AAPS Advances in the Pharmaceutical Sciences Series 34

Hiten J. Gutka
Harry Yang
Shefali Kakar
Editors

Biosimilars

Regulatory, Clinical, and Biopharmaceutical Development

aaps®

Springer

AAPS Advances in the Pharmaceutical Sciences Series

Volume 34

Series Editor

Yvonne Perrie, Strathclyde Institute of Pharmacy, University of Strathclyde, Bearsden, Dunbartonshire, UK

The AAPS Advances in the Pharmaceutical Sciences Series, published in partnership with the American Association of Pharmaceutical Scientists, is designed to deliver volumes authored by opinion leaders and authorities from around the globe, addressing innovations in drug research and development, and best practice for scientists and industry professionals in the pharma and biotech industries.

More information about this series at http://www.springer.com/series/8825

Hiten J. Gutka • Harry Yang • Shefali Kakar
Editors

Biosimilars

Regulatory, Clinical, and Biopharmaceutical Development

Editors
Hiten J. Gutka
Oncobiologics Inc.
Cranbury, NJ, USA

Harry Yang
Statistical Sciences
MedImmune LLC
Gaithersburg, MD, USA

Shefali Kakar
Novartis (United States)
East Hanover, NJ, USA

ISSN 2210-7371 ISSN 2210-738X (electronic)
AAPS Advances in the Pharmaceutical Sciences Series
ISBN 978-3-319-99679-0 ISBN 978-3-319-99680-6 (eBook)
https://doi.org/10.1007/978-3-319-99680-6

Library of Congress Control Number: 2018961588

© American Association of Pharmaceutical Scientists 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This book is dedicated to my family (Sejal, Rishabh, Sushilaben, and Jayantibhai) for their support, unconditional love, and affection. The book is also dedicated to gurus and thought leaders who have inspired me and had a transformational influence on my life.

Hiten J. Gutka PhD

This book is dedicated to my family.

Harry Yang PhD

This book is dedicated to "Science is better without Politics."

Shefali Kakar PhD

Preface

A biosimilar is a biotherapeutic, intended to be highly similar to an already marketed reference product and approved through a regulatory process. Since the first biosimilar approved, over a decade ago, the biosimilar drug development has transformed. The regulatory perspectives have shifted with evolving science; however, the key principles have remained the same to demonstrate that the proposed biosimilar will not have any clinically meaningful differences in terms of its safety and efficacy, compared to the reference product.

A small molecule intending to be a generic product also undergoes a similar exercise; however, it is well established that for most small molecules, an identical molecule can be synthesized. This is both due to the simplicity of the molecule and the ability to characterize that the generic molecule is identical to the reference product. A battery of in vitro assays and a clinical pharmacokinetic bioequivalence study confirm biological equivalence (bioequivalence). The in vitro tests and the bioequivalence studies serve as a surrogate for the confirmation that the small-molecule generic will not have any meaningful differences in terms of its safety and efficacy to the small-molecule reference product. Given the complexity and heterogeneity of most biologics, the term identical is not well suited, as it is not possible for any two batches even from the same manufacturer to be "identical." The term equivalent is therefore replaced with "similar." The requirements for demonstration of this "similarity" are significantly more for a biologic, as compared to that for small molecules.

A book on biosimilars may seem premature in an ever-evolving field of transformational science and shifting regulatory policies. This book is therefore not intended to be a static exercise in reviewing biosimilar development but comprehensive overview to develop a biopharmaceutical, regulatory, and development strategy for future biosimilars. The key principles for developing a biosimilar are, and will remain, to ensure that the proposed biosimilar will have similar safety and efficacy as the reference product. What will continue to evolve is, however, the extent of evidence and the quality of evidence required to determine the above criteria.

The science of biosimilars is defined on principles often borrowed from new drug development. Most individuals entering the biosimilar field are already experts in a particular aspect of drug development. This book has been written with a focus for an individual looking to get a comprehensive overview of the biosimilar regulatory framework, and the development process, while also ensuring that the pharmaceutical professional in a field can learn to apply their expertise from novel drug development process to biosimilar development. The current book has painstakingly laid out the current state of the art in each of the "subsciences." Each chapter, however, is written as a stand-alone reference, should the reader need to focus on a particular aspect of the biosimilar paradigm. The development path of a biosimilar is just as unique as a development path of a new drug, tailored by the mechanism of action, the quality of the molecule, the published information on the reference product, the current competitive environment, the target market and regulatory guidances, and, most importantly, the emerging totality of evidence for the proposed biosimilar during development.

The more than 20 chapters in this book follow the journey of a biosimilar development. The book is divided in six parts each comprising of three to seven chapters. The first part comprising of four chapters lays the foundation of the biosimilars with an introduction to the biosimilars, economics, and intellectual property landscape for biosimilars. A key strategic decision facing a biosimilar developer is the target market for the biosimilar. This is largely due to the differences in the requirements for demonstration of "similarity" across the various regulatory authorities. Part II provides an overview of the current opinions from the US, Canadian, and European health authorities, to set the tone for requirements of a global biosimilar development program. Given the complexity and heterogeneity of most biologics, biosimilars must be systematically engineered to match the reference product. This is achieved by understanding the reference product to tailor the manufacturing process for both biosimilar drug substance and drug product by building in quality by design (QbD) (Part III). The foundation of the biosimilar assessment is in the analytical assessment of the molecule, which is captured in detail in Part IV. Part IV not only provides a rational similarity assessment approach but also discusses several aspects of key analytical test methods used in biosimilarity assessment. The results during this assessment form the basis of the residual uncertainty that will guide the clinical program, the totality of evidence, and the justification for extrapolation to other indications (Part V). The final destination in the journey is the global experiences in safety and pharmacovigilance (Part VI).
This book owes a special thanks to Dr. Carol Kirchoff the Past Chair of the American Association of Pharmaceutical Scientists (AAPS) Biosimilar Focus Group for the ideation and initiating the discussion for creation of the book and for her immense focus and exhaustive efforts to identify and follow through the subject matter experts and thought leaders for the specific scientific chapters. We would like to thank AAPS Biosimilar Focus Group leadership and members for being the community, where

the scientific debates, exchanges, and diverse passionate energy come together to move the science forward. Lastly, we are grateful to AAPS and Springer in helping us realize that the time for this book is now.

