replace animals completely; however, they help to significantly reduce animal numbers needed.

There is knowledge that alternative methods also have their specific advantages and

drawbacks. For example, cell cultures are criticized for not providing enough information about the complex interactions of living systems, computer simulations for using data from prior animal experiments, and microdosing for not revealing information about toxicity or toxicology. Thorough knowledge of the strengths and limitations of one's model is therefore crucial for its appropriate use and interpretation of results.

For cells and tissue cultures, in vitro tests are recommended in **OECD guidelines**. Examples are:

TG 428	Skin absorption: in vitro method
TG 430	In vitro skin corrosion: transcutaneous electrical resistance (TER)
TG 431	In vitro skin corrosion: human skin model test
TG 432	In vitro 3 T3 NRU phototoxicity test
TG 437	Bovine corneal opacity and permeability test method for identifying ocular corrosives and severe irritants
TG 438	Isolated chicken eye test method for identifying ocular corrosives and severe irritants
TG 439	In vitro skin irritation: reconstructed human epidermis (RhE) test method

Various types of cultures like cell culture, callus culture, tissue culture, organ culture, or separated cellular components are used for various purposes. For instance, for safety testing, bovine corneal organ culture can replace rabbits eye irritancy test, or models of human skin derived from cultured human skin (Corrositex®, EPISKIN™, EpiDerm™) can replace animal-based skin irritative and corrosive studies. Test systems based on the activation of human monocytes or monocytoid cell lines have been developed that take advantage of the role of these cells in the fever response and can replace rabbit pyrogen test. Similarly, mouse fibroblast (3 T3) and normal human keratinocyte (NHK) cells can be used in basal cytotoxicity test (e.g., phototoxicity) and support to determine the starting dose for the acute oral systemic toxicity test method and thereby reducing overall animal use requirements (NTP 2017). Cell cultures are further used to measure the rate of chemical absorption by the skin or phototoxic reactions and cultured cells have been developed to create monoclonal antibodies (Hester et al. 2006; Doke and Dhawale 2015).

Another example is represented by **tissue models**. For example, in-vitro metabolism studies have traditionally involved cells cultured into monolayers. However, because the interactions of cells with their surrounding environment can greatly affect shape, cell function and gene expression, two-dimensional or three-dimensional models have been developed. These models are supposed to better mimic mechanisms such as cell-to-cell adhesion and resistance to drug-induced apoptosis. Among the 3D-tissue reconstruction models are models of epidermis, full-thickness skin models, respiratory epithelia, keratinocyte eye cornea, vaginal epithelia, oral epithelia, and even models of the blood–brain barrier or three-dimensional models such as placenta, lymph node, and liver (Liebsch et al. 2011).

Organs on a Chip

Organ on a chip is a multichannel 3D microfluidic cell culture chip that simulates to some extent the activities, mechanics, and physiological response of entire organs. The chip is formed by small chambers containing a sample of tissue from a particular organ. When nutrients, air, blood and test compounds, such as experimental drugs, are pumped through the chambers, the cells replicate some of the key functions of that organ, just as they do in the body. By recapitulating the multicellular architectures, tissue-tissue interfaces, physicochemical microenvironments, and vascular perfusion of the body, these devices produce levels of tissue and organ functionality not possible with conventional 2D or 3D-culture systems. Biochemical, genetic, and metabolic activities of the cells are then measured by sensors and transferred for computer analysis. In the context of drug discovery and development, this technology is valuable for the study of molecular mechanisms of action, prioritization of lead candidates, toxicity testing,

and biomarker identification (Bhatia and Ingberg 2014; Prot and Leclerc 2012).

Validation of Alternative Methods

Alternative approaches may produce relevant and reliable results, but all new methods must be confirmed as suitable for its scientific **and** regulatory purpose. These methods are used routinely and repeatedly; they should be acceptable across countries; formal validation is a necessity. Therefore, it is recommended to involve regulators already in the process of definition of performance standards. Such cooperation will facilitate regulatory acceptance and help to implement new test methods (Liebsch et al. 2011).

The need of new alternative and validated methods is expressed in the EU legislation. Directive 2010/63/EU describes the coordination of formal validation studies at EU level to facilitate rapid uptake of new methods and approaches to replace reliance on animal testing as one of its key tasks. For this purpose, the European Union Reference Laboratory for alternatives to animal testing, **EURL ECVAM**, was established by this Directive. Through its network of laboratories (EU-NETVAL, European Union Network of Laboratories for the Validation of Alternative Methods), EURL ECVAM focuses on the validation of 3Rs methods for safety testing and efficacy/potency testing of chemicals, biologicals, and vaccines. It offers to research laboratories to scientifically validate alternative methods to animal testing. Through a dialogue with the stakeholder community and provision of information systems (DataBase service on Alternative Methods, DB-ALM, QSAR Model database and TSAR tracking system on alternative methods), EURL ECVAM further promotes the use and acceptance of new alternative methods in industry, academia, and by regulators. Examples are non-animal approaches for skin sensitization (allergy) testing, or co-developing two new (VICH) guidelines for the reduction of animal tests for the quality control of veterinary vaccines (EURL ECVAM 2017).

Method validation relates mainly to:

1. The repeatability and reproducibility of results obtained
2. The test's relevance for measuring or predicting relevant biological effects

Validity assessment can include general knowledge of the method, the scientific principles on which it is based, historical data from using the method, and the use of pilot studies (when using in vivo methods) with smaller numbers of animals before embarking on a full scale study (European Commission 2016). The formal validation process involves multiple phases including preparatory method refinement, small-scale transfer studies, and finally large-scale international collaborative studies with manufacturers and national control laboratories (EMA 2016a). Alternatively, testing approaches that have sufficiently demonstrated their scientific validity according to the criteria described but have not been assessed in a formal validation process can also be evaluated on a case-by-case basis by the competent authorities (EMA 2016a).

