Chapman & Hall/CRC Biostatistics Series

Biosimilars

Design and Analysis of Follow-on Biologics

Shein-Chung Chow

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Biosimilars

Design and Analysis of Follow-on Biologics

Chapman & Hall/CRC Biostatistics Series

Editor-in-Chief

Shein-Chung Chow, Ph.D.
Professor
Department of Biostatistics and Bioinformatics
Duke University School of Medicine
Durham, North Carolina

Series Editors

Byron Jones
Biometrical Fellow
Statistical Methodology
Integrated Information Sciences
Novartis Pharma AG
Basel, Switzerland

Jen-pei Liu
Professor
Division of Biometry
Department of Agronomy
National Taiwan University
Taipei, Taiwan

Karl E. Peace
Georgia Cancer Coalition
Distinguished Cancer Scholar
Senior Research Scientist and
Professor of Biostatistics
Jiann-Ping Hsu College of Public Health
Georgia Southern University
Statesboro, Georgia

Bruce W. Turnbull
Professor
School of Operations Research
and Industrial Engineering
Cornell University
Ithaca, New York

Chapman & Hall/CRC Biostatistics Series

Adaptive Design Methods in Clinical Trials, Second Edition
Shein-Chung Chow and Mark Chang

Adaptive Design Theory and Implementation Using SAS and R
Mark Chang

Advanced Bayesian Methods for Medical Test Accuracy
Lyle D. Broemeling

Advances in Clinical Trial Biostatistics
Nancy L. Geller

Applied Meta-Analysis with R
Ding-Geng (Din) Chen and Karl E. Peace

Basic Statistics and Pharmaceutical Statistical Applications, Second Edition
James E. De Muth

Bayesian Adaptive Methods for Clinical Trials
Scott M. Berry, Bradley P. Carlin, J. Jack Lee, and Peter Muller

Bayesian Analysis Made Simple: An Excel GUI for WinBUGS
Phil Woodward

Bayesian Methods for Measures of Agreement
Lyle D. Broemeling

Bayesian Methods in Health Economics
Gianluca Baio

Bayesian Missing Data Problems: EM, Data Augmentation and Noniterative Computation
Ming T. Tan, Guo-Liang Tian, and Kai Wang Ng

Bayesian Modeling in Bioinformatics
Dipak K. Dey, Samiran Ghosh, and Bani K. Mallick

Biosimilars: Design and Analysis of Follow-on Biologics
Shein-Chung Chow

Biostatistics: A Computing Approach
Stewart J. Anderson

Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation
Mikel Aickin

Clinical Trial Data Analysis using R
Ding-Geng (Din) Chen and Karl E. Peace

Clinical Trial Methodology
Karl E. Peace and Ding-Geng (Din) Chen

Computational Methods in Biomedical Research
Ravindra Khattree and Dayanand N. Naik

Computational Pharmacokinetics
Anders Källén

Confidence Intervals for Proportions and Related Measures of Effect Size
Robert G. Newcombe

Controversial Statistical Issues in Clinical Trials
Shein-Chung Chow

Data and Safety Monitoring Committees in Clinical Trials
Jay Herson

Design and Analysis of Animal Studies in Pharmaceutical Development
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bioavailability and Bioequivalence Studies, Third Edition
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bridging Studies
Jen-pei Liu, Shein-Chung Chow, and Chin-Fu Hsiao

Design and Analysis of Clinical Trials with Time-to-Event Endpoints
Karl E. Peace

Design and Analysis of Non-Inferiority Trials
Mark D. Rothmann, Brian L. Wiens, and Ivan S. F. Chan

Difference Equations with Public Health Applications
Lemuel A. Moyé and Asha Seth Kapadia

DNA Methylation Microarrays: Experimental Design and Statistical Analysis
Sun-Chong Wang and Arturas Petronis

DNA Microarrays and Related Genomics Techniques: Design, Analysis, and Interpretation of Experiments
David B. Allison, Grier P. Page, T. Mark Beasley, and Jode W. Edwards

Dose Finding by the Continual Reassessment Method
Ying Kuen Cheung

Elementary Bayesian Biostatistics
Lemuel A. Moyé

Frailty Models in Survival Analysis
Andreas Wienke

Generalized Linear Models: A Bayesian Perspective
Dipak K. Dey, Sujit K. Ghosh, and Bani K. Mallick

Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries
Daryl S. Paulson

Interval-Censored Time-to-Event Data: Methods and Applications
Ding-Geng (Din) Chen, Jianguo Sun, and Karl E. Peace

Joint Models for Longitudinal and Time-to-Event Data: With Applications in R
Dimitris Rizopoulos

Measures of Interobserver Agreement and Reliability, Second Edition
Mohamed M. Shoukri

Medical Biostatistics, Third Edition
A. Indrayan

Meta-Analysis in Medicine and Health Policy
Dalene Stangl and Donald A. Berry

Monte Carlo Simulation for the Pharmaceutical Industry: Concepts, Algorithms, and Case Studies
Mark Chang

Multiple Testing Problems in Pharmaceutical Statistics
Alex Dmitrienko, Ajit C. Tamhane, and Frank Bretz

Optimal Design for Nonlinear Response Models
Valerii V. Fedorov and Sergei L. Leonov

Randomized Clinical Trials of Nonpharmacological Treatments
Isabelle Boutron, Philippe Ravaud, and David Moher

Randomized Phase II Cancer Clinical Trials
Sin-Ho Jung

Sample Size Calculations in Clinical Research, Second Edition
Shein-Chung Chow, Jun Shao and Hansheng Wang

Statistical Design and Analysis of Stability Studies
Shein-Chung Chow

Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis
Kelly H. Zou, Aiyi Liu, Andriy Bandos, Lucila Ohno-Machado, and Howard Rockette

Statistical Methods for Clinical Trials
Mark X. Norleans

Statistics in Drug Research: Methodologies and Recent Developments
Shein-Chung Chow and Jun Shao

Statistics in the Pharmaceutical Industry, Third Edition
Ralph Buncher and Jia-Yeong Tsay

Survival Analysis in Medicine and Genetics
Jialiang Li and Shuangge Ma

Translational Medicine: Strategies and Statistical Methods
Dennis Cosmatos and Shein-Chung Chow

Chapman & Hall/CRC Biostatistics Series

Biosimilars

Design and Analysis of Follow-on Biologics

Shein-Chung Chow

Duke University School of Medicine
Durham, North Carolina, USA

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

Chow, Shein-Chung, 1955-

Biosimilars : design and analysis of follow-on biologics / Shein-Chung Chow.

pages cm. -- (Chapman & Hall/CRC biostatistics series ; 60)

Includes bibliographical references and index.

ISBN 978-1-4665-7969-9 (hardback) 1. Pharmaceutical biotechnology. 2. Pharmaceutical biotechnology industry. 3. Drugs--Generic substitution. 4. Pharmaceutical policy. 5. Biological products. I. Title.

RS380.C45 2014

615.1'9--dc23

2013012926

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2014 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20130625

International Standard Book Number-13: 978-1-4665-7970-5 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

Biologic drug products are therapeutic moieties that are manufactured using a living system or organism. These are important life-saving drug products for patients with unmet medical needs. They also comprise a growing segment in the pharmaceutical industry. In 2007, for instance, worldwide sales of biological products reached $94 billion, accounting for about 15% of the pharmaceutical industry's gross revenue. Meanwhile, many biological products face losing their patents in the next decade. Attempts have been made therefore to establish an abbreviated regulatory pathway for approval of biosimilar drug products, that is, follow-on (or subsequent entered) biologics of the innovator's biological products in order to reduce cost. However, due to the complexity of the structures of biosimilar products and the nature of the manufacturing process, biological products differ from traditional small-molecule (chemical) drug products. Although the concepts and principles for bioequivalence and interchangeability could be the same for both chemical generics and biosimilar products, scientific challenges remain for establishing an abbreviated regulatory pathway for approval of biosimilar products due to their unique characteristics.