Cranbury, NJ, USA Hiten J. Gutka
Gaithersburg, MD, USA Harry Yang
East Hanover, NJ, USA Shefali Kakar

Contents

Part I Business, Health Economics and Intellectual Property Landscape for Biosimilars

1 Biosimilars 101: An Introduction to Biosimilars 3
Anna Rose Welch
The Path to Biosimilars: How We Got Here 4
What Are Biosimilars? The Three Tenets of Biosimilarity 5
Unpacking Biosimilar Development and the Regulatory Pathway 6
Biosimilar Regulatory Challenges 9
The Future of Biosimilars 16
References 19

2 Innovation, Patents and Biologics: The Road to Biosimilar Competition: Factors Influencing Investment, Business Decisions and Marketing of Biosimilars 23
Erwin A. Blackstone and Joseph P. Fuhr
Introduction 23
Biologics R&D Costs 24
Biosimilar Barriers to Entry 26
FDA Lag in Approvals 29
Economic Justification for Patents and Exclusivity 29
Data and Market Exclusivity 30
Patents as Barriers to Entry 31
Patent Litigation Before Application 32
Litigation Concerning the Patent Dance 32
Other Patent Litigation 36
Other Patent Issues 37
Interchangeability 37
Pricing and Market Opportunities 38
U.S. Market 39
Potential Gains from Biosimilars 41
Biosimilar Risks and Return on Investment 42

The Future of Biosimilars 43
Conclusion 44
References 44

3 The Changing US Reimbursement Landscape and Biosimilars 49
Molly Burich
Reimbursement Overview 50
Medicare Reimbursement and Biosimilars 52
Medicaid Reimbursement and Biosimilars 61
Commercial Reimbursement and Biosimilars 63
Spotlight on the European Experience 65
Reimbursement Landscape Moving Forward 70
References 72

4 Litigation-Related Issues Under the Biologics Price Competition and Innovation Act 75
Brian D. Coggio, Ron Vogel, and Tasha Francis
Introduction 76
The BPCIA and What It Amends 76
Types of Biosimilar Products/Naming Biosimilar Products 77
BPCIA Statute 78
BPICA Comparison to the Hatch-Waxman Act 84
The Hatch-Waxman Safe Harbor 84
Preparation for BPCIA Litigation 86
Litigation 90
IPRs Are Another Avenue to Challenge Biologic Patents 98
Conclusion 104
References 104

Part II Regulatory Aspects of Development and Approval for Biosimilars

5 Design and Implementation of Successful Regulatory Strategies in Biosimilar Development 107
Kamali Chance
Introduction 107
Countries of Interest/Competitive Intelligence 108
Product Target Profile 110
Biosimilar Regulatory Pathways 110
Chemistry, Manufacturing, and Controls 116
Nonclinical Studies 117
Clinical Studies 117
Bridging Studies 119
Pulling It All Together 120
Conclusion 120
References 121

6 Clinical Information Requirements for Biosimilar Biologic Drug Approvals in Canada ... 123
Bradley J. Scott and Jian Wang
Introduction ... 123
Regulatory Framework ... 125
Reference Biologic Drug ... 126
Development Paradigm ... 128
Non-Clinical Development ... 130
Clinical Development ... 130
Authorizing Indications (Extrapolation) ... 137
Switching and Interchangeability ... 139
Post-Market Monitoring ... 142
Conclusions ... 142
References ... 143

7 EU Perspective on Biosimilars ... 145
Pekka Kurki and Niklas Ekman
Introduction ... 146
Evolution of the EU Policy on Biosimilars ... 147
Guidelines for Biosimilars ... 148
Assessment of Biosimilars ... 149
Reference Product ... 151
Quality ... 151
Non–Clinical Development ... 153
Clinical Studies ... 155
Product Information ... 165
Post-Marketing Surveillance ... 167
References ... 167

Part III Biopharmaceutical Development and Manufacturing of Biosimilars (Chemistry and Manufacturing Controls (CMC), Quality by Design (QbD), Quality Target Product Profile (QTPP))

8 Biosimilars Drug Substance Development and Manufacturing: Effective CMC Strategy ... 173
Adriana E. Manzi and Michiel E. Ultee
Understanding the Reference Drug ... 174
Expression Systems and Clonal Selection ... 175
Development of Upstream and Downstream Processes and Scale-Up to Manufacturing ... 177
Effect of Raw Materials on Product Quality ... 179
Effect of Manufacturing Conditions on Product Quality and Control Strategy for Biosimilars ... 180
Implementing a Successful CMC and Analytical Strategy for the Development of Biosimilars ... 182

Conclusion 185
References 185

9 **QbD in Biopharmaceutical Manufacturing and Biosimilar Development** 187
Christina Vessely and Christopher Bussineau
Introduction 188
Application of QbD to Biosimilar Product Development 196
QbD and Analytical Development 212
QbD and Formulation Development 214
Conclusion 217
References 218

10 **Drug Product Considerations for Biosimilars** 221
Satish K. Singh, Susanne Jörg, and Hanns-Christian Mahler
Introduction 221
Development Strategies for Biosimilars Drug Products 223
Summary 254
References 255

Part IV Analytical Similarity Considerations for Biosimilars (Protein characterization, Biophysical Tools, Biological Assays, Statistical Methods)

11 **Principles of Analytical Similarity Assessment** 261
Kristof Vandekerckhove and Russell Reeve
Introduction 261
Choice of Test Materials 262
Selection and Ranking of Product Quality Attributes 275
Analytical Testing Plan 281
Assessment of Analytical Similarity Data 285
Forced Degradation 290
Conclusion 301
References 301

12 **Application of an Adaptive Analytical Characterization Strategy to Support Development and Approval of Biosimilars** 305
John P. Gabrielson, Jared A. Young, and Brent S. Kendrick
Introduction 306
Analytical Characterization of Protein Attributes 307
Lifecycle Management of Analytical Characterization Methods 314
Future Outlook 316
References 319

13 **Higher Order Structure Methods for Similarity Assessment** 321
Jared A. Young and John P. Gabrielson
Introduction 322

Current Industry Landscape ... 322
Industry Perception of HOS Characterization ... 324
Method Selection Considerations ... 326
Method Strategy ... 330
Conclusion ... 335
References ... 336

14 Protein Conformational Array Technology for Biosimilar Higher Order Structure Analysis ... 339
Xing Wang, Wen-Rong Lie, and Jehangir Mistry
Introduction ... 340
PCA Development ... 341
Applications in Biosimilar mAb HOS Comparability Analysis ... 343
Applications in Bioprocess Development ... 344
Applications in Formulation Development and Accelerated/Stress Condition Stability Testing ... 348
HOS Polymorphism Revealed with PCA Technology and its Relevance to Immunogenicity ... 352
The PCA Technology Adapted to Luminex xMAP® Multiplex Technology Platform ... 355
HOS Characterization of Biosimilar Rituximab with the Luminex Multiplex System ... 357
Conclusion ... 360
References ... 360