Minimalize Bias in Experimental Data and Mind Good Research Practices

Animal studies can elucidate normal biology and improve the understanding for the pathogenesis of a disease, a deficiency often appearing when developing therapeutic interventions. However, animal studies produce insights only if tests are carefully designed, critically interpreted, and thoroughly reported. These quality features amplify good laboratory practices (GLP), compliant to which many animal studies (e.g., safety studies) should be performed. GLP ensures traceability and uniform, reproducible quality, but it does not guarantee the quality of the animal model or scientific valuable interpretation of the outcome for human purposes.

The lack of methodological rigor in preclinical studies acts as a barrier to translation of research findings and represents a major source of reduced attrition rates in drug development (Glasziou

2014; Green 2015). The Lancet published 2009 a review on the production and reporting of biomedical research in which it was calculated that 85% of basic and clinical research is wasted because of inadequate or inappropriate design, nonpublication, and poor reporting (Chalmers and Glasziou 2009). This represents an estimated annual loss of over $100 billion research funding. Clinical trials erroneously based on poorly conducted preclinical safety studies may lead to unnecessary exposure of trial participants to potentially harmful agents or to prevent them from participating in other trials with possibly effective products (Landis et al. 2012).

Particularly widespread are deficiencies in reporting key methodological parameters and poor experimental designs, both correlating with overstated findings (Landis et al. 2012; Gulin et al. 2015). Scientists from hematology and oncology department at the biotechnology firm Amgen (Begley and Ellis 2012) tried to confirm published findings related to their work and despite efforts to avoid technical differences they could confirm scientific findings in only 11% of cases. Reproducible studies were mainly those, in which authors had paid close attention to controls, reagents, investigator bias and describing complete data set. In the other cases, results could not be reproduced, the data were not routinely analyzed by investigators blinded to the experimental versus control groups and/or only selected experimental results supporting an underlying hypothesis were presented (Begley and Ellis 2012). Corresponding results were reported by Bayer HealthCare who could validate only about 25% of published preclinical studies (Prinz et al. 2011).

The recognition grows, that the use of techniques that assess the impact of publication and study-quality biases on estimates of efficacy in animal experiments is necessary (Sena et al. 2007). An adoption of newly and better defined quality standards would lead to improved effectiveness and efficiency in the selection of promising candidate drugs.

There are a number of sources of experimental bias which reduce the quality of the research.

Bias from Poor Reporting

Reporting details of a study including methods of statistical analyses used, sample sizes, inclusion/exclusion criteria, methods of randomization, blinding, gender, strain, species selection, and age of animals is essential to avoid publication bias, assist replication, and justify the research. Meanwhile, several guidelines have been issued to improve poor reporting, among them the ARRIVE guidelines ("Animal Research: Reporting of In-Vivo Experiments," 2010), the GSPC ("Gold Standard Publication Checklist," 2011), or the checklist of the Nature Journal (2013). Although the guidelines list suggestions for improved reporting, lack of pressure to apply these suggestions and report comprehensively and uniformly leads to noticeably inconsistency, obstructing correct assessment of reported results (Green 2015). Here, journal editors and regulators/assessors of clinical trial applications can support improvements considerably.

Bias from Nonpublication

Selective reporting is another reason for the lack of translation from basic research to the clinical situation. An increasing number of studies demonstrate publication bias that only about 50% of animal research results are published. The main motivation seems to be the lack of statistical significance as there is relatively little incentive for journals to publish negative, non-novel, or repeated findings (Korevaar et al. 2011; Sena et al. 2010; Ter Riet et al. 2012; Tsilidis et al. 2013). Nonpublication causes unnecessary duplication of research and poses a serious problem for performing valid literature syntheses. There should be a plead to publish **all** results regardless of whether the outcomes are positive or negative; all studies (equivalent to existing registers of clinical trials) should be registered in professional circles (Kimmelman and Anderson 2012). Registration of animal trials would impede retrospective changes of endpoints and study protocols and not publishing negative or unfavorable results. And today such publications of failures should be easy since safety studies are always done under GLP conditions and collecting data and archiving is performed on local computers.

Bias from Using Inappropriate Animal Model

Prestige, economy, convenience, and poor awareness of the translation of basic research into medical practice influences decisions on animal studies more than scientific rigor and patient need (Green 2015). For instance, laboratory mice are disproportionately more often used than any other animal species (JAXmice® alone stocks tens of thousands of types of strains of mouse models to choose from), and the common practice of using inbred rodent strains completely ignores genetic variation of target populations. Bennani (2012) points out that for some conditions (e.g., influenza, bacterial, and fungal infections, measuring CVD and LDL and simple blood chemistry) animal models are more reliable predictors, whereas for other diseases (oncology, immunology, psychiatry, HIV, etc.), animal models are to large extent nonpredictive of clinical outcome. The importance of selecting the best possible animal model should be therefore not underestimated.

Bias from the Regulation of Animal Research

Regulatory agencies require sometimes preclinical investigations that use animal models known to have no predictive value. Among such problematic disease areas are oncology, immunology, or diseases of the central nervous system (Bennani 2012). In addition, compliance with the 3Rs and animal welfare are in many countries controlled by veterinary inspectors and ethic committees. The assessment process is however neither open nor transparent and relies on individual opinions of the experts (Green 2015). Pressure to apply the 3Rs principles may be overstretched; it may reduce the statistical power of experiments under meaningful values. Scott et al. (2008) demonstrated that the failure of murine amyotrophic lateral sclerosis treatments to translate to the clinic was due to small group size numbers and underpowered experiments.

Many good research principles are actually long known, it seems that they sometimes get forgotten in the complex process as R&D of new medicines. Awareness of quality guidelines for biomedical research should therefore reanimated, examples are, for example, good research practice system of the World Health Organization (2006), guidelines published by the Research Quality Association in the UK (2008), or the Quality Assurance Toolkit developed at the University of Minnesota, USA (Michelson Prize and Grants 2014). Few details should be stressed.

Planning an Experimental Protocol

The methodological quality of an animal study starts with preparing a detailed experimental protocol. Checklist of factors listed, for example, in the ARRIVE guidelines, can be meaningful. Variations in the experiments must be considered and outlined in the protocol. Results from control animals need to be known and interpretation should benefit from these historical data. Study directors should seek consultancy from interdisciplinary interactions of the primary investigative team with experts in ancillary disciplines (statistics, laboratory animal science, pathology, etc.) and include the data generation and collection process (Everitt 2015). The experimental hypothesis to be tested must be well explained and defined as well as the experimental aims, design, and endpoints.