This book is intended to be the first book entirely devoted to the design and analysis of biosimilarity and drug interchangeability and includes tests for comparability in important quality attributes at critical stages of manufacturing processes of biological products. It covers most of the statistical issues that one may encounter in biosimilar studies under various study designs at different stages of research and development of biological products. The goal of this book is to provide a useful desk reference and describe the state of the art to (1) scientists and researchers engaged in pharmaceutical/clinical research and development of biological products, (2) those in government regulatory agencies who have to make decisions in the review and approval process of biological regulatory submissions, and (3) biostatisticians who provide statistical support to the assessment of biosimilarity and drug interchangeability of biosimilar products. I hope that this book can serve as a bridge among the pharmaceutical/biotechnology industry, government regulatory agencies, and academia.

The scope of this book is restricted to scientific factors and practical issues related to the design and analysis of biosimilar studies that are commonly seen in biosimilar research and development. Also, since regulatory requirements for assessment of biosimilar products between the European Medicines Agency (EMA) and the United States Food and Drug Administration (FDA) are similar but slightly different, this book primarily focuses on regulatory requirements from FDA. The book contains 17 chapters. Chapter 1 provides a background of pharmaceutical/clinical

development of biosimilar products and describes commonly seen scientific factors and practical issues in biosimilar clinical research and development. Chapter 2 reviews past experience for generic approval of small-molecule drug products. Chapter 3 summarizes regulatory requirements for assessment of biosimilar products (or follow-on biologics) and includes a review of recently published FDA draft guidances on biosimilar products. Criteria for assessment of biosimilarity, which are available in the regulatory guidances/guidelines and/or literature, are described in Chapter 4. Chapter 5 introduces statistical methods for assessing average biosimilarity based on the concept of relative distance between a test product and a reference product as compared to the distance between the reference product and itself. Chapter 6 proposes a general approach based on biosimilarity index (reproducibility probability) for the assessment of biosimilar products. Chapter 7 explores the relationship between the concept of testing non-inferiority and testing for equivalence. Chapter 8 deals with statistical tests for assessment of biosimilarity in variability of biosimilar products. Formulas or procedures for sample size calculations for comparing variabilities under a crossover design or a parallel design with or without replicates are given in Chapter 9. Chapter 10 studies the impact of variability on biosimilarity limits for assessing biosimilar products. Chapter 11 investigates the feasibility/applicability of the assessment of interchangeability (in terms of the concepts of switching and alternating among biosimilar products) and describes useful study designs that address switching and/or alternation in biosimilar studies. The issue of immunogenicity in biosimilar studies is examined in Chapter 12. Chemistry, manufacturing, and control (CMC) requirements for biological products in regulatory submission are discussed in Chapter 13. Chapter 14 provides statistical methods for testing comparability of important quality attributes at various critical stages of a manufacturing process of biosimilar products. Stability design and analysis of biosimilar products are dealt with in Chapter 15. Chapter 16 discusses statistical tests for assessment of biosimilarity using biomarker data. Current issues for assessing biosimilarity and interchangeability of biosimilar products are discussed in the Chapter 17.

From Taylor & Francis Group, I would like to thank David Grubbs for providing me the opportunity to work on this book. I wish to express my gratitude to my wife Annpey Pong, PhD, for her understanding, constant encouragement, and support during the preparation of this book. I thank Laszlo Endrenyi, PhD, of the University of Toronto for his constructive comments and editing, which have led to a significant improvement of the book. I would also like to thank colleagues from the Statistical Scientific Advisory Board (SSAB) on Biosimilars (sponsored by Amgen, Inc.), Amgen, Inc., the Department of Biostatistics and Bioinformatics, Duke Clinical Research Institute (DCRI), Duke Clinical Research Unit (DCRU), and the Center for AIDS Research (CFAR) of Duke University School of Medicine as well as many friends from academia, the pharmaceutical industry, and

regulatory agencies for their support and discussions during the preparation of this book.

Finally, the views expressed are those of the author and not necessarily those of Duke University School of Medicine. I am solely responsible for the contents and errors of this book. Any comments and suggestions will be very much appreciated.

Shein-Chung Chow, PhD

School of Medicine
Duke University
Durham, North Carolina

1

Introduction

1.1 Background

In the United States (U.S.), for small-molecule drug products, when an innovative (brand-name) drug product is going off patent, pharmaceutical and/or generic companies may file an abbreviated new drug application (ANDA) for the approval of the generic copies of the brand-name drug. In 1984, the United States Food and Drug Administration (FDA) was authorized to approve generic drug products under the *Drug Price Competition and Patent Term Restoration Act*, which is also known as the *Hatch-Waxman Act*. For the approval of generic (small-molecule) drug products, the FDA requires that evidence of *average* of bioavailability, which is measured in terms of the rate and extent of drug absorption, be provided through the conduct of bioequivalence studies. As indicated by Chow and Liu (2008), the assessment of bioequivalence as a surrogate for evaluation of drug safety and efficacy is based on the so-called *Fundamental Bioequivalence Assumption* that if two drug products are shown to be bioequivalent in average bioavailability, it is assumed that they will reach the same therapeutic effect or that they are therapeutically equivalent. Many practitioners interpret that approved generics and the brand-name drug can, in most cases, be used interchangeably since they are therapeutically equivalent. Under the Fundamental Bioequivalence Assumption, regulatory requirements (e.g., FDA guidances), study design (e.g., a standard two-sequence, two-period crossover design), acceptance criteria (e.g., the 80/125 rule based on log-transformed data), and statistical methods (e.g., Shuirmann's two one-sided tests procedure or the confidence interval approach) for the assessment of bioequivalence have been well established over the past several decades (see, e.g., Schuirmann, 1987; FDA, 2001, 2003; Chow and Liu, 2008).

Unlike small-molecule drug products, a generic version of a biological products is only a similar biological drug product (SBDP) in comparison with the originator biological product. It should be noted that the SBDPs are *not* like the small-molecule generic drug products, which are usually referred to as containing *identical* active ingredient(s) as the innovative drug product. The concept for the development of SBDPs, which are made of living cells or

organisms, is very different from that of the (small-molecule) generic drug products. The SBDPs are usually referred to as *biosimilars* by the European Medicines Agency (EMA) of the European Union (EU), *follow-on biologics* (FOB or FoB) by the U.S. FDA, and subsequent entered biologics (SEB) by the Health Canada. Throughout this book, unless otherwise stated, the term biosimilars or follow-on biologics will be used. Note that experience with biosimilar development worldwide can be found in McCamish and Woollett (2011).

Webber (2007) defines follow-on (protein) biologics as products that are intended to be *sufficiently similar* to an approved product to permit the applicant to rely on certain existing scientific knowledge about the safety and efficacy of an approved reference product. It should be noted that the generic (small-molecule) drug products are fundamentally different from biosimilar (large-molecule) drug products. For example, biosimilar products are made of living cells and have heterogeneous structures (usually mixtures of related molecules) which are difficult to characterize. In addition, biosimilar products are often variable and sensitive to environmental conditions such as light and temperature. A small change or variation at any critical stage of a manufacturing process of a biological product could result in a drastic change in clinical outcomes. Thus, the current standard methods for bioequivalence assessment of generic drug products may not be appropriate for the assessment of biosimilar products due to these fundamental differences.