15 Protein Particulates and Biosimilar Development: Analytical Tools and Therapeutic Implications ... 363
Aaron B. Krueger and Matthew D. Brown
Introduction ... 364
Immunogenicity of Biotherapeutics: Evidence and Therapeutic Consequences ... 366
FDA Guidance on Subvisible Particles, Immunogenicity, and Biosimilar Development ... 368
Instrumentation and Characterization of Subvisible Particles (2–100 μm): Light Obscuration and Flow Imaging Methods ... 370
Instrumentation and Characterization of Submicron Particles: Nanoparticle Tracking Analysis and Dynamic Light Scattering ... 378
Novel Technologies for the Assessment and Characterization of Particles ... 381
Examples of Particle Characterization in Biosimilarity and Biocomparability Studies for Licensed Products ... 386
Conclusion ... 388
References ... 388

16 Biological Activity Assays for Antibody Therapeutics 393
Xu-Rong Jiang and Anthony Mire-Sluis
Introduction 394
Definition of Bioassay and Potency Assays 395
Types of Biological Activity Assays for Antibody Therapeutics 395
Biological Assay Selection Based on Their Therapeutic
Mechanism of Actions (MoA) 399
Biological Assays in Biosimilar Product Life-Cycle Management........ 402
Special Considerations of Biological Assays Comparing Innovative
Biologics and Biosimilar 408
Case Studies... 418
Concluding Remarks... 428
References... 429

**17 Statistical Considerations for Demonstration of Analytical
Similarity** 431
Harry Yang, Richard K. Burdick, Aili Cheng, and Richard O. Montes
Introduction 432
Background 432
Stepwise Approach 433
Analytical Similarity Assessment....................................... 434
EMA Statistical Approach... 435
FDA Statistical Approach 437
Alternative Approaches to Address Criticisms of FDA Approach 455
Concluding Remarks... 466
References... 466

**Part V Clinical Aspects of Biosimilar Development
(Clinical Trials, Immunogenicity, Extrapolation
and Interchangeability)**

**18 Comparative Clinical Studies for Biosimilars as Part
of a Stepwise Approach** 471
Thomas E. Gwise
Background 471
Biosimilar Evaluation Process 472
Analytical Similarity.. 473
Comparative Clinical Studies 474
Interchangeability... 483
Discussion... 484
References... 485

**19 Immunogenicity Assessment of Biosimilars:
A Multidisciplinary Perspective**...................................... 489
Paul Chamberlain and Pekka Kurki
Introduction 490
Immune Response to Therapeutic Proteins 491

Differential Factors for Biosimilar Candidates 492
Role of Non-clinical Evaluation 494
Clinical Study Design Considerations for the Biosimilarity Exercise 495
Extrapolation of Conclusions on Relative Immunogenicity to Different Therapeutic Indications 513
Switching and Interchangeability 519
Post-authorization Risk Mitigation 527
Future Perspectives 530
References 533

20 Interchangeability Study Design and Analysis 543
Harry Yang, Dongyue Fu, and Lorin Roskos
Introduction 544
Background 544
Interchangeability, Switching and Substitution 545
Interchangeability in the European Union 545
Interchangeability in the United States 546
Interchangeability Worldwide 546
Switching Study 547
Other Considerations 565
Concluding Remarks 567
References 569

21 Extrapolation 571
Peijuan Zhu and Andrej Skerjanec
Introduction 572
The Regulatory Requirement for Extrapolation 574
Scientific Considerations on Extrapolations 576
Case Study 1: Infliximab Biosimilar CT-P13, Biosimilar of Remicade® 582
Case Study 2: Adalimumab Biosimilar ABP501, Biosimilar of Humira® 592
Conclusions 596
References 598

22 Totality of Evidence and the Role of Clinical Studies in Establishing Biosimilarity 601
Hillel P. Cohen, William C. Lamanna, and Martin Schiestl
Totality of Evidence is the Regulatory Paradigm for Biosimilar Approval 602
Totality of Evidence is Product-Specific 605
Analytical Data 614
Non-clinical Animal Data 616
Clinical Studies 618

Confirmatory Safety and Efficacy Studies 619
Non-US Sources of Information Supporting the “Totality of Evidence” for Approved Biosimilars 620
Extrapolation 621
Future Outlook 625
References 626

Part VI Biosimilars: Global Development and Clinical Experience

23 Pharmacovigilance of Biosimilars: Global Experience and Perspective 631
Thomas Felix, Binakumari Patel, Brian D. Bradbury, and Gustavo Grampp
Introduction 632
Background on Pharmacovigilance 633
Importance of Pharmacovigilance 634
Importance of Pharmacovigilance in a Multisource Biologics Market 635
Components of Pharmacovigilance 637
Pharmacovigilance in Other Areas of Medicine 645
Conclusions 646
References 646

24 Development and Commercialization of Biosimilars in India: Current Regulatory and Clinical Experience 653
Anurag S. Rathore, Srishti Joshi, Ankita Bhargava, and Neh Nupur
Introduction 654
Overview of Biosimilars Regulations 659
Analytical Similarity of Biosimilar Products 664
Clinical Experience 669
Conclusions 671
References 671

25 Immunogenicity and Adverse Reactions to Biosimilar Erythropoietin Products in Thailand: The Significance of Science and Quality Driven Process for Approval 675
Kearkiat Praditpornsilpa, Anunchai Assawamakin, and Kriang Tungsanga
Introduction 676
Immunogenicity of Recombinant Human Erythropoietin 677
Risk Factors of Immunogenicity 679
Pharmacovigilance Study in Thailand 681
Product Quality Issues 683
Conclusion 688
References 689

Index 693

Contributors

Anunchai Assawamakin Department of Pharmacology, Faculty of Pharmaceutical Science, Mahidol University, Bangkok, Thailand

Ankita Bhargava Department of Chemical Engineering, Indian Institute of Technology, New Delhi, India

Erwin A. Blackstone Temple University, Philadelphia, PA, USA

Brian D. Bradbury Center for Observational Research, Amgen Inc., Thousand Oaks, CA, USA

Matthew D. Brown Malvern Panalytical, Westborough, MA, USA

Richard K. Burdick Burdick Statistical Consulting, Longmont, CO, USA

Molly Burich Boehringer Ingelheim, Pharmaceuticals Inc., Danbury, CT, USA

Christopher Bussineau Manufacturing and Regulatory, Biologics Consulting Group, Inc., Alexandria, VA, USA

P. Chamberlain NDA Advisory Services Ltd, Leatherhead, Surrey, UK

Kamali Chance BioSciencesCorp, Graham, NC, USA

Aili Cheng Pharmaceutical Sciences and Manufacturing Statistics, Pfizer, Andover, MA, USA