Recognizing Sources of Variation

Sources of variation can include inherent factors of the animal (e.g., stock/strain/substrain, source, sex, age, weight, source, pathogen status, etc.) as well as the animal facility environment (diet, bedding, housing, water delivery, lighting, noise, vibration, temperature, humidity, etc.). For this reason, harmonization of international standards for animal care would already reduce one important source of internal variation. Other factors are the methods used, dose form and timing of dose administration, types and preparation of excipients and vehicles, blood and tissue sampling sites and methods, handling of subjects, etc. (Everitt 2015). Example is the significant difference in the serum hepatic enzyme, alanine transaminase, which can occur if mice are handled by the body instead of the tail (Swaim et al. 1985). Similarly, significant differences have been reported in research endpoints, such as cytokine

concentrations, depending on method/site of blood removal (Mella et al. 2014).

Collecting Data

Although usually inbred rodent strains with minimal genetic differences are used, data show how important randomization is. For instance, in a systematic review of hypothermia in experimental stroke, nonrandomized studies overstated the reduction in infarct volume by 27% and studies without blinded outcome assessment overstated efficacy by 19% when they were compared to randomized and blinded studies, respectively (van der Worp et al. 2005). To minimize bias resulting from internal variation in the data, following steps should be always taken:

- **Randomization**: Animals should be assigned randomly to the various experimental groups and the method of randomization reported. Information on the allocation, treatment and handling of animals across study groups, the selection and source of control animals, including whether they are true littermates of the test groups should be provided. Data should be collected and processed randomly or appropriately blocked.
- **Blinding:** The investigator should be unaware of the group to which the next animal taken from a cage will be allocated (allocation concealment). Animal caretakers and investigators conducting the experiments should be blinded to the allocation sequence (blinded conduct of the experiment). Investigators assessing, measuring, or quantifying experimental outcomes should be blinded to the intervention (blinded assessment of outcome). This may hold true for all instances of the experiment, including also post-mortal investigations like macroscopical and pathohistological inspections and assessments.
- **Sample-size estimation**: Underpowered experiments with low predictive value may either falsely conclude that interventions are without efficacy or provide falsely positive results leading to needless subsequent studies building upon the incorrect results. Too large studies will be unnecessarily costly. Both cases mean wasted resources in terms of time, money, and animals. An appropriate sample size should be therefore computed when the study is being designed and the statistical method of computation reported, which would also provide some assurance that sample size has not been increased incrementally in the light of ongoing analyses. Statistical methods that take into account multiple evaluations of the data should be used when an interim evaluation is carried out (Sena et al. 2007; Landis et al. 2012).
- **Data handling**: Rules for stopping data collection should be defined in advance. Also criteria for inclusion and exclusion of data should be established prospectively. How outliers will be defined and handled should be decided when the experiment is being designed, and any data removed before analysis should be reported. The primary endpoint should be prospectively selected. If multiple endpoints are to be assessed, then appropriate statistical corrections should be applied. Pseudoreplicate issues need to be considered before determining study design and analysis. For example, when analyzing effects of pollutants on reproductive health, multiple sampling from a litter, regardless of how many littermates are quantified, provides data from only a single biologic replicate. Investigators should also report how often a particular experiment was performed and whether results were substantiated by repetition under a range of conditions. Additionally, it should not be forgotten that a significant result does not provide information on the magnitude of the effect and thus does not necessarily mean that the effect is robust and highly reproducible (Landis et al. 2012).
- **Fighting experimental noise sound**: Irrelevant animals like those that die for reasons unrelated to disease (such as mishandling) should not be counted in results. Reasons for exclusion should be well documented. Whenever possible, numbers of males and females should be balanced because they can show sex-dependent differences in symptoms that

obscure modest drug effects. Littermates should be splitted among experimental groups (Perrin 2014).

- **Retrospective primary end-point selection**: Selection of a primary end-point only after data have been analyzed inflates the type-I error (false-positive results). This can be avoided by specifying a primary end point before the study is undertaken, the time(s) at which the end point will be assessed, and the method(s) of analysis. Significant findings for secondary end-points can and should be reported but should be delineated as exploratory in nature.
- **Reporting of individual data**: Nonrodent data are usually reported and interpreted on the basis of individual observations, reactions, and results. With rodent data using considerable higher numbers of animals, the statistical results often prevail and rare individual reactions get lost. The rule should be that preclinical investigations are handled like clinical results: Individual by individual, and not as a group mean. Hereby, possibly human relevant, but rare reactions do not get lost.
- **Avoid publication bias. Register all experiments. Use systematic reviews**

Systematic review (SR) (Sanderscock and Roberts 2002) is a simple technique developed to provide summary information by combining results from different sources and to make judgments on possible translation into clinical trials. In contrast to a narrative review which has no standardized methodology, the SR is a type of review that is structured, thorough, and transparent. Performing such appraisal can save resources and improve safety for participants in clinical trials achieved (van Lujik et al. 2014; Ritskes-Hoitinga et al. 2014; Vesterinen et al. 2014). Examples of the use of SR include, for instance, the study of Horn et al. (2001), who found no evidence to justify the start of clinical trials of nimodipine for focal cerebral ischemia in humans. The study emerged, however, only after 7665 patients participated in clinical trials. Comparably, Pound et al. (2004) demonstrated that drug side effects (in this case, excess risk of intracranial hemorrhage after thrombolysis treatment for acute stroke) found during a clinical trial could have been identified beforehand if a SR of preclinical animal studies had been performed.