On March 23, 2010, the *Biologics Price Competition and Innovation* (BPCI) *Act* (as part of the Affordable Care Act) was written into law, which has given the FDA the authority to approve similar biological drug products. As indicated in the BPCI Act, a biosimilar product is defined as a product that is *highly similar* to the reference product notwithstanding minor differences in clinically inactive components and there are no clinically meaningful differences in terms of safety, purity, and potency. However, little or no discussion regarding how similar is considered highly similar is given in the BPCI Act. As stated in Subsection 351(k)(4), a biological product is considered to be *interchangeable* with the reference product if (1) the biological product is biosimilar to the reference product; and (2) it can be *expected* to produce the same clinical result in *any given patient*. In addition, for a biological product that is administered more than once to an individual, the risk in terms of safety or diminished efficacy of alternating or switching between use of the biological product and the reference product is not greater than the risk of using the reference product without such alternation or switch. Thus, by definition, there is a clear distinction between biosimilarity and interchangeability. In other words, biosimilarity does not imply interchangeability, which is much more stringent. The BPCI Act also states that if a test product is judged to be interchangeable with the reference product, then it may be substituted, even alternated, without a possible intervention, or even notification, of the health care provider. However, as noted earlier, interchangeability is expected to produce the *same* clinical result in *any given patient*, which can be interpreted as that the same clinical result

can be expected in *every single patient*. In reality, conceivably, lawsuits may be filed if adverse effects are recorded in a patient after switching from one product to another.

Following the passage of the BPCI Act, in order to obtain input on specific issues and challenges associated with the implementation of the BPCI Act, the U.S. FDA conducted a 2 day public hearing on the *Approval Pathway for Biosimilar and Interchangeability Biological Products* held on November 2–3, 2010, at the FDA in Silver Spring, Maryland. Several scientific factors were raised and discussed at the public hearing. These scientific factors included criteria for assessing biosimilarity, study design and analysis methods for the assessment of biosimilarity, and tests for comparability in quality attributes of the manufacturing process and/or immunogenicity (see, e.g., Chow et al., 2010). These issues primarily focused on the assessment of biosimilarity. The issue of interchangeability in terms of the concepts of alternating and switching was also mentioned and discussed. The discussions of these scientific factors have led to the development of regulatory guidances. On February 9, 2012, the U.S. FDA circulated three draft guidances on the demonstration of biosimilarity for comments. These draft guidances are

1. Scientific Considerations in Demonstrating Biosimilarity to a Reference Product (FDA, 2012a)
2. Quality Considerations in Demonstrating Biosimilarity to a Reference Protein Product (FDA, 2012b)
3. Biosimilars: Questions and Answers Regarding Implementation of the BPCI Act of 2009 (FDA, 2012c)

Subsequently, the FDA hosted another public hearing on the discussion of these draft guidances at the FDA on May 11, 2012.

As patents of a number of biological products are due to expire in the next few years, the subsequent production of follow-on products has aroused interest within the pharmaceutical industry as biosimilar manufacturers strive to obtain part of an already large and rapidly growing market. The potential opportunity for price reductions versus the originator biological products remains to be determined, as the advantage of a slightly cheaper price may be outweighed by the hypothetical increased risk of side effects from biosimilar molecules that are not exact copies of their originators. In this chapter, we shall focus not only on the fundamental differences between small-molecule drug products and biological products but also on practical issues surrounding the assessment of biosimilar products, including scientific factors on biosimilarity, drug interchangeability, quality, and comparability in manufacturing process, and clinical efficacy and side effects.

The rest of this chapter is organized as follows. In the next section, fundamental differences between small-molecule drug products and biological drug products are briefly described. Section 1.3 provides a brief summary of the current regulatory requirements for the approval of biosimilars in the

European Union and the United States. The concepts and corresponding issues regarding biosimilarity, drug interchangeability, and the quality and comparability in the manufacturing process are discussed in Sections 1.4, 1.5, and 1.6. Note that basic concepts and issues are briefly introduced here. These basic concepts and issues will be discussed in greater detail in later chapters. The aim and scope of the book are given in the last section of this chapter.

1.2 Fundamental Differences

Biosimilars are fundamentally different from small-molecule generic drugs. Some of the fundamental differences between biosimilars and generic drugs are summarized in Table 1.1. As can be seen from the table, for example, small-molecule drug products are made by chemical synthesis, while large-molecule biologics are made of living cells or organisms. Small-molecule drug products have well-defined structures which are easy to characterize, while biosimilars have heterogeneous structures with mixtures of related molecules which are difficult to characterize. Small-molecule drug products are usually relatively stable, while biosimilars are known to be variable and very sensitive to environmental conditions such as light and temperature. A small change or variation during the manufacturing process may translate to a drastic change in clinical outcomes (e.g., safety and effectiveness). Small-molecule drug products which are often taken orally are generally prescribed by general practitioners, while biosimilars which are usually injected are often prescribed by specialists. In addition, unlike small-molecule drug products, biosimilars may induce unwanted immune responses which may cause a loss of efficacy or change in their safety profile. Moreover, with differences

TABLE 1.1

Fundamental Differences between Chemical Drugs and Biologics

Chemical Drugs	Biologics
Made by chemical synthesis	Made of living cells or organisms
Defined structure	Heterogeneous structure Mixtures of related molecules
Easy to characterize	Difficult to characterize
Relatively stable	Variable Sensitive to environmental conditions such as light and temperature
No issue of immunogenicity	Issue of immunogenicity
Usually taken orally	Usually injected
Often prescribed by a general practitioner	Usually prescribed by specialists

in the size and complexity of the active substance, important differences also include the nature of the manufacturing process.

As indicated by Kuhlmann and Covic (2006), biological products are usually recombinant-protein molecules manufactured in living cells. Thus, manufacturing processes for biological products are highly complex and require hundreds of specific isolation and purification steps. As a result, in practice, it is impossible to produce an *identical* copy of a biological products, as changes to the structure of the molecule can occur with changes in the production process. Since a protein can be modified during the process (e.g., a side chain may be added, the structure may have changed due to protein misfolding, and so on), different manufacturing processes may lead to structural differences in the final product, which result in differences in efficacy and safety, and may have a negative impact on the immune responses of patients. It should be noted that these issues may also occur during the post-approval changes of the innovator's biological products.

Biosimilar products are not generic products since they are *not* identical to their originator products. Thus, biosimilars should not be brought to market using the same procedure applied to generics. This is partly a reflection of the complexities of manufacturing and safety and efficacy controls of biosimilars when compared to their small-molecule generic counterparts (see, e.g., Chirino and Mire-Sluis, 2004; Schellekens, 2004; Crommelin et al., 2005; Roger and Mikhail, 2007). Instead, for investigating biological products, including biosimilars, the state-of-the-art of analytical procedures should be applied.

1.3 Regulatory Requirements

For the approval of biosimilars in the EU community, the EMA has issued a new guideline describing general principles for the approval of similar biological medicinal products, or biosimilars. The guideline is accompanied by several concept papers that outline areas in which the agency intends to provide more targeted guidance (EMA, 2003a,b, 2006a–g). Specifically, the concept papers discuss approval requirements for several classes of human recombinant products containing erythropoietin, human growth hormone, granulocyte-colony stimulating factor, and insulin. The guideline consists of a checklist of documents published to date relevant to the data requirements for biological pharmaceuticals. It is not clear what specific scientific requirements will be applied to biosimilar applications. In addition, it is not clear how the agency will treat the innovator data contained in the dossiers of the reference product. The guideline provides a useful summary of the biosimilar legislation and previous EU publications, but it provides few answers to the issues.