Brian D. Coggio Fish & Richardson P.C., New York, NY, USA

Hillel P. Cohen Sandoz Inc., Princeton, NJ, USA

Niklas Ekman Finnish Medicines Agency, Biotechnology, Helsinki, Finland

Thomas Felix Global Regulatory and R&D Policy, Amgen Inc., Washington, DC, USA

Tasha Francis Fish & Richardson P.C., Minneapolis, MN, USA

Dongyue Fu Statistical Sciences, MedImmune, LLC, Gaithersburg, MD, USA

Joseph P. Fuhr Widener University, Chester, PA, USA

Thomas Jefferson University College of Population Health, Philadelphia, PA, USA

John P. Gabrielson Elion Labs, KBI Biopharma, Inc., Louisville, CO, USA

Gustavo Grampp Corporate Product Quality, Amgen Inc., Thousand Oaks, CA, USA

Thomas E. Gwise FDA, Silver Spring, MD, USA

Susanne Jörg Lonza Pharma and Biotech, Drug Product Services, Basel, Switzerland

Xu-Rong Jiang BioVentures, AstraZeneca, Gaithersburg, MD, USA

Development Quality Biologics, AstraZeneca, Gaithersburg, MD, USA

Srishti Joshi Department of Chemical Engineering, Indian Institute of Technology, New Delhi, India

Brent S. Kendrick Elion Labs, KBI Biopharma, Inc., Louisville, CO, USA

Aaron B. Krueger KBI Biopharma Inc., Boulder, CO, USA

Pekka Kurki Clinical Immunology, University of Helsinki, Helsinki, Finland

William C. Lamanna Sandoz Biopharmaceuticals, Sandoz GmbH, Kundl, Austria

Wen-Rong Lie MilliporeSigma Corp., St. Charles, MO, USA

Hanns-Christian Mahler Lonza Pharma and Biotech, Drug Product Services, Basel, Switzerland

Adriana E. Manzi Atheln, Inc., San Diego, CA, USA

Anthony Mire-Sluis Development Quality Biologics, AstraZeneca, Gaithersburg, MD, USA

Jehangir Mistry MilliporeSigma Corp., St. Charles, MO, USA

Neh Nupur Department of Chemical Engineering, Indian Institute of Technology, New Delhi, India

Richard O. Montes Biosimilars Pharmaceutical Sciences, Pfizer, Lake Forest, IL, USA

Binakumari Patel Global Patient Safety & Labelling, Amgen Inc., Uxbridge, UK

Kearkiat Praditpornsilpa Division of Nephrology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Anurag S. Rathore Department of Chemical Engineering, Indian Institute of Technology, New Delhi, India

Russell Reeve Advisory Analytics, IQVIA, Durham, NC, USA

Lorin Roskos Clinical Pharmacology & DMPK, MedImmune, LLC, Gaithersburg, MD, USA

Martin Schiestl Sandoz Biopharmaceuticals, Sandoz GmbH, Kundl, Austria

Bradley J. Scott Biologics and Genetic Therapies Directorate, Health Canada, Ottawa, ON, Canada

Satish K. Singh Lonza Pharma and Biotech, Drug Product Services, Basel, Switzerland

Andrej Skerjanec Sandoz AG, Basel, Switzerland

Kriang Tungsanga Division of Nephrology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Michiel E. Ultee Ulteemit BioConsulting, LLC, Hillsborough, NJ, USA

Kristof Vandekerckhove Strategic Drug Development, IQVIA, Vilvoorde, Belgium

Christina Vessely Analytical and Formulation Development, Biologics Consulting Group, Inc., Alexandria, VA, USA

Ron Vogel Fish & Richardson P.C., New York, NY, USA

Jian Wang Biologics and Genetic Therapies Directorate, Health Canada, Ottawa, ON, Canada

Xing Wang Array Bridge Inc., St. Charles, MO, USA

Anna Rose Welch Biosimilar Development, VertMarkets Publication, Erie, PA, USA

Harry Yang Statistical Sciences, MedImmune, LLC, Gaithersburg, MD, USA

Jared A. Young Elion Labs, KBI Biopharma, Inc., Louisville, CO, USA

Peijuan Zhu Janssen Research and Development, Inc., Raritan, NJ, USA

About the Editors

Hiten J. Gutka is an Associate Director Formulation Development at Oncobiologics Inc. where he works on biosimilar drug product development. He worked on formulation and stabilization of novel insulin analogues at Thermalin Diabetes LLC. He did a co-op at MedImmune LLC, in the Formulation and Drug Product Development group. Hiten got his PhD in Pharmacognosy (Pharmaceutical Biotechnology track) from the University of Illinois at Chicago (UIC), College of Pharmacy. Dr. Gutka worked at Biocon, USV, and Reliance Biopharmaceuticals in India, on analytical and pharmaceutical development of biosimilar insulins and other biosimilars including cytokines and hormones. He got his Bachelor's in Pharmaceutical Sciences from the University of Mumbai, India, and Master's in Pharmaceutical Sciences from Mumbai University Institute of Chemical Technology, Mumbai, India. Hiten is a member of the American Association of Pharmaceutical Scientists (AAPS). He serves on the editorial board of the journal *mAbs*.

Harry Yang is Senior Director and Head of Statistical Sciences at MedImmune. He has participated in industry initiatives related to quality by design, analytical method development and validation, and sterile drug product bioburden testing. He has 23 years of experience in drug discovery, preclinical and clinical study design and analysis, translational sciences, biopharmaceutical development and manufacturing, and regulatory submissions. He has authored 5 statistical books and published 14 book chapters and more than 90 papers. He received his Bachelor's and Master's degrees in Applied Mathematics from Peking University and PhD in Statistics from the University of Pittsburgh.

Shefali Kakar is an Executive Director, Global Head Respiratory, in the Novartis Pharmacokinetic Sciences group. She is responsible for overseeing all aspects of pharmacokinetics sciences (ADME, PK/PD, and clinical pharmacology) for the respiratory portfolio ranging from early discovery to full development. Dr. Kakar has over 15 years of industrial experience and has been involved in discovery and development of over 40 small molecules and protein therapeutics, including biosimilars. Prior to joining Novartis, Dr. Kakar worked at Pfizer and also served as an Adjunct Faculty at the Brown University. She received her PhD in Pharmacology from the University of Michigan.