When performing a SR, it is important to evaluate the quality of data collected by other researches. Its relevance can be illustrated on Alzheimer's disease, a condition which is despite of decades of experimental research known for a lack of effective disease modifying interventions. Egan et al. (2016) performed a SR and a meta-analysis of interventions tested in transgenic mouse model of the disease and after analyzing 427 publications describing 357 interventions in 55 transgenic models, involving 11,118 animals in 838 experiments, the authors found that the quality of these experiments was relatively poor – less than one in four publications reported blinded assessment of outcome or random allocation to group and no study reported a sample size calculation. Additionally, "trim and fill" analyses suggested that one in seven pathological and neurobehavioral experiments remained unpublished.

Likewise, Tsilidis et al. (2013) evaluated 4445 animal studies or 160 candidate treatments of neurological disorders and observed that 1719 of them had a "positive" result, whereas only 919 studies would a priori be expected to have such a result. From these 160 treatments, only 8 should have been subsequently tested in humans. These examples illustrate not only historical methodological weaknesses in preclinical animal testing but also insufficient critical appraisal of existing animal data before starting clinical research. Considering ethical issues and enormous financial costs related to clinical trials this is a rather alarming finding.

The bias resulting from not publishing could be significantly reduced by registering all experiments in a system similar to the one established for clinical trials. In this way, negative data would be published, unnecessary duplication of experiments would be prevented, investigators would receive credit for their work done and those seeking to summarize what is known would have access to all relevant data. The registration to the system could be flexible, with information

embargoed for a time to protect intellectual property (Macleod 2011).

To facilitate assessment of data collected, and to point out critical factors several study-quality checklists have been proposed. Among these are the CAMARADES checklist (Collaborative Approach to Meta-Analysis and Review of Animal Data in Experimental Stroke), Macleod et al. 2004), Stroke Therapy Academic Industry Roundtable (STAIR 1999), Amsterdam criteria (Horn et al. 2001), Utrecht criteria (van der Worp et al. 2005), ARRIVE Guidelines (Animal Research: Reporting of In Vivo Experiments), Kilkenny et al. 2010), and the "Guidance for the Description of Animal Research in Scientific Publications" (National Research Council [US] Institute for Laboratory Animal Research 2011). Factors itemized on the checklists are, e.g., publication in peer-reviewed journal, assessment of functional and histological outcome, replication in two laboratories, testing both males and females, behavioral outcome measured for at least 1 month, assessment made in acute and chronic phase, randomization of treatment or control, blinded assessment of outcome, sample-size calculation before start of an experiment, and others.

Conclusions and Outlooks

Preclinical development and especially here animal studies have been identified as possible factors, responsible for the insufficient efficiency of nonclinical pharmaceutical R&D. Quantitative analyses of publicly available animal toxicity studies revealed that their results were inconsistent predictors of undesirable or toxic responses in humans. There is a lack of powerful data and sometimes only a poor basis available for deciding whether a compound should proceed to clinical testing (e.g., Bailey et al. 2014). The selection and justification of the studies is frequently based on regulatory principles 50 years old. The way how they are designed and performed is in many cases decided rather on habit and tradition than on modern, scientifically sound justifications. Yet, although new approaches and technologies are being developed in a fast pace, their integration into drug development is rather slow.

But progress and optimization is required and can be achieved by opening and accepting new pathways: better animal models are needed, and better predictive non-animal models required. Provided that fit-for-purpose animal models are used and the design and execution of the testing is implemented according to stringent quality criteria in vitro and in vivo experiments can provide valuable information for the clinical performance of the drug. Unfortunately, besides a few exceptions, like development of humanized experimental animals, investment in development of more predictive animal models has been during the last decades considerably lower than in development of new technologies in areas such as molecular biology or clinical trial biomarkers (Denayer et al. 2014).

Better predictive state-of-the-art in vitro assays and in silico data, applied during early stages of drug discovery, can facilitate the long-term process of drug development. Replacing current acute and selected chronic in vivo regulatory toxicology studies by validated in vitro replacements would result in reduced animal use in pharmaceutical development of individual compounds. Following such strategy can be already observed in, e.g., some OECD test guidelines. The guideline no. 404 (Acute Dermal Irritation/Corrosion) recommends the conduct of in vitro assays (EOECD TG 430, 431) to limit the severity of toxicity for compounds that progress to in vivo evaluation.

The concept of the 3Rs exists since almost 60 years but its value has been mostly perceived only as a European regulatory issue (Chapman et al. 2013). Recognition is growing during recent years that there are benefits for improving the quality of research and reducing costs. The quality, reliability, and predictive value of many well established methods have not sufficiently been validated, but to abandon them is associated with insecurity and reduced trust in the view of remaining data. Lack of confidence in novel approaches is rooted in limited experience among researchers and Agencies and additionally in the lack of historical data. There is fear to change long established pathways, which have been successful

in the past. Joined effort of all parties involved is therefore needed to achieve progress and better acceptance of new approaches. Sharing knowledge (positive as well as negative experience) among stake-holders would facilitate the selection of the most promising methods. Such cooperation would fasten the transition towards novel approaches and reveal gaps for future research. To achieve the objective of such a paradigm change is only successful with full governmental support: the common goal should be to develop, optimize, and validate new translational tools, to revize some of the older guidelines and harmonize their acceptance on a global level. On the regulatory side, first attempt to catch up with progress is already happening: example the publication of the new guidelines on the evaluation of biosimilars on the European level. The International Conferences on Harmonization continue their awareness of scientific advances and their Expert Working Groups modify and improve the important recommendations of their global guidelines.