On the other hand, for the approval of follow-on biologics in the United States, its path depends on whether the biological products is approved under the United States Food, Drug, and Cosmetic Act (US FD&C) or if it is licensed under the United States Public Health Service Act (US PHS). As indicated, some proteins are licensed under the PHS Act, while some are approved under the FD&C Act. For products approved under a New Drug Application (NDA, under the US FD&C Act), generic versions of products can be approved under an ANDA, for example, under Section 505(b)(2) of the FD&C Act. For products that are licensed under a Biologics License Application (BLA, under the US PHS Act), there exists no abbreviated BLA. As pointed out by Woodcock et al. (2007), for the assessment of similarity of follow-on biologics, the FDA would consider the following factors:

1. The robustness of the manufacturing process
2. The degree to which structural similarity could be assessed
3. The extent to which the mechanism of action was understood
4. The existence of valid, mechanistically related pharmacodynamic (PD) assays
5. Comparative pharmacokinetics (PK)
6. Comparative immunogenicity
7. The amount of available clinical data
8. The extent of experience with the original product

A typical example would be the recent regulatory approval of Omnitrope (somatropin), which was approved in 2006 under Section 505(b)(2) of the FD&C Act. Omnitrope was approved based on the following evaluations:

1. Physicochemical testing that established highly similar structure to Genotropin
2. New non-clinical pharmacology and toxicology data specific to Omnitrope
3. PK, PD, and comparative bioavailability data
4. Clinical efficacy and safety data from comparative controlled trials and from long-term trials with Omnitrope
5. Vast clinical experience and a wealth of published literature concerning the clinical effects (safety and effectiveness) of human growth hormone

The approval of Omnitrope was based on an ad hoc, case-by-case review of an individual biosimilar application. In practice, there is a strong industrial interest and desire for the regulatory agencies to develop review standards and an approval process for biosimilars instead of an ad hoc, case-by-case review of

individual biosimilar applications. For this purpose, the FDA has established three committees to ensure consistency in the FDA's regulatory approach and guidance to applicants regarding development programs for proposed biosimilar biological products which are intended for submission under the new section 351(k) of the PHS Act. The three committees involve the two centers of FDA which actively review submissions on new biosimilars: the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER). The committees to review applications for biosimilars are the CDER/CBER Biosimilar Implementation Committee (BIC), the CDER Biosimilar Review Committee (BRC), and the CBER Biosimilar Review Committee. The CDER/CBER BIC will focus on the cross-center policy issues related to the implementation of the BPCI Act. The CDER BRC and CBER BRC committees are responsible for considering applicant requests for advice about proposed development programs for biosimilar products, reviewing Biologic License Applications (BLAs) that are submitted under section 351(k) of the PHS Act, and managing related issues. Thus, the CDER BRC (CBER BRC) review process steps include the following:

1. An applicant submits a request for advice.
2. Internal review team meeting.
3. Internal CDER BRC (or CBER BRC) meeting.
4. Internal post-BRC meeting.
5. CDER (CBER) meeting with the applicant.

As mentioned earlier, the FDA has circulated, on February 9, 2012, three draft guidances on the assessment of biosimilar products. The first draft guidance regarding scientific considerations is intended to assist sponsors in demonstrating that a proposed therapeutic protein product is biosimilar to a reference product for the purpose of a submission for a marketing application under section 351(k) of the PHS Act. The second draft guidance on quality considerations describes the Agency's current thinking on the factors to consider when demonstrating that a proposed protein product is highly similar to a reference product. Specifically, the guidance is intended to provide recommendations to applicants on scientific and technical information on the chemistry, manufacturing, and controls (CMC) section of a marketing application for a proposed biosimilar product. The third draft guidance provides answers to common questions from sponsors interested in developing proposed biosimilar products, biologics license application (BLA) holders, and other interested parties regarding FDA's interpretation of the BPCI Act.

It should be noted that the three draft guidances do not describe the FDA's current position on drug interchangeability. In order to obtain public input and comments on the draft guidances and drug interchangeability, the FDA also hosted a public hearing at FDA on May 11, 2012. The thinking on drug interchangeability in terms of the concepts of switching and alternating was

explored while some useful study designs and statistical methods were proposed and discussed at the public hearing. More details regarding individual regulatory requirements for assessing biosimilarity of biosimilar products from the EU, the United States, and Japan and a discussion regarding the comparison and harmonization of these regulatory requirements are given in Chapter 3.

1.4 Biosimilarity

1.4.1 Definition and Basic Principles

As indicated earlier, the BPCI Act defines a biosimilar product as a product that is *highly similar* to the reference product, notwithstanding minor differences in clinically inactive components. There are no clinically meaningful differences between a biosimilar and an originator biological product in terms of safety, purity, and potency. Based on this definition, we would interpret that a biological medicine is biosimilar to a reference biological medicine if it is highly similar to the reference in *safety*, *purity*, and *potency*. Here purity may be related to some important *quality* attributes at critical stages of the manufacturing process, and potency has something to do with the *stability* and *efficacy* of the biosimilar product. However, little or no discussion regarding how similar is considered highly similar (or how close is considered sufficiently close) was mentioned in the BPCI Act.

The BPCI Act seems to suggest that a biosimilar product should be highly similar (sufficiently close) to the reference drug product in all spectrums of good drug characteristics such as identity, strength (potency), quality, purity, safety, and stability as described in the U.S. Pharmacopeia and National Formulary (see, e.g., USP/NF, 2000). In practice, however, it is almost impossible to demonstrate that a biosimilar product is highly similar to the reference product in *all* aspects of good drug characteristics in a *single* study. Thus, to ensure that a biosimilar product is highly similar to the reference product in terms of these good drug characteristics, different biosimilar studies may be required. For example, if safety and efficacy are the concern, then a clinical trial must be conducted to demonstrate that there are no clinically meaningful differences in terms of safety and efficacy between a biosimilar product and the innovator biological product. On the other hand, to ensure that important quality attributes are highly similar, critical stages of the manufacturing process, assay development/validation, process control/validation, and product specification of the reference product should be necessarily established through the conduct of relevant studies. In addition, studies need to be conducted for testing the comparability in the manufacturing process (raw materials, in-use materials, and end-product) between the biosimilars and the reference product. This is extremely important because biological

products are known to be sensitive to small changes or variations in environmental factors such as light and temperature, during the manufacturing process. In some cases, if a surrogate endpoint such as PK, PD, or a genomic marker is predictive of the primary efficacy/safety clinical endpoint, then a PK/PD or genomic study may be used to assess biosimilarity.

The current regulatory requirements are guided on a case-by-case basis by the following basic principles:

1. The extent of the physicochemical and biological characterization of the product
2. The nature or possible changes in the quality and structure of the biological product due the changes in the manufacturing process (and their unexpected outcomes)
3. Clinical/regulatory experiences with the particular class of the product in question
4. Several factors that need to be considered for biocomparability

Most recently, in its recent draft guidance on *Scientific Considerations in Demonstrating Biosimilarity to a Reference Product*, the FDA has suggested that considerations and reviews of biosimilarity should be based on the totality-of-the-evidence. This indicates that the FDA is interested in demonstrating *global* similarity in all aspects related to safety, purity, and potency of the biosimilar products.

1.4.2 Criteria for Bioequivalence/Biosimilarity

The BPCI Act defines a biosimilar product as a biological product that is highly similar to the reference drug product. However, no criteria for assessing biosimilarity were mentioned in the Act. Statistically, one could refer to as *similarity* between two drug products as similarity in average, variability, or distribution of the response of a specific study endpoint of interest. In practice, the assessment of similarity in the average of the response of a specific study endpoint is often considered. A typical example is the assessment of average bioequivalence in terms of drug absorption (which is measured by the study endpoint of area under the blood or plasma concentration time curve or maximum concentration) for the regulatory approval of generic drug products. In this book, unless otherwise stated, we shall focus on the biosimilarity in the *average* response of the study endpoint of interest in a given biosimilar study. More details regarding the bioequivalence experience for small-molecule drug products are discussed in the next chapter.