Part I
Business, Health Economics and Intellectual Property Landscape for Biosimilars

Chapter 1
Biosimilars 101: An Introduction to Biosimilars

Anna Rose Welch

Abstract The surging costs of biological medicines worldwide have necessitated the development of biosimilars. These highly similar versions of off-patent biological products entered the EU in 2006 and the U.S. in 2015. Unlike small molecule generic drugs, which are chemically identical to their originator products, biosimilar products cannot be identical because of the nature of biological molecules. However, as more than a decade of experience in Europe has demonstrated, the slight differences between a biosimilar and its originator do not result in clinically-meaningful differences in the drug's efficacy and safety. The complexity of biosimilars has led the major regulatory agencies to establish unique biosimilar regulatory protocols. In order to earn approval for these products, biosimilar developers must present thorough analytical characterization packages, pharmacokinetic and pharmacodynamic profiles, and comparative clinical trial data to eliminate any residual uncertainty. Beyond development and regulatory complexities, much of the fascination with biosimilars stems from ongoing efforts to establish unique commercialization blueprints, educate stakeholders, and collect and present real-world evidence from ongoing treatment and post-marketing "switching trials" to demonstrate biosimilars' safety and efficacy in everyday use. Varying healthcare and reimbursement frameworks worldwide have given rise to dynamic case studies highlighting the diversity of the burgeoning biosimilar market.

Keywords Biosimilarity · Analytical Similarity · Extrapolation · Switching · Interchangeability · Regulatory Framework · Reimbursement · Infliximab · Etanercept · U.S. FDA · EMA

A. R. Welch (✉)
Biosimilar Development, VertMarkets Publication, Erie, PA, USA
e-mail: anna.welch@jamesonpublishing.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_1

The Path to Biosimilars: How We Got Here

The U.S. Drug Price Competition and Patent Term Restoration Act of 1984, otherwise known as the Hatch-Waxman Act, ushered in the revolutionary small molecule generics industry (Milne and Cairns 2003). This legislation established the regulatory pathway for pharmaceutical companies to create identical, more cost-effective copies of small molecule brand name drugs that were facing patent expiration. Though uptake of these lower-cost pharmaceuticals progressed slowly in the first 10–15 years, The Association for Accessible Medicines (AAM) reveals that generic drugs now account for upwards of 90% of drugs dispensed in the U.S., but only 26% of total drug costs (AAM 2017). In fact, between 2007 and 2017 alone, small molecule generics saved the U.S. healthcare system roughly $1.7 trillion (AAM).

However, it's now the age of biologics. As research, scientific knowledge, and technology advanced, the global pharmaceutical industry began directing R&D efforts into the development of novel biological products derived from human, animal, or microorganism sources. These treatments include protein-based hormones, enzymes, monoclonal antibodies, vaccines, blood products, and gene and cellular therapies. The first biological products were approved in the 1980s (EC 2014).

But this evolution in drug development has, inevitably, introduced more complex manufacturing procedures and higher clinical development costs. In fact, an often-cited and widely debated 2014 study from The Tufts Center for the Study of Drug Development pegged the costs of drug development to be $2.6 billion (TCSDD 2014). However, a separate study published in the *Journal of Health Economics* found that this total increases to roughly $2.9 billion with the inclusion of post-approval R&D costs (DiMasi et al. 2016). Though these figures do not break down the costs in terms of small vs. large molecule/biologic drug development costs, it's impossible to overlook the impressive growth of the biologics market as a whole—and, with it, the drastic rise in healthcare costs.

According to Deloitte's (2016) *2016 Global Life Sciences Outlook*, biologics spending hit $289 billion in 2014 and is projected to reach $445 billion by 2019. By 2020, Evaluate Pharma's *World Preview 2014: Outlook to 2020* forecasts that biologics will account for 50% of the top 100 prescription product sales, a statistic many in the industry argue is indicative of a "biologics tidal wave" sweeping through the pharmaceutical and healthcare space (Evaluate 2014). As such, it was only a matter of time before there would be a need for a "pressure valve" to counter the rising costs of biologics and growing healthcare spends.

What Are Biosimilars? The Three Tenets of Biosimilarity

Biosimilars, which have been marketed in the EU since 2006 and gained a foothold in the U.S. in 2015, are expected to decrease healthcare costs like small molecule generics have. But unlike generics, which are chemically identical to the original brand product, the nature of biologic drugs dictates these products cannot and will not be molecularly identical to the previously marketed biologic (known as the reference product).

Three basic tenets must be met for a product to be regulated and approved as a biosimilar. As the European Medicines Agency (EMA 2014) outlines, the product must be a biologic. It must also contain a version of the active substance of the reference product (EMA). Finally, it must embody highly similar quality characteristics, biological activity, and safety and efficacy profiles to those of its reference product (EMA).

Despite semantic differences, definitions from several regulatory agencies, including the U.S. FDA, the EMA, Japan's Pharmaceutical and Medical Devices Agency (PMDA), and Brazil's Agência Nacional de Vigilância Sanitária (ANVISA) reveals notions of biosimilarity are comparable globally and operate within the three previously outlined tenets.

Take, for instance, the EMA definition of a biosimilar compared to ANVISA's and PMDA's definitions. In the EMA, a biosimilar is defined as:

> A medicinal product that contains a version of the active substance of an already authorized original biological medicinal product in the European Economic Area (EEA). Similarity to the reference medicinal product in terms of quality characteristics, biological activity, safety, and efficacy based on a comprehensive comparability exercise needs to be established (EMA 2014).

In addition to defining a biosimilar as "a biologic medicine with known biologic activity that contains no new molecules, already licensed in Brazil," ANVISA also emphasizes that the product go through "all the production steps (including formulation, vialing, freeze drying, labeling, packaging, storage, quality control, and biologic product lot release) (Tsuruta et al. 2015)".

Even more succinctly, the PMDA specifies that a biosimilar be "a biotechnological drug product developed by a different company to be comparable to an approved biotechnology-derived product of an innovator" (Tsuruta et al. 2015).

The U.S. FDA has also taken its own unique stance to defining a biosimilar. While the agency's definition is comparable to the other major regulators, the FDA took it a step farther by creating four classifications for biosimilars in terms of analytical similarity.

The FDA defines biosimilars or biosimilarity to mean that "the biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components and that there are no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product" (FDA 2017b).

Following comparative analytical characterization, a biosimilar is then assessed within a development-phase continuum and determined to embody one of four levels of similarity: insufficient analytical similarity, analytical similarity with residual uncertainty, tentative analytical similarity, or fingerprint-like analytical similarity (FDA 2016). The outcome of the comparative analytical characterization ultimately determines a company's next steps in proving biosimilarity.

For instance, a molecule that reveals insufficient analytical similarity may not be recommended for continued development via the biosimilar 351(k) pathway unless manufacturing process modifications are made that will eliminate the problematic differences. However, a biosimilar that demonstrates tentative analytical similarity—meaning it provides high confidence in analytical similarity—will allow a developer to move forward with the appropriate clinical studies to eliminate any residual uncertainty.