The objective of this chapter is to propose several approaches that can contribute to improving efficiency and translational value of non-clinical testing. The suggestions target at:

1. Enlarge in silico data bases and improve their accessibility.
2. For researchers and editors: Publish **all** data, knowledge and experience. Include **all** data, positive and negative results.
3. Expand options for in vitro methods and elucidate their advantages and limitations.
4. Improve the selection of validated approaches and document the real values of animal models and applicability of methods.
5. Improve the design and execution of experiments and use fully randomization, blinding, optimal statistical interpretation etc.
6. Reflect clinical conditions in preclinical studies: Standard programs, biomarkers, specified endpoints. Understand the mechanisms of disease.
7. Identify weaknesses of methods: Which are the most relevant and predictive models and methods for human conditions?.
8. Introduce state-of-the-art methods into daily practice: GLP, statistics, combine functions with morphology, use non-invasive methods, provide support by kinetic data, etc.
9. Improve quality of reporting: analyze and assess all results, from **all** studies, focus on mean and individual effects.
10. Conscientious review of literature: built up "weight of evidence" approach, use information from Quality, Safety and Efficacy.
11. Encourage an open dialogue among researchers from all disciplines in industry and agencies.
12. Use scientific advice offered by Agencies to facilitate the decisions for best strategy during all phases of development.
13. Gain meaningful information of animal usage for human conditions at every step of development, get human data early as possible, use expedited explorations.
14. Better prediction of drug reactions in humans based on modern intelligent complex approaches will fasten access to efficient and safe drugs. For all these objectives, courage should be stimulated to swing from preconceived concepts to new methods and pathways. Only frank imaginative discussions will open the doors to an optimum way forward.

References and Further Reading

Akhtar A (2015) The flaws and human harms of animal experimentation. Camb Q Healthc Ethics 24(4):407–419

Arlington S (2012) From vision to decision: pharma 2020. Price Waterhouse Coopers (PwC), London. Available at https://www.pwc.com/gx/en/pharma-life-sciences/pharma2020/assets/pwc-pharma-success-strategies.pdf. Accessed on 3 July 2017

ARRIVE – Animal Research: Reporting of In Vivo Experiments. Available at https://www.nc3rs.org.uk/arrive-animal-research-reporting-vivo-experiments. Accessed on 3 July 2017

Arrowsmith J (2011) Trial watch: phase III and submission failures: 2007–2010. Nat Rev Drug Discov 10:87

Bailey J, Thew M, Balls M (2013) An analysis of the use of dogs in predicting human toxicology and drug safety. Altern Lab Anim 41(5):335–350

Bailey J, Thew M, Balls M (2014) An analysis of the use of animal models in predicting human toxicity and drug safety. Altern Lab Anim 42:181–199

Baumann A, Flagella K, Forster R, de Haan L, Kronenberg S, Locher M et al (2014) New challenches and opportunities in nonclinical safety testing of biologics. Regul Toxicol Pharmacol 69(2):226–233

Begley C, Ellis L (2012) Drug development: raise standards for preclinical cancer research. Nature 483:531–533

Bennani YL (2012) Drug discovery in the next decade: innovation needed ASAP. Drug Discov Today 16(17–18):779–792

Bhatia SN, Ingberg DE (2014) Microfluidic organs-on-chips. Nat Biotechnol 32:760–772

Bluemel J (2012) Considerations for the use of nonhuman primates in nonclinical safety assessment. In: Weinbauer GF, Vogel F (eds) Challenges in nonhuman primate research in the 21st century, Waxman (2012). Charles River Publication Series, pp 59–70. ISBN 978-3-8309-2839-3

Bode G, Van der Laan JW (2016) Paradigm change in cancerogenicity, Presentation: German Pharm-Tox Summit, Berlin, 29-02 to 03-03-2016, Berlin, ACS Publications, pubs.acs.org./crt, Symposium 19

Bode G, Clausing P, Gervais F, Loegsted J, Luft J, Nogues V, Sims J et al (2010) The utility of the minipig as an animal model in regulatory toxicology. J Pharmacol Toxicol Methods 62(3):196–220

Bussiere JL, Martin P, Horner M, Couch J, Flaherty M, Andrews L, Beyer J, Horvath C (2009) Alternative strategies for toxicity testing of species-specific biopharmaceuticals. Int J Toxicol 28(3):230–253

CAMARADES: Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies. Available at http://www.dcn.ed.ac.uk/camarades/. Accessed 3 July 2017

Casty F, Wieman M (2013) Drug development in the 21st century. The synergy of public, private, and international collaboration. Ther Innov Regul Sci 47(3):375–383

Cavagnaro J, Lima BS (2015) Regulatory acceptance of animal models of disease to support clinical trials of medicines and advanced therapy medicinal products. Eur J Pharmacol 759:51–62

Chalmers I, Glasziou P (2009) Avoidable waste in the production and reporting of research evidence. Lancet 374:86–89

Chapman KL, Pullen N, Andrews L, Ragan I (2010) The future of non-human primate use in mAb development. Drug Discov Today 15:235–242

Chapman K, Andrews L, Bajramovic JJ, Baldrick P et al (2012) The design of chronic toxicology studies of monoclonal antibodies: implications for the reduction in use of non-human primate. Regul Toxicol Pharmacol 62:347–354

Chapman KL, Holzgrefe H, Black LE, Brown M, Chellman G, Copeman C, Couch J et al (2013) Pharmaceutical toxicology: designing studies to reduce animal use, while maximizing human translation. Regul Toxicol Pharmacol 66:88–103

Cherkasov A, Muratov E, Fourches D, Varnek A, Baskin I, Cronin M et al (2014) QSAR modeling: where have you been? Where are you going to? J Med Chem 57(12):4977–5010

Clarke J, Hurst C, Martin P, Vahle J, Ponce R, Mounho B, Heidel S, Andrews L, Reynolds T, Cavagnaro J (2008) Duration of chronic toxicity studies for biotechnology-derived pharmaceuticals: is 6 months still appropriate? Regul Toxicol Pharmacol 50:2–22

Tufts center for study of drug development (CSDD) (2015) Outlook 2015. Available at: http://csdd.tufts.edu/files/uploads/Outlook-2015.pdf. Accessed 5 Feb 2017

Davis PR, Head E (2014) Prevention approaches in a preclinical canine model of Alzheimer's disease: benefits and challenges. Front Pharmacol 5:47

Denayer T, Stöhr T, van Roy M (2014) Animal models in translational medicine: validation and prediction. New Horiz Transl Med 2:5–11

DiMasi J, Kim J, Getz K (2014) The impact of collaborative and risk-sharing innovation approaches on clinical and regulatory cycle times. Ther Innov Regul Sci 48(4):482–487

DiMasi J, Grabowski H, Hansen R (2016) Innovation in the pharmaceutical industry: new estimates of R&D costs. J Health Econ 47:20–33