In practice, the terms of biosimilarity (similarity), bioequivalence (equivalence), comparability, biocomparability, and consistency are alternately used in biopharmaceutical/biotechnology research and development. For comparisons between drug products, some criteria for the assessment of

bioequivalence (e.g., the comparison of drug absorption profiles), similarity (e.g., the comparison of dissolution profiles), and comparability or consistency (e.g., comparisons between manufacturing processes) are available in either regulatory guidelines/guidances or the literature. These criteria, however, can be classified into the following categories:

1. Absolute change versus relative change
2. Aggregated versus disaggregated criteria
3. Moment-based versus probability-based criteria
4. Scaled versus unscaled criteria
5. Weighted versus unweighted criteria

In what follows, these categories of criteria are briefly reviewed. While the criteria have been applied to bioequivalence studies, they are equally relevant to investigations of biosimilarity.

1.4.2.1 Absolute Change versus Relative Change

In clinical research and development, for a given study endpoint, post-treatment absolute change from baseline or post-treatment relative change (% change) from a baseline is usually considered for comparisons between treatment groups. A typical example would be the study of weight reduction in an obese patient population. In practice, it is not clear whether a clinically meaningful difference in terms of absolute change from the baseline can be translated to a clinically meaningful difference in terms of relative change from the baseline. Sample size calculations based on power analysis in terms of absolute change from the baseline or relative change from the baseline could lead to a very different result.

For generic approval, current U.S. regulation adopts a one size-fits-all criterion based on *relative change* for bioequivalence assessment. In other words, we conclude (average) bioequivalence between a test product and a reference product if the 90% confidence interval for the ratio of geometric means of the primary endpoint (e.g., a PK response such as the area under the blood or plasma concentration time curve) between the two drug products is (in%) completely within 80% and 125%. Note that regulatory agencies suggest that a log-transformation be performed before data analysis for the assessment of bioequivalence.

1.4.2.2 Aggregated versus Disaggregated Criteria

As indicated by Chow and Liu (2008), bioequivalence can be assessed by evaluating differences, *separately*, in averages, intra-subject variabilities, and variance due to subject-by-formulation interaction between drug products. Individual criteria for the assessment of differences in averages,

intra-subject variabilities, and variance due to subject-by-formulation interaction are referred to as *disaggregated criteria*. If the criterion is a single summary measure composed of these individual criteria, it is called an *aggregated criterion*.

For the assessment of bioequivalence in average bioavailability (ABE), most regulatory agencies recommend the use of a disaggregated criterion based on average bioavailability. That is, bioequivalence is concluded if the average bioavailability of the test formulation is within (80%, 125%) that of the reference formulation, with a certain assurance. Note that the EMA (2001) and WHO (2005) use the same equivalence criterion of 80%–125% for the log-transformed PK responses such as the area under the blood or plasma concentration time curve (AUC).

Aggregated criteria for population bioequivalence (PBE) and individual bioequivalence (IBE) were presented in an FDA guidance (FDA, 2001). PBE and IBE will be discussed in greater detail in Chapter 4. It is noted here only that both procedures rely on aggregated criteria. PBE evaluates *jointly* the differences between the means and between the total variances of the two drug products. (Total variances are the sums of the between- and within-subject variances.) Similarly, IBE assesses *jointly* the differences between the means and between the intra-subject variances as well as the variance component of the subject-by-product interaction (FDA, 2001). These examples of aggregated criteria will be considered later.

For aggregated criteria, the FDA proposes the use of an individual bioequivalence (IBE) criterion (IBC) for addressing drug switchability and population bioequivalence (PBE) criterion (PBC) for addressing drug prescribability (FDA, 2001). For the assessment of IBE, the IBC, denoted by θ_I, can be expressed as

$$\theta_I = \frac{\delta^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{\max\{\sigma_{W0}^2, \sigma_{WR}^2\}}, \tag{1.1}$$

where

$\delta = \mu_T - \mu_R$, σ_{WT}^2, σ_{WR}^2, σ_D^2 are the true differences between means, the intra-subject (within-subject) variabilities of the test product and the reference product, and the variance component due to subject-by-formulation interaction between drug products, respectively

σ_{W0}^2 is a scale parameter specified by the user

Similarly, the PBC for the assessment of PBE, denoted by θ_P, suggested in the FDA guidance (FDA, 2001) is given by

$$\theta_P = \frac{\delta^2 + \sigma_{TT}^2 - \sigma_{TR}^2}{\max\{\sigma_{T0}^2, \sigma_{TR}^2\}}, \tag{1.2}$$

where

σ_{TT}^2, σ_{TR}^2 are the total variances for the test product and the reference product, respectively

σ_{T0}^2 is a scale parameter specified by the user

A typical approach is to construct a one-sided 95% confidence interval for $\theta_I(\theta_P)$ for the assessment of individual (population) bioequivalence. If the one-sided 95% upper confidence limit is less than the bioequivalence limit of $\theta_I(\theta_P)$, we then conclude that the test product is bioequivalent to that of the reference product in terms of individual (population) bioequivalence. More details regarding individual and PBE can be found in Chow and Liu (2008).

1.4.2.3 Moment-Based versus Probability-Based Criteria

Schall and Luus (1993) proposed the moment-based and probability-based measures for the expected discrepancy in PK responses between drug products. The moment-based measure compares the expectation of the (squared) difference between responses of the test and reference products (*T* versus *R*) with that of the (squared) difference between two administrations of the reference formulation (*R* versus *R′*). The probability-based approach makes the same comparison but utilizes the probabilities for occurrence of such differences. Details of the approaches will be provided in Chapter 4. The moment-based measure suggested by Schall and Luus (1993) is based on the following expected mean-squared differences:

$$d(Y_j; Y_{j'}) = \begin{cases} E(Y_T - Y_R)^2 & \text{if } j = T \text{ and } j' = R \\ E(Y_R - Y_R')^2 & \text{if } j = R \text{ and } j' = R. \end{cases} \tag{1.3}$$

For some pre-specified positive number *r*, one of the probability-based measures for the expected discrepancy is given as (Schall and Luus, 1993)

$$d(Y_j; Y_{j'}) = \begin{cases} P\{|Y_T - Y_R| < r\} & \text{if } j = T \text{ and } j' = R \\ P\{|Y_R - Y_R'| < r\} & \text{if } j = R \text{ and } j' = R. \end{cases} \tag{1.4}$$

$d(Y_T; Y_R)$ measures the expected discrepancy for some PK metric between the test and reference formulations, and $d(Y_R; Y_R')$ provides the expected discrepancy between the repeated administrations of the reference formulation. The role of $d(Y_R; Y_R')$ in the formulation of the bioequivalence criteria is to serve as a control. The rationale is that the reference formulation should be bioequivalent to itself. Therefore, for the moment-based measures, if the test formulation is indeed bioequivalent to the reference formulation, then $d(Y_T; Y_R)$

should be very close to $d(Y_R; Y_R')$. It follows that if the criteria are functions of the difference (or ratio) between $d(Y_T; Y_R)$ and $d(Y_R; Y_R')$, bioequivalence is concluded if they are smaller than some pre-specified limit. On the other hand, for probability-based measures, if the test formulation is indeed bioequivalent to the reference formulation, as measured by $d(Y_R; Y_R')$, then comparison $d(Y_T; Y_R)$ should not be much larger. As a result, bioequivalence is concluded if the criterion based on the probability-based measure is larger than some pre-specified limit.