Since the release of the EMA and the World Health Organization's (WHO's) biosimilar guidelines in 2006 and 2009, respectively, a handful of other countries have established their own regulatory pathways for biosimilars. A majority of these pathways are closely aligned, if not identical to the EMA and/or WHO. However, it's important to note nomenclature for biosimilars within these guidelines varies by regulator. Terms that have often been used in place of "biosimilars" include "subsequent entry biologics," "biogenerics," "similar biological medicinal products," "follow-on proteins," "follow-on biologics," and "similar biotherapeutic products"(Wadhwa and Thorpe 2012).

Unpacking Biosimilar Development and the Regulatory Pathway

Understanding Small vs. Large Molecules

Despite their 10-plus-year tenure in the European market, there are still several basic misconceptions about what biosimilars are and how they are to be developed and regulated. But one thing that has been regularly reiterated throughout the pharmaceutical and healthcare industries is that biosimilars are not to be equated with small molecule generics—especially on a scientific and molecular level.

As their names imply, one primary difference between a *small molecule* pharmaceutical/generic and a *large molecule* biologic/biosimilar has to do with the size and complexity of the molecule in question. A small molecule typically weighs less than 1000 daltons (Da), while a large molecule is greater than 1000 Da. Take the small molecule aspirin, for instance, which is 180 Da and comprises 21 atoms. A monoclonal antibody can be 150-plus kDa and comprise 20,000-plus atoms (GaBI 2012a; Christl et al. 2017).

The differing size and complexity of these molecules is also reflected within manufacturing processes, administration methods, and immunogenic profiles. The

size of a small molecule enables it to be formulated into tablets or capsules, taken orally, and easily absorbed into the bloodstream (Bayer n.d.). There is also little to no chance that a small molecule will invoke immunogenicity, an immune response which could lead to severe adverse events for the patient or render the drug ineffective over a period of time (GaBI 2012a).

Biologics and biosimilars, on the other hand, are created out of living cells, for instance bacteria, yeast, and mammalian cells (Bayer n.d.). These cells are genetically modified to express a specific protein which will selectively bind to a specific disease target (e.g. cancer cell receptors) (Bayer). By nature, biologics are highly variable, and as such, are sensitive to manufacturing changes, heat, and cold—hence the importance of refrigeration and a well-developed cold chain for biologics. Unlike small molecules, which are can be taken orally because they are not sensitive to the acidic gastric environment, biologics drugs are administered via injection or intravenously. A biological molecule—the basic structure of which is a peptide bond—is highly susceptible to acidity, hence the need to bypass the digestive system and administer the drug directly into the bloodstream (Bayer). Another key difference between small and large molecules is that biologics hold the potential to be immunogenic. Large enough differences in molecular structure could potentially alter the drug's safety and efficacy over time.

Given these fundamental differences in development and molecular structures, there are also markedly different costs and development timelines to bring small and large molecules to market. For instance, the U.S. Federal Trade Commission has estimated it can cost anywhere from $1 million to $5 million over the course of three to five years to bring a generic drug to market (2009). But it can take as long as 8–10 years and cost $100 million to $200 million to bring a biosimilar to market (FTC 2009).

Small vs. Large Molecule Regulation in the U.S.

Though biosimilars and generics are quite different on a molecular level, the regulatory approval process is similar in one key way: both regulatory pathways are abbreviated to decrease the costs of development and encourage a quicker path to market.

For instance, in the U.S., a novel small molecule is approved under a New Drug Application (NDA). A pharmaceutical company must file an NDA providing data on safety and efficacy of the drug from clinical trials in all applied-for indications. But when it comes to a generic copy, a drugmaker filing an Abbreviated New Drug Application (ANDA) for a generic does not need to independently gather evidence of the safety and effectiveness (U.S. FDA 2017a). Instead, a generic drug relies on the safety and efficacy of the previously approved reference drug to prove it is the same. To be considered a generic, the drug must contain the same active ingredient(s), conditions of use, route of administration, dosage form, strength, and labeling as the reference product (U.S. FDA 2017a) An ANDA must also

demonstrate the generic drug is bioequivalent to its reference product using in vivo or in vitro testing, or both, depending on the product (FDA n.d.-a).

A company seeking approval for a novel biologic in the U.S. must file a Biologics License Application (BLA), which comprises manufacturing information, pre-clinical and clinical study data, and labeling information (FDA n.d.-b). Like a novel small molecule drug, most of the safety and efficacy data is gleaned from a large complement of human clinical trials.

A biosimilar is also approved using a BLA—however, via the FDA's 351(k) pathway, which was established in the Biologics Price Competition and Innovation Act (BPCIA) of 2009 specifically for the approval of biosimilars (FDA 2015). This pathway dictates that a biosimilar maker does not need to recreate all the same steps as the reference product sponsor. In order to demonstrate that a molecule is highly similar, a biosimilar sponsor must submit analytical studies showing molecular similarity to the reference product; animal studies, including a toxicity assessment; and one or more studies in at least one indication to demonstrate safety, purity, and potency (FDA n.d.-c). (One study must at least be an immunogenicity and pharmacokinetics [PK] or pharmacodynamics [PD] study) (FDA n.d.-c). But rather than testing the biosimilar in all indications for which the reference product is licensed, it is only necessary for a biosimilar to be tested in one. Common regulatory protocol requires a sponsor to investigate the candidate in a Phase 1 PK/PD study with healthy volunteers and a Phase 3 clinical study comparing its safety, efficacy, and immunogenicity to that of the reference product in one of the applied-for indications.

Biosimilarity Relies on "Totality of the Evidence"

In order to approve either a biologic and a biosimilar, a company needs to amass four types of data: analytical, nonclinical, clinical pharmacology, and clinical studies. Together, this data provides a thorough characterization of the molecule's structure, quality, and efficacy and forms the basis of the FDA's "totality-of-the-evidence" approach (FDA 2015).

The "totality-of-the-evidence" approach for biosimilars is best visualized by the popular two pyramid approach comparing the 351(a) and 351(k) pathway (as described in Chap. 8). The diagram depicts two pyramids, each representing the development of a novel biologic and a biosimilar. Each pyramid is broken into four sections representing the data needed for approval. Each section of these pyramids is sized differently to demonstrate in which area(s) the bulk of the development work falls. While the largest part of the reference product's pyramid is the clinical studies section, the analytical section of the biosimilar triangle is the largest. Clinical studies data is represented by the small tip of the biosimilar triangle, and it is the largest portion of the novel biologic's development triangle.