Directive 65/65/EEC of 26 January 1965 on the approximation of provisions laid down by Law, Regulation or Administrative Action relating to proprietary medicinal products. Off J Eur Union 22, 09/02/1965, pp 369–373

Directive 2001/82/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to veterinary medicinal products. Off J Eur Union L 311, 28/11/2001, pp 1–66

Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. Off J Eur Union L276, 20/10/2010, pp 33–79

Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. Off J Eur Union L 311, 28/11/2001, pp 37–128

Doke SK, Dhawale SC (2015) Alternatives to animal testing: a review. Saudi Pharm J 23(3):223–229

Egan KJ, Vesterinen HM, Beglopoulos V, Sena ES, Macleod MR (2016) From a mouse: systematic analysis reveals limitations of experiments testing interventions in Alzheimer's disease mouse models. Evid-Based Preclin Med 3(1):12–23

EMA (2008) Note for guidance on non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals. CPMP/ICH/286/95

EMA (2010) Guideline on repeated dose toxicity. CPMP/SWP/1042/99 Rev 1 Corr

EMA (2012) Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues. EMA/CHMP/BMWP/403543/2010

EMA (2014) Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. EMEA/CHMP/BMWP/42832/2005 Rev1

EMA (2016a) Guideline on the principles of regulatory acceptance of 3R (replacement, reduction, refinement) testing approaches. EMA/CHMP/CVMP/JEG-3Rs/450091/2012

EMA (2016b) Reflection paper providing an overview of the current regulatory testing requirements for medicinal products for human use and opportunities for implementation of the 3R (EMA/CHMP/CVMP/JEG3Rs-3Rs/742466/2015) – published for consultation

EMA (2016c) Reflection paper providing an overview of the current regulatory testing requirements for veterinary medicinal products and opportunities for implementation of the 3Rs (EMA/CHMP/CVMP/JEG-3Rs/164002/2016) – published for consultation

EMA (2017a) Guideline on strategies to identify and mitigate risks for first-in-human clinical trials with investigational medicinal products (EMEA/CHMP/SWP/28367/07 Rev. 1

EMA (2017b) Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017) – published for consultation

EURL ECVAM (2017) Homepage of the European Commission, The European Union Reference Laboratory for alternatives to animal testing (EURL-ECVAM). Available at https://eurl-ecvam.jrc.ec.europa.eu/. Accessed 1 Sept 2017

European Commission 2013: Report from the Commission to the Council and the European Parliament: seventh report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union, SEC(2010) 1107, 15pp. European Commission, Brussels. Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013DC0859&from=EN. Accessed 13 Jan 2017

European Commission (2016) Animals used for scientific purposes. Available at http://ec.europa.eu/environment/chemicals/lab_animals/3r/acceptance_en.htm. Accessed 4 July 2017

Evens R (2016) Pharma success in product development – does biotechnology change the paradigm in product development and attrition. AAPS J 18:281

Everitt JI (2015) The future of preclinical animal models in pharmaceutical discovery and development: a need to bring in cerebro to the in vivo discussions. Toxicol Pathol 43:70–77

FDA (2004) Innovation or stagnation: challenge and opportunity on the critical path to new medical products. In: Challenges and Opportunities Report US Department of Health and Human Services. Available at: https://www.fda.gov/downloads/scienceresearch/specialtopics/criticalpathinitiative/criticalpathopportunitiesreports/ucm113411.pdf. Accessed 13 Sept 2017

Friedrich A, Olejniczak K (2011) Evaluation of carcinogenicity studies of medicinal product for human use authorized via the European centralized procedure (1995–2009). Regul Toxicol Pharmacol 60:225–248

Ganderup NC, Harvey W, Mortensen JT, Harrouk W (2012) The minipig as nonrodent species in toxicology – where are we now? Int J Toxicol 31(6):507–528

Getz K (2011) Transforming legacy R&D through open innovation. Monit 4(3):16–21

Getz KA, Kaitin KI (2015) Why is the pharmaceutical and biotechnology industry struggling? In: Schüler P, Buckley B (eds) 2015: re-engineering clinical trials. Best practices for streamlining the development process. Academic Press/Elsevier. ISBN: 9780124202467

Glasziou P (2014) The role of open access in reducing waste in medical research. PLoS Med 11:e1001651

Greaves P, Williams A, Eve M (2004) First dose of potential new medicines to humans: how animals help. Nat Rev Drug Discov 3:226–236

Greek R, Hansen LA (2013) The strengths and limits of animal models as illustrated by the discovery and development of antibacterials. Biol Syst 2:109

Green SB (2015) Can animal data translate to innovations necessary for a new era of patient-centred and individualised healthcare? Bias in preclinical animal research. BMC Med Ethics 16:53

Groenink L, Folkerts G, Schuurman HJ (2015) Editorial, special issue on translational value of animal models. Eur J Pharmacol 759:1–2

Gulin JE, Rocco DM, García-Bournissen F (2015) Quality of reporting and adherence to ARRIVE guidelines in animal studies for Chagas disease preclinical drug research: a systematic review. PLoS Negl Trop Dis 9(11):e0004194

Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J (2014) Clinical development success rates for investigational drugs. Nat Biotechnol 32:40–51

Hester RE, Harrison M, Illing P (2006) General overview of the safety evaluation of chemicals. In: Hester RE, Harrison M (eds) Alternatives to animal testing. RSC Publishing. ISBN 978-0-85404-211-1

Heywood R (1981) Target organ toxicity. Toxicol Lett 8:349–358

Hirano E, Knutsen RH, Sugitani H, Ciliberto CH, Mecham RP (2007) Functional rescue of elastin insufficiency in mice by the human elastin gene: implications for mouse models of human disease. Circ Res 101(5):523–531

Horn J, de Haan RJ, Vermeulen M, Luiten PGM, Limburg M (2001) Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. Stroke 32:2433–2438

ICH guideline S3A: Note for guidance on toxicokinetics: the assessment of systemic exposure in toxicity studies. Oct 1994