Chow et al. (2010) compared the moment-based criterion with the probability-based criterion for the assessment of bioequivalence or biosimilarity under a parallel group design. The results indicate that the probability-based criterion is not only much more stringent but also sensitive to small changes in variability. This justifies the use of the probability-based criterion for the assessment of biosimilarity if a certain level of precision and reliability of biosimilarity is desired.

1.4.2.4 Scaled versus Unscaled Criteria

Scaled criteria are usually referred to as criteria that are adjusted for the intra-subject variability of the reference product or for the therapeutic index. For example, the IBC criterion, to be discussed in Chapter 4, is adjusted, depending on the circumstances, either for a constant variance or for the within-subject variability. The PBC criterion is adjusted correspondingly and thereby becomes also a scaled criterion. Scaled criteria adjusting for the variability of the reference product do not penalize good generic or biosimilar products having smaller variability.

As indicated by the FDA, a drug product is considered a highly variable drug if its intra-subject coefficient of variation (CV) is higher than or equal to 30%. It should be noted that, by applying the regulatory criterion for average BE, it may be difficult to demonstrate bioequivalence or biosimilarity between highly variable test and reference drug products. Alternatively, Haidar et al. (2008) described a procedure using scaled average bioequivalence (SABE) for the assessment of bioequivalence for highly variable drug products. The procedure has been, in effect, adopted by the FDA for bioequivalence assessment of highly variable drug products. As a result, SABE has attracted much attention for possible application for the assessment of biosimilarity of follow-on biologics since biological products are usually highly variable.

1.4.2.5 Weighted versus Unweighted Criteria

Weighted criteria are aggregated criteria with different weights of each component (e.g., of the difference between means and of the variance components). For example, the three components of IBE (the difference between the means, the difference between the within-subject variances, and the

variance component for the subject-by-product interaction) may be considered to have differing weights. However, this could further complicate an already complicated criterion. In practice, it is a challenging question to select an appropriate weight for each component, which will then have an impact on the assessment of bioequivalence or biosimilarity. Besides, it is difficult to interpret the selected weights for each component since there are masking effects among differences in means and variance components (Chow, 1999). Note that assessments of biosimilarity assessments are based on the totality-of-the-evidence. The FDA seems to suggest a weighted criterion or weighted scoring system (across different functional areas or domains) for global similarity.

In summary, for the assessment of bioequivalence of small-molecule drug products, the FDA recommends aggregated, moment-based, scaled, and unweighted criteria based on relative change. This has led to SABE for average bioequivalence of highly variable drug products and also, earlier, to the criteria for IBE and PBE. For the assessment of biosimilarity, on the other hand, Chow et al. (2010) suggested a disaggregated, probability-based, scaled, and weighted criterion based on relative distance (the distance between "*T* versus *R*" and "*R* versus *R*") being considered. This has led to the development of the (totality) biosimilarity index for the assessment of biosimilarity and drug interchangeability, which are further discussed in Chapter 6 (biosimilarity) and Chapter 11 (interchangeability).

1.4.3 Biosimilarity versus Non-inferiority

As indicated in the 2012 FDA draft guidance on *Scientific Considerations in Demonstrating Biosimilarity to a Reference Product*, in some cases, a one-sided test (non-inferiority design) may be appropriate for comparing safety and effectiveness and also advantageous as it could generally allow for a smaller sample size than an equivalence (two-sided) design (see, e.g., Chow et al., 2008). The FDA draft guidance provided the following example. If doses of the reference product higher than those recommended in its labeling do not create safety concerns, then a one-sided test may be sufficient for comparing the efficacy of certain protein products. The FDA draft guidance indicated that it is generally important to demonstrate that a proposed product has no more risk in terms of safety and immunogenicity than the reference product. For this purpose, a one-sided test may also be adequate in a clinical study which evaluates immunogenicity or other safety endpoints, as long as it is clear that lower immunogenic or other adverse events would not have implications for the effectiveness of a protein product. For a non-inferiority design, the FDA draft guidance indicated that a non-inferiority margin should be pre-specified with scientific justification.

The approaches of non-inferiority, superiority, equivalence, and similarity will be presented in detail in Chapter 7.

Statistically, testing for non-inferiority includes testing for equivalence and testing for superiority. In practice, we may test for equivalence or test for superiority once the non-inferiority has been established. Thus, non-inferiority does not imply equivalence. It should be noted that testing for non-inferiority/superiority is often employed based on a one-sided test procedure at the 5% level of significance, which is equivalent to a two-sided test procedure at the 10% level of significance. In practice, it is suggested that a one-sided test procedure at the 2.5% level of significance should be applied for testing non-inferiority; it is equivalent to a two-sided test procedure at the 5% level of significance. Similarly, testing for superiority includes testing for equivalence and testing for non-inferiority. In other words, we may test for equivalence or test for non-inferiority if we fail to reject the null hypothesis of non-superiority. It should also be noted that superiority does not imply equivalence. In practice, it is also suggested that a one-sided test procedure at the 2.5% level of significance, which is equivalent to a two-sided test procedure at the 5% level of significance, should be used for testing superiority.

Since non-inferiority is regarded as one-sided equivalence, we may consider establishing *non-inferiority* first and then test for *non-superiority* for the assessment of biosimilarity by utilizing the concept of *asymmetric* equivalence limits (α). This proposal deals with distinct values of α_1 and α_2 rather than $\alpha_1 = \alpha_2$. This enables us to adopt flexible biosimilarity criteria. However, the selection of the non-inferiority margin and the choices of α_1 and α_2 are controversial issues. Consideration of spending functions could be helpful. In any case, consensus among the regulatory agency, pharmaceutical/biotechnology industry, and academia should be reached based on appropriate and valid scientific/statistical justification. More details regarding testing for non-inferiority versus testing for equivalence or similarity are given in Chapter 7.

1.4.4 Practical Issues

In practice, the following questions are often asked when assessing biosimilarity between biosimilars and an innovative drug product.

How similar is considered highly similar?—Current criteria for assessment of bioequivalence may, in some cases, be useful for determining whether a biosimilar product is similar to a reference product. However, they do not provide additional information regarding the *degree* of similarity. As indicated in the BPCI Act, a biosimilar product is defined as a product that is *highly similar* to the reference product. However, little or no discussion regarding the degree of similarity for achieving *highly similar* was provided. It may well be that, in addition to the demonstration of similarity, on the average, of a study endpoint, demonstration of similarity in variability of a study endpoint should also be considered for achieving *highly similar* as defined in the BPCI Act.

What criteria should be used for assessing biosimilarity?—As indicated earlier, several criteria for the assessment of similarity are available in the published regulatory guidelines/guidances and the literature. The question regarding what criteria should be considered for assessing biosimilarity has become interesting. However, no systematic comparisons have been undertaken among these criteria in terms of their relative advantages and limitations. In practice, it is of interest to investigate

1. Whether these criteria will lead to the same conclusion?
2. Which criterion is superior (or more efficient) in comparison with others for a fixed sample size?
3. Can these criteria translate to one another?
4. Which criterion is telling the truth?

Further research is needed in order to address these questions.

Is a one-size-fits-all criterion feasible?—The use of one size-fits-all criterion for bioequivalence assessment has been criticized in the past several decades. The major-criticism is that it ignores the variability associated with the response. In practice, it would be difficult, if not impossible, to demonstrate, with the usual criterion for average bioequivalence, that a test product is bioequivalent to a reference product if the reference product is highly variable. The one size-fits-all criterion is also criticized for penalizing good products having lower variability. Thus, it has been suggested that the one-fits-all criterion be flexible by adjusting for the intra-subject variability of the reference product and/or the therapeutic window whenever possible. This has led to the approach of the SABE criterion which can be applied to the assessment of bioequivalence for highly variable drug products. Since most biological products are considered highly variable, the application of SABE for assessment of biosimilar products is being studied (see, e.g., Zhang et al., 2013).