The FDA takes into account the quality and quantity of the different data, also referred to as the totality-of-the-evidence-approach, when making a determination

of a biosimilar's efficacy (FDA 2015). As the FDA's *Scientific Considerations in Demonstrating Biosimilarity To A Reference Product* guidance explains:

> A sponsor may be able to demonstrate biosimilarity even though there are formulation or minor structural differences, provided that the sponsor provides sufficient data and information demonstrating that the differences are not clinically meaningful and the proposed product otherwise meets the statutory criteria for biosimilarity (2015).

Arguably, the most important takeaway about the biosimilar development and regulatory process is that the biosimilar does not need to independently reestablish safety and efficacy. Requiring biosimilar makers to do so would not only lead to higher development costs, in turn eliminating incentive for developers, but it could also lead to higher overall drug prices and less market competition. The intricacies of the CMC, analytical, and manufacturing-related aspects of biosimilar development are discussed in-depth in Parts III and IV of this book.

Biosimilar Regulatory Challenges

A Brief History of Biosimilar Regulation

As data protection and key patents for some of the first marketed biologics began to expire, the EMA took the first steps toward establishing an abbreviated regulatory pathway for biosimilars. Following a mandate from the Committee for Medicinal Products for Human Use (CHMP), the Biosimilar Medicinal Products Working Party (BMWP) and the Biologics Working Party (BWP) established an abbreviated regulatory pathway for "similar biological medicinal products," or biosimilars (EMA 2013). The overarching regulatory guideline, which provides the nonclinical and clinical requirements for biosimilar products, was originally released in 2005 and revised in 2014 (EMA 2014). Over the past 11 years, the BWP has also produced guidances for biosimilar development in a handful of different product classes, including epoetins, filgrastims, insulins, growth hormones, alfa interferons, monoclonal antibodies, beta interferons, follitropins, and low-molecular-weight heparins (LMWHs) (EMA 2013).

Throughout 2017, the EU has held the record for the most biosimilar approvals and market launches. Since the first few biosimilars—somatropin, erythropoietin, and filgrastim—were approved in the EU in 2006, 2007, and 2008, respectively (Tsuruta et al. 2015), there have been more than 30 biosimilars approved and 26 launched in the EU as of July 2017 (Chocolis 2017).

In October of 2009, the United Nation's WHO expert committee on biological standardization took its own steps into the biosimilar space, launching its Guidelines on Evaluation of Similar Biotherapeutic Products (GaBI 2012b). Along with the EMA guidelines, WHO guidance has been a primary influence for many global regulatory agencies as they align biosimilar-related legislation with WHO's internationally accepted standards (GaBI 2012b). Though true harmonization amongst regulations and agency reviews has not been achieved, many of the established

biosimilar guidelines are closely aligned, if not identical to the EMA and WHO guidelines. Some of the countries that have established abbreviated biosimilar regulatory pathways include Australia in 2008 (Power 2013); Brazil in 2010 (Generics and Biosimilars Initiative Journal (GaBI Journal) 2012); Canada in 2009 (Welch 2016a); China in 2014 (GaBI 2014); India in 2012 (Loeb 2012), which was then revised in 2016 (Government of India 2016); Japan in 2009 (Ministry of Health 2009); South Korea in 2009 (Jeewon 2015); South Africa in 2012 (Leng et al. 2015); and Thailand in 2013 (Adcock and Homhuan 2016).

Four years after the EU opened its biosimilar pathway, the U.S. established the Biologics Price Competition and Innovation Act (BPCIA). This legislation, which was written into The Patient Protection and Affordable Care Act of 2009 (ACA), created the U.S.' own biosimilar regulatory pathway via section 351(k) of the Public Health Service Act (FDA 2015). However, it wasn't until 2015 that the U.S. saw the approval of its first biosimilar, Sandoz's Zarxio, a biosimilar of Amgen's Neupogen (filgrastim). Though FDA approvals continued at a relatively steady clip through 2017, only three biosimilars have been made available for patient use as of August 2017.

Controversial Regulatory Concepts

As is to be expected in any new industry, biosimilars have introduced a paradigm shift in terms of how they're developed and assessed by regulators. In addition, these new drugs have raised several key regulatory debates. The three most prominent debates in the last few years have centered around biosimilar extrapolation, switching, and interchangeability.

Extrapolation

One of the most complicated concepts for biosimilar stakeholders to grasp is extrapolation, a scientific rationale in which data garnered in one indication is used to justify approval in all other indications (Tesser et al. 2017). This critically important rationale, which will be thoroughly addressed in Chap. 21, is a major advantage of the biosimilar 351(k) pathway.

To date, companies seeking biosimilar approval have sought approval for all the reference products' indications (excluding orphan indications), even though comparative clinical trials have not been performed in each condition. In turn, regulators have relied upon the concept of extrapolation to approve these biosimilars for treatment in all applied-for indications. Though regulatory approval without comparative clinical data in all therapeutic areas has been controversial, it's hardly a novel concept in the biologics world. In fact, extrapolation has become routine for all biologics following major changes in their manufacturing process (EMA and EC 2017).

It's an overlooked and understated fact that even reference products exhibit clinically non-meaningful differences from batch to batch—especially given the variability of biologic molecules and the need to alter manufacturing processes. Following these process changes, it's customary for a biologics company to perform analytical and non-clinical characterization and comparability on the molecule to ensure it still lies within the approved limits and will not alter clinical performance (Gerrard et al. 2015). The biologic is not required to re-demonstrate safety and efficacy in clinical trials.

This reliance on analytical methods to demonstrate similarity following manufacturing changes essentially justifies biosimilar extrapolation. Because a biosimilar must arrive on the market at a lower price than the originator, it's imperative the costs of development remain manageable. A particularly important means of lowering development costs is to eliminate large comparative clinical trials in every indication as the originator.

One of the most integral parts of biosimilar development is the analytical characterization of multiple lots of the innovator product and the biosimilar. These chemical, physical, and biological comparisons reveal the primary amino acid sequence, the tertiary structure of each molecule, and the mechanism of action (Gerrard et al. 2015). The key to understanding extrapolation lies in recognizing that the structure of the protein informs the clinical performance of the drug (Gerrard et al.). In other words, should the structure of the biosimilar molecule be highly similar to the reference product, the biosimilar will work the same way in treating all the reference product's licensed conditions. Though no two biological molecules can be identical, the differences between a biosimilar molecule and its reference product are minimal. In fact, these differences are smaller than those that occur following manufacturing process changes in a biologic's life cycle (Gerrard et al.).