ICH guideline S1A: Need for carcinogenicity studies of pharmaceuticals. Nov 1995

ICH guideline S1B: Testing for carcinogenicity of pharmaceuticals. July 1997

ICH guideline S 7 B: The nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals. May 2005
ICH guideline S1C(R2): Dose selection for carcinogenicity studies of pharmaceuticals. March 2008
ICH guideline M3(R2): Non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals. June 2009
ICH Guideline S2(R1): Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use. Nov 2011
ICH guideline S3B Pharmacokinetics: guidance for repeated dose tissue distribution studies. Oct 1994
ICH Guideline S5(R2): Detection of toxicity to reproduction for medicinal products & toxicity to male fertility. June 1993
ICH guideline S6 (R1): Preclinical safety evaluation of biotechnology-derived pharmaceuticals. June 2011
ICH guideline S7A: Safety pharmacology studies for human pharmaceuticals. Nov 2000
ICH guideline S9: Nonclinical evaluation for anticancer pharmaceuticals. Oct 2009
Kaitin K, DiMasi J (2011) Pharmaceutical innovation in the 21st century: new drug approvals in the first decade 2000–2009. Clin Pharm Therap 89(2):183–188
Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG (2010) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol 8:e1000412
Kimmelman J, Anderson JA (2012) Should preclinical studies be registered? Nat Biotechnol 30(6):488–489
Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–715
Kooijman M (2013) Why animal studies are still being used in drug development. Altern Lab Anim 41(6):79–81
Kooijman M, van Meer P, Moors E, Schellekens H (2012) Thirty years of preclinical safety evaluation of biopharmaceuticals: did scientific progress lead to appropriate regulatory guidance? Expert Opin Drug Saf 11(5):797–801
Korevaar DA, Hooft L, Ter Riet G (2011) Systematic reviews and meta-analyses of preclinical studies: publication bias in laboratory animal experiments. Lab Anim 45:225–230
Lamberti M, Mathias A, Myles J, Howe D, Getz K (2012) Evaluating the impact of patient recruitment and retention practices. Drug Inf J 46(5):573–580
Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG et al (2012) A call for transparent reporting to optimize the predictive value of preclinical research. Nature 490:187–191
Langley G, Austin CP, Balapure AK, Birnbaum LS, Bucher JR, Fentem J, Fitzpatrick SC et al (2015) Lessons from toxicology: developing a 21st-century paradigm for medical research. Environ Health Perspect 123:A268–A272
Lee GS, Lee KY, Choi KC, Ryu YH, Paik SG, Oh GT, Jeung EB (2007) Phenotype of a calbindin-D9k gene knockout is compensated for by the induction of other calcium transporter genes in a mouse model. J Bone Miner Res 22(12):1968–1978
Liebsch M, Grune B, Seiler A, Butzke D, Oelgeschläger M et al (2011) Alternatives to animal testing: current status and future perspectives. Arch Toxicol 85(8):841–858
Macleod M (2011) Why animal research need to improve. Nature 477:511
Macleod MR, O'Collins T, Howells DW, Donnan GA (2004) Pooling of animal experimental data reveals influence of study design and publication bias. Stroke 35:1203–1208
Matthews R (2008) Medical progress depends on animal models – doesn't it? J R Soc Med 101(2):95–98
Mella JR, Chiswick EL, King E, Remick DG (2014) Location, location, location: cytokine concentrations are dependent on blood sampling site. Shock 42:337–342
Michelson Prize & Grants (2014) The quality assurance toolkit. Available at http://www.michelsonprizeandgrants.org/resources/qa-toolkit. Accessed 29 Aug 2017
Morgan SJ, Elangbam CS, Berens S, Janovitz E, Vitsky A, Zabka T, Conour L (2013) Use of animal models of human disease for nonclinical safety assessment of novel pharmaceuticals. Toxicol Pathol 41(3):508–518
National Research Council [US] Institute for Laboratory Animal Research 2011: Guidance for the Description of Animal Research in Scientific Publications
NIH, National Human Genome Research Institute (2009) Knockout mice. Available at: https://www.genome.gov/12514551/. Accessed 2 Aug 2017
NTP (2017) https://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/acute-systemic-tox/in-vitro-validation/index.html
Nuffield Council on Bioethics (2005) The ethics of research involving animals. Available at: https://nuffieldbioethics.org/wp-content/uploads/The-ethics-of-research-involving-animals-full-report.pdf. Accessed 12 Sept 2017
OECD (2015) Guidance document on revisions to OECD genetic toxicology test guidelines. Available at https://www.oecd.org/chemicalsafety/testing/Genetic%20Toxicology%20Guidance%20Document%20Aug%2031%202015.pdf
Olson H, Betton G, Robinson D, Thomas K, Monro A, Kolaja G, Lilly P, Sanders J et al (2000) Concordance of the toxicity of pharmaceuticals in humans and in animals. Regul Toxicol Pharmacol 32(1):56–67
Paul S, Mytelka D, Dunwiddie C, Persinger C, Munos B, Lindborg S, Schacht A (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 9:203–214
Perrin S (2014) Preclinical research: make mouse studies work. Nature 507:7493
PhRMA (Pharmaceutical Research and Manufacturers of America) 2015: 2015 biopharmaceutical research industry profile.

Available at: http://www.phrma.org/sites/default/files/pdf/2015_phrma_profile.pdf. Accessed 4 Feb 2017

Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I (2004) Where is the evidence that animal research benefits humans? BMJ 328(7438):514–517

Prinz F, Schlange T, Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 10(9):712

Prot J, Leclerc E (2012) The current status of alternatives to animal testing and predictive toxicology methods using liver microfluidic biochips. Ann Biomed Eng 40(6):1228–1243

Pwc (PricewaterhouseCoopers) 2012: Pharma 2020: The vision. Which path will you take? Available at: http://www.pwc.com/gx/en/industries/pharmaceuticals-life-sciences/pharma-2020/pharma-2020-vision-path.html. Accessed 4 Feb 2017

Research Quality Association (2008) Guidelines for quality in non-regulated scientific research. Available at: http://www.therqa.com/publications/booklets/guidelines-quality-non-regulated-scientific-research/. Accessed 29 Aug 2017