Should similarity in variability or distribution of response be considered?—As discussed earlier, the one size-fits-all criterion, based on the average response of the study endpoint, suffers from the following disadvantages:

1. It ignores the variability associated with the response
2. It may penalize good products with lower variability

The use of SABE for highly variable drugs is an attempt to fix the problem. In practice, it is of interest to establish similarity in variability or distribution for the response of the study endpoint for achieving the ultimate goal of high similarity (see, e.g., Chow and Liu, 2010). For this purpose, many authors have explored the potential application of IBE or PBE to assess biosimilarity

(see, e.g., Hsieh et al., 2010). Hsieh et al. (2010) suggested that the similarity in variability of the response of the study endpoint be evaluated because the assessment of similarity in variability is more stringent than that for assessing the biosimilarity in average and, consequently, a higher degree of similarity can be achieved.

What endpoints should be used for the assessment of biosimilarity?—As indicated in the BPCI Act, a biosimilar product should not only be highly similar to that of a reference product but also there should be no clinically meaningful differences in terms of the drug characteristics of safety, purity, and potency. Thus, an easy answer to this question would depend upon which good drug characteristics one would like to show high similarity. For example, if we are to show that there are no clinically meaningful differences in terms of safety and potency (efficacy), then clinical endpoints for safety and efficacy should be used for the assessment of biosimilarity.

Should a clinical trial always be conducted?—If one would like to show that the safety and efficacy of a biosimilar product are highly similar to those of the reference product, then a clinical trial may be required. In some cases, clinical trials for the assessment of biosimilarity may be waived if there is substantial evidence that surrogate endpoints or biomarkers are predictive of the clinical outcomes. On the other hand, clinical trials are required for the assessment of drug interchangeability in order to show that the safety and efficacy between a biosimilar product and a reference product are similar in any given patient of the patient population under study.

What if a biosimilar product turns out to be superior to the reference product?—It should be noted that superiority (including both statistical superiority and clinical superiority) is not biosimilarity. Thus, if a biosimilar product has been shown to be superior to the reference product, then it is suggested that it should be considered as a new biological product. Thus, it is a controversial issue that a biosimilar product should go through the lengthy regulatory review/approval process for similar indications if it is shown to be superior to the innovative product.

Is there a unified approach for the assessment of biosimilarity?—Chow et al. (2010) proposed a unified approach, which is referred to as the biosimilarity index, for the assessment of biosimilarity. The method of biosimilarity index is robust with respect to criteria for biosimilarity and the study design used. The proposed biosimilarity index can be extended to a totality biosimilarity index, which can be used to provide the totality-of-the-evidence across functional areas or domains for the assessment of biosimilarity as suggested in the FDA draft guidance on scientific considerations. More details regarding the development and application of the biosimilarity index for assessing biosimilarity can be found in Chapter 6.

1.5 Interchangeability of Biological Drug Products

As indicated in the Public Health Act Subsection 351(k)(4), that is, in Subsection (k)(4) of the BPCI Act, the term *interchangeable* or *interchangeability* in reference to a biological product means that the biological product may be substituted for the reference product without the intervention of the health care provider who prescribed the reference product. Along this line, in what follows, the definition and basic concepts of interchangeability (in terms of switching and alternating) are given.

1.5.1 Definition and Basic Concepts

As stated in the Public Health Act Subsection 351(k)(4), a biological product is considered to be interchangeable with the reference product if (1) the biological product is biosimilar to the reference product, and (2) it can be expected to produce the same clinical result in *any given patient*. In addition, for a biological product that is administered more than once to an individual, the risk in terms of safety or diminished efficacy of alternating or switching between use of the biological product and the reference product is not greater than the risk of using the reference product without such alternation or switch.

Thus, by the definition of the BPCI Act, there is a clear distinction between biosimilarity and interchangeability. In other words, biosimilarity does not imply interchangeability which is much more stringent. According to the BPCI Act, if a test product is judged to be interchangeable with the reference product, then it may be substituted, even alternated, without a possible intervention, or even notification, of the health care provider. However, interchangeability is expected to produce the *same* clinical result in *any given patient*, which can be interpreted as expecting the same clinical result in *every single patient*. In reality, conceivably, lawsuits may be filed if adverse effects are recorded in a patient after switching from one product to another.

It should be noted that when the FDA declares the biosimilarity of two drug products, it may not be assumed that they are interchangeable. Therefore, labels ought to state whether for a follow-on biologic which is biosimilar to a reference product, interchangeability has or has not been established. However, payers and physicians may, in some cases, switch products even if interchangeability has not been established.

1.5.2 Switching and Alternating

Unlike the interchangeability of small-molecule drug products (in terms of prescribability and switchability) (Chow and Liu, 2008), the FDA has slight perception of drug interchangeability for biosimilars. From the FDA's perspective, interchangeability includes the concepts of switching and alternating between an innovative biological products (R) and its follow-on

biologics (*T*). The concept of switching involves the switch from not only "*R* to *T*" or "*T* to *R*" (narrow sense of switchability) but also "*T* to *T*" and "*R* to *R*" (broader sense of switchability). As a result, in order to assess switching, biosimilarity for "*R* to *T*," "*T* to *R*," "*T* to *T*," and "*R* to *R*" needs to be assessed based on some biosimilarity criteria under a valid study design.

On the other hand, the concept of alternating is referred to as either the switch from *T* to *R* and then switch back to *T* (i.e., "*T* to *R* to *T*") or the switch from *R* to *T* and then switch back to *R* (i.e., "*R* to *T* to *R*"). Thus, the difference between "the switch from *T* to *R*" or "the switch from *R* to *T*" and "the switch from *R* to *T*" or "the switch from *T* to *R*" needs to be assessed for addressing the concept of alternating.

1.5.3 Study Design

For the assessment of bioequivalence for chemical drug products, a standard two-sequence, two-period (2 × 2) crossover design is often considered, except for drug products with relatively long half-lives. Since most biosimilar products have relatively long half-lives, it is suggested that a parallel-group design should be considered. However, the parallel-group design does not provide independent estimates of variance components such as inter-subject and intra-subject variabilities and the variability due to subject-by-product interaction. Thus, it is a major challenge for assessing biosimilarity and interchangeability (in terms of the concepts of switching and alternating) of biosimilar products under parallel-group designs.

For the assessment of switching, a switching design should allow the assessment of biosimilarity for the switch from "*R* to *T*," "*T* to *R*," "*T* to *T*," and "*R* to *R*" in order to determine whether there is a risk when a switch occurs. For this purpose, Balaam's 4 × 2 crossover design, that is, *TT, RR, TR, RT*, may be useful. Similarly, for addressing the concept of alternating, a two-sequence, three-period dual design, that is, *TRT, RTR*, may be useful since the designs allow the assessment of the switch from *T* to *R* and then back to *T*, that is, "*T* to *R* to *T*" and from *R* to *T* and then back to *R*, that is, "*R* to *T* to *R*." For addressing both concepts of switching and alternating for drug interchangeability of biosimilars, a modified Balaam's crossover design, that is, *TT, RR, TRT, RTR*, is recommended.

More details and further discussions regarding the design and analysis of drug interchangeability in terms of switching and alternating are given in Chapter 11.

1.5.4 Remarks

With small-molecule drug products, bioequivalence generally reflects therapeutic equivalence. Drug prescribability, switching, and alternating are generally considered reasonable. With biological products, however, variations are often higher (other than PK factors may be sensitive to small changes

in conditions). Thus, often only parallel-group design rather than crossover kinetic studies can be performed. It should be noted that very often, with follow-on biologics, biosimilarity does *not* reflect therapeutic comparability. Therefore, switching and alternating should be pursued with extreme caution.

1.6 Scientific Factors

Following the passage of the BPCI Act, in order to obtain input on specific issues and challenges associated with the implementation of the BPCI Act, the U.S. FDA conducted a 2 day public hearing on the *Approval Pathway for Biosimilar and Interchangeability Biological Products* held on November 2–3, 2010, at the FDA in Silver Spring, Maryland, United States. In what follows, some of the scientific factors and practical issues are addressed.

1.6.1 Fundamental Biosimilarity Assumption

Similar to the Fundamental Bioequivalence Assumption for the assessment of bioequivalence, Chow et al. (2010) proposed the following Fundamental Biosimilarity Assumption for follow-on biologics:

> *When a biosimilar product is claimed to be biosimilar to an innovator's product based on some well-defined product characteristics, it is therapeutically equivalent, provided that the well-defined product characteristics are validated and are reliable predictors of safety and efficacy of the products.*

For the chemical generic products, the well-defined product characteristics are the exposure measures for early, peak, and total portions of the concentration–time curve. The Fundamental Bioequivalence Assumption assumes that equivalence in the exposure measures implies therapeutical equivalence. However, due to the complexity of the biosimilar drug products, one has to verify that some validated product characteristics are indeed reliable predictors of safety and efficacy. It follows that the design and analysis of studies for the evaluation of similarity between a biosimilar drug product and an innovator product are substantially different from those for the chemical generic products.

1.6.2 Consistency in Manufacturing Process/Quality Control

Tse et al. (2006) proposed a statistical quality control (QC) method to assess a proposed index to test the consistency between raw materials (which are from different resources) and/or between final products manufactured by

different manufacturing processes. The consistency index is defined as the probability that the ratio of the characteristics (e.g., potency) of the drug products produced by two different manufacturing processes is within a pre-specified limit of consistency. The consistency index close to 1 indicates that the characteristics of the drug products from the two manufacturing processes are almost identical. The idea for testing consistency is to construct a 95% confidence interval for the proposed consistency index under a sampling plan. If the constructed 95% confidence lower limit is larger than a pre-specified QC lower limit, then we claim that the final products produced by the two manufacturing processes are consistent.

Let U and W be characteristics of the drug products from two different manufacturing processes, where $X = \log U$ and $Y = \log W$ follow normal distributions with means μ_X, μ_Y and variances V_X, V_Y, respectively. Similar to the idea of using $P(X < Y)$ to assess reliability in statistical QC (Church and Harris, 1970; Enis and Geisser, 1971), Tse et al. (2006) proposed the following probability as an index to assess the consistency between the two different manufacturing processes:

$$p = P\left(1-\delta < \frac{U}{W} < \frac{1}{1-\delta}\right),$$

where $0 < \delta < 1$ and is defined as a limit that allows for consistency. Tse et al. (2006) refer to p as the consistency index. Thus p tends to 1 as δ tends to 1. For a given δ, if p is close to 1, characteristics U and W are considered to be nearly identical. It should be noted that a small δ implies the requirement of high degree of consistency between the characteristics U and W. In practice, it may be difficult to meet this narrow specification for consistency. Tse et al. (2006) proposed the following QC criterion. If the probability that the lower limit $LL(\hat{p})$ of the constructed $(1 - \alpha) \times 100\%$ confidence interval of p is larger than or equal to a pre-specified QC lower limit, say, QC_L, exceeds a pre-specified number β (say $\beta = 80\%$), then we claim that U and W are consistent or similar. In other words, U and W are consistent or similar if $P(QC_L \le LL(\hat{p})) \ge \beta$, where β is a pre-specified constant.

1.6.3 Biosimilarity in Biological Activity

Pharmacological or biological activity is an expression describing the beneficial or adverse effects of a drug on living matter. When the drug is a complex chemical mixture, this activity is exerted by the substance's active ingredient or pharmacophore but can be modified by the other constituents. The main kind of adverse biological activity is a substance's toxicity. Activity is generally dosage-dependent and it is not uncommon to have effects ranging from beneficial to adverse for one substance when going from low to high doses. Activity depends critically on the fulfillment of the ADME (absorption, distribution, metabolism, and elimination) criteria.

Biosimilarity refers to comparisons between a reference product and a biosimilar product (the new EU "pharmaceutical review" legislation published on April 30, 2004, amended the EU community code on medicinal products to provide for the approval of biosimilars based on fewer preclinical and clinical data than had been required for the original reference product.)

The complexity of the protein and knowledge of its structure–function relationships determine the types of information needed to establish similarity.

1.6.4 Similarity in Size and Structure

In practice, various in vitro tests such as the assessments of the primary amino acid sequence, charge, and hydrophobic properties are performed to compare the structural aspects of biosimilars with their originator molecules. However, it is a concern whether in vitro tests can be predictive of biological activity in vivo due to the fact that there are significant differences in biological activity despite similarities in size and structure. Besides, it is difficult to assess biological activity adequately as few animal models are able to provide data that can be extrapolated for an accurate and reliable prediction of biological activity in humans. Thus, controlled clinical trials remain the most reliable means of demonstrating therapeutic similarity between a biosimilar molecule and the originator product.

1.6.5 Issues of Immunogenicity

The immune system consists of a diverse and complex set of cells and organs that have complicated interactions with each other and with other physiological systems. These complexities make the detection and evaluation of drug-induced immunogenicity difficult. The use of biosimilar products could have unwanted immune responses. An unwanted immune reaction could result in a clinical consequence of severe life-threatening conditions. Thus, the assessment of potential immunogenicity on the immune system is an important component of the overall evaluation of the safety (toxicity) of biosimilar products. However, although immunogenicity findings could indict a biosimilar product for some types of clinical investigations or certain indications, these findings appear to be rare (FDA, 2002).

Since all biological products are biologically active molecules derived from living cells and have the potential to evoke an immune response, immunogenicity is probably the most critical safety concern for the assessment of biosimilarity of follow-on biologics. The immune responses to biological products can lead to loss in efficacy and change in safety profile such as

1. Anaphylaxis
2. Injection site reactions
3. Flu-like syndromes
4. Allergic responses

The risk of immunogenicity can be reduced through stringent testing of the products during their development. More details regarding issues on immunogenicity are provided in Chapter 12.

1.6.6 Comparability/Consistency of Manufacturing Processes

Unlike small-molecule drug products, biological products are made of living cells. Thus, manufacturing of biological products is a very complicated process, which involves the steps of

1. Cell expansion
2. Cell production (in bioreactors)
3. Recovery (through filtration or centrifugation)
4. Purification (through chromatography)
5. Formulation

A small discrepancy at each step (e.g., purification) could lead to a significant difference in the final product, which might cause drastic change in clinical outcomes. Thus, process control and validation play an important role for the success of the manufacturing of biological products. In addition, since at each step (e.g., purification) different methods may be used at different biological manufacturing processes (within the same company or at different biotech companies), tests for consistency are necessarily performed. Note that at the step of purification, the following chromatography media or resin are commonly considered:

1. Gel filtration
2. Ion exchange
3. Hydrophobic interaction
4. Reversed phase normal phase
5. Affinity

Thus, at each step of the manufacturing process, primary performance characteristics should be identified, controlled, and tested for consistency for process control and validation.

Issues involving the comparability and the assessment of consistency for manufacturing processes are presented in Chapter 14.

1.6.7 Other Practical Issues

There are many critical attributes of a potential patient's response to follow-on biologics. For a given critical attribute, valid statistical methods are necessarily to be developed under a valid study design and a given set of criteria for