Ritskes-Hoitinga M, Leenaars M, Avey M, Rovers M, Scholten R (2014) Systematic reviews of preclinical animal studies can make significant contributions to health care and more transparent translational medicine[editorial]. Cochrane Database Syst Rev (3). https://doi.org/10.1002/14651858.ED000078

Rowell JL, McCarthy DO, Alvarez CE (2012) Dog models of naturally occurring cancer. Trends Mol Med 17(7):380–388

Russell WM, Burch RL (1959) The principles of humane experimental technique. Methuen, London. ISBN 9780900767784

Sandercock P, Roberts I (2002) Systematic reviews of animal experiments. Lancet 360(9333):586

Scannel J, Blanckley A, Boldon H, Warrington B (2012) Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov 11:191–200

Schein PS, Davis RD, Carter S, Newman J, Schein DR, Rall DP (1970) The evaluation of anticancer drugs in dogs and monkeys for the prediction of qualitative toxicities in man. Clin Pharmacol Ther 11:3–40

SCHER (The Scientific Committee on Health and Environmental, The European Commission) (2009) Opinion on "The need for non-human primates in biomedical research, production and testing of products and devices". Available at: http://ec.europa.eu/health/scientific_committees/opinions_layman/en/non-human-primates/l-3/2-research-safety-testing.htm#0p0. Accessed 7 Aug 2017

Schueler P, Buckley B (eds) (2014) Re-engineering clinical trials: best practices for streamlining the development process. Academic Press. ISBN-10: 0124202462

Scott S, Kranz JE, Cole J, Lincecum JM, Thompson K, Kelly N et al (2008) Design, power, and interpretation of studies in the standard murine model of ALS. Amyotroph Lateral Scler 9:4–15

Sena E, van der Worp B, Howells D, Macleod M (2007) How can we improve the pre-clinical development of drugs for stroke? Trends Neurosci 30(9):433–439

Sena ES, van der Worp HB, Bath PM, Howells DW, Macleod MR (2010) Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol 8:e1000344

STAIR (Stroke Therapy Academic Industry Roundtable) (1999) Recommendations for standards regarding preclinical neuroprotective and restorative drug development. Stroke 30:2752–2758

Swaim LD, Taylor HW, Jersey GC (1985) The effects of handling techniques on serum ALT in mice. J Appl Toxicol 5:160–162

Taylor K, Gordon N, Langley G, Higgins W (2008) Estimates for worldwide laboratory animal use in 2005. Altern Lab Anim 36:327–342

Te Koppele J, Witkamp R (2008) Use of animal models of disease in the preclinical safety evaluation of biopharmaceuticals. In: Cavagnaro JA (ed) Preclinical safety evaluation of biopharmaceuticals: a science-based approach to facilitating clinical trials. Wiley, pp 293–308. ISBN: 978-0-470-10884-0

Teelmann K, Hohbach C, Lehmann H (1986) Preclinical safety testing of species specific proteins produced with recombinant DNA-techniques. An attempt to transfer current experience into future testing strategies. Arch Toxicol 59:195–200

Ter Riet G, Korevaar DA, Leenaars M, Sterk PJ, VanNoorden CJF, Bouter LM, Lutter R, Elferink RPO, Hooft L (2012) Publication bias in laboratory animal research: a survey on magnitude, drivers, consequences and potential solutions. PLoS One 7:e43404

Tiwari A (2014) Microdosing: a new approach to clinical drug development. Available at: https://de.slideshare.net/drashutoshtiwari/microdosing-phase-0-studies. Accessed on 22 Sept 2017

Tsilidis KK, Panagiotou OA, Sena ES, Aretouli E, Evangelou E, Howells DW, Salman RA, Macleod MR, Ioannidis JPA (2013) Evaluation of excess significance bias in animal studies of neurological diseases. PLoS One 11:e1001609

Van Aerts LA, De Smet K, Reichmann G, van der Laan JW, Schneider CK (2014) Biosimilars entering the clinic without animal studies: a paradigm shift in the European Union. MAbs 6(5):1155–1162

van der Worp HB, de Haan P, Morrema E, Kalkman CJ (2005) Methodological quality of animal studies on neuroprotection in focal cerebral ischaemia. J Neurol 252:1108–1114

Van Luijk J, Bakker B, Rovers MM, Ritskes-Hoitinga M, de Vries RB, Leenaars M (2014) Systematic reviews of animal studies; missing link in translational research? PLoS One 9:e89981

Van Meer P (2013) The scientific value of non-clinical animal studies in drug development. PhD thesis, Utrecht University

Van Meer P, Gispen-de Wied C, Moors E, Schellekens H (2012) The ability of animal studies to detect serious post marketing adverse events is limited. Regul Toxicol Pharmacol 64(3):345–349

Van Meer PJ, Kooijman M, van der Laan JW, Moors EH, Schellekens H (2013) The value of non-human primates in the development of monoclonal antibodies. Nat Biotechnol 31(10):882–883

Van Meer PJK, Graham ML, Schuurman HJ (2015) The safety, efficacy and regulatory triangle in drug development: impact for animal models and the use of animals. Eur J Pharmacol 759:3–13

Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, Antonic A, Howells DW, Macleod MR (2014) Meta-analysis of data from animal studies: a practical guide. J Neurosci Methods 221:92–102

Weaver ML, Blak Grossi A, Schützsack J, Parish J et al (2016) Vehicle systems and excipients used in minipig drug development studies. Toxicol Pathol 44(3):367–372

WHO (World Health Organization) (2006) Handbook: quality practices in basic biomedical research. Available at: http://www.who.int/tdr/publications/training-guideline-publications/handbook-quality-practices-biomedical-research/en/. Accessed 29 Oct 2017

Zbinden G (1987) Biotechnology products intended for human use, toxicological targets and research strategies. Prog Clin Biol Res 235:143–159

Zheng Y, Tesar DB, Benincosa L, Birnböck H (2012) Minipig as a potential translatable model for monoclonal antibody pharmacokinetics after intravenous and subcutaneous administration. MAbs 4(2):243–255

Index

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-68864-0

B

Q

T