Despite the fact that extrapolation pertains to all biologics, it has been exceptionally challenging for physicians and patients to grasp the concept of approving a "highly similar but not identical" biologic for multiple conditions without seeing clinical data in all those indications. For instance, a 2016 survey of U.S. specialty physicians by the Biosimilars Forum revealed only 12% of 1201 physicians were comfortable with the concept of extrapolation (Cohen et al. 2016). But the authors of the survey also determined through a series of questions around extrapolation that many physicians were simply unfamiliar with the concept (Cohen et al.).

If the European experience has demonstrated anything, these numbers have already improved or are likely to improve with several more years of experience. When the first infliximab biosimilar Remsima arrived in the EU in 2013, the European Crohn's and Colitis Organization (ECCO) released a position statement on the use of biosimilars in inflammatory bowel disease (IBD). The organization's position was that the efficacy of a biosimilar in one indication would not necessarily carry over to a different indication (Danese and Gomollon 2013). As such, the ECCO articulated its desire to see biosimilars tested specifically in patients with IBD (Danese & Gomollon).

Similarly, a 2013 survey of physicians revealed a lack of awareness and confidence in biosimilar use in IBD (Danese et al. 2016). However, in 2016,

just three years later, the ECCO surveyed physicians again to determine their knowledge of and comfort with biosimilars. The results showed a staggering reversal of physicians' original concerns and hesitance to accept biosimilar use in IBD. A meager 19.5% of physicians had little-to-no confidence in biosimilars usage compared to the 63% that expressed caution in 2013 (Danese et al.).

Though the FDA and EMA were open to extrapolation from the start, the concept of extrapolation took a bit longer to be accepted by Health Canada, which, arguably, led to one of the most controversial regulatory decisions in the biosimilar space. In 2014, the agency approved Pfizer's Inflectra (infliximab, also known as Celltrion's Remsima) for the same rheumatology conditions as Remicade. But despite the EMA's approval a year prior for all indications, Health Canada took a different path and denied approving the biosimilar for gastrointestinal diseases. It took additional data from Pfizer and roughly two years before the agency approved the biosimilar for Crohn's disease and colitis in June 2016. Though the agency eventually reversed its original decision against extrapolation for Inflectra, this example goes a long way in demonstrating how challenging extrapolation has been for patients, physicians, and regulators alike.

Switching vs. Interchangeability

Another complex and controversial topic related to biosimilars is interchangeability, which is addressed in Chap. 20. As was discussed in the first part of this introduction, biologics, by nature, have the potential to be immunogenic, meaning they could be recognized by the immune system as a threat. In such cases, the immune system may choose to attack the drug, causing mild to severe adverse events in patients and/or rendering the drug ineffective at treating the patient. Because it can be a challenge to find an effective treatment for chronic conditions, physicians, patients, and regulators have approached the prospect of switching patients from the reference product to the biosimilar with great caution. As the biosimilar industry took root abroad and in the U.S., it was not unusual to hear doctors specify they'd prefer to put newly diagnosed patients on the biosimilar, while keeping current biologics patients on the reference product.

But it's important to understand the concepts of switching and interchangeability, especially considering that these terms have led to such great stakeholder confusion. Switching is used to describe a single switch from the reference product to the biosimilar (Amgen n.d.). In many cases, this switch is overseen by a physician. However, as U.S. payers have begun adopting biosimilars and, in some cases, prioritizing them on formularies, doctors and patients have created the term "nonmedical switching" to express their concerns about switching treatments for financial reasons. Many patients have expressed frustration that an insurance company can trump the doctor-patient relationship in treatment decisions.

Though there are some lingering concerns about switching, there has been growing acceptance of the concept thanks to the publication of several switching

studies. One notable example is the NOR-SWITCH study, which was sponsored by the Norwegian government. The goal of the study was to determine the impact of a single switch from Remicade to Remsima in patients with rheumatoid arthritis, spondyloarthritis, psoriatic arthritis, ulcerative colitis, Crohn's disease, and chronic plaque psoriasis. As results released in October 2016 reveal, switching half of the 500 patients to biosimilars did not lead to significant safety or efficacy concerns in any of the six indications (Celltrion 2016).

However, it's important to differentiate between the single switch and multiple switches between a biosimilar and an originator that would likely occur as patient health plans and payer formularies change in response to price competition. Though a biosimilar may be switched for the reference product once in the course of treatment and be safe, stakeholders still express hesitancy about multiple switches back and forth between a biosimilar and reference product. Given the immunogenic potential of biologics, there are concerns that multiple switches could, theoretically, impact the long-term efficacy of that biologic treatment.

Section 351(k)(4) of the PHS Act granted the FDA authority to determine if a biosimilar is interchangeable with its reference product (FDA 2017b). In its 2017 interchangeability draft guidance, the FDA defines an interchangeable product as a "biosimilar to an FDA-approved reference product . . . [that] can be expected to produce the same clinical result as the reference product in any given patient." In other words, physicians must be certain that switching between the biosimilar and its reference product will be just as safe and effective as keeping the patient on the reference product without any switches. Section 351(i) of the PHS Act also designates that any biosimilar meeting these requirements for interchangeability "may be substituted [at the pharmacy level] for the reference product without the intervention of the health care provider who prescribed the reference product" (FDA 2017b). In order to be deemed interchangeable, however, the FDA has stipulated the biosimilar must undergo additional testing and clinical trials.

The U.S.' move to regulate interchangeability garnered quite a bit of attention. Since small molecule generics hit the market, they have been considered interchangeable and substituted at the pharmacy level without attaining additional data and a second regulatory approval. So not only does the FDA's decision to regulate interchangeability mark another striking departure from the small molecule space, but it also branches away from the EMA's approach to interchangeability. The EMA chose not to take a stance on interchangeability. Rather, each EU country has been granted the right to determine its own stance on biosimilar interchangeability, switching, and substitution.

Indeed, in the EU, the term interchangeability generally means a biosimilar can be expected to produce the same clinical effect as the reference biologic. Interchangeability is very rarely linked to automatic substitution at the pharmacy level. So far, only several European countries have ruled biologics, including biosimilars, are legally substitutable, or substitution is applied to biological medicines, including biosimilars, at the retail level. These countries include the Czech Republic, Estonia, France (still pending for naïve patients only), Latvia, Lithuania, the Netherlands, Poland, Portugal, and Turkey (Medicines for EU 2017). In the cases of Latvia,

Lithuania, Poland, and Turkey, prior permission from the physician is not required for substitution, and the pharmacist is required to notify the patients that a switch has occurred; however, physicians and patients can refuse the substitution (Medicines for EU).

After a decade of successful biosimilar use, interchangeability (not automatic substitution) has become a more readily accepted concept amongst regulators in the EU. In a widely publicized opinion piece in *BioDrugs* in April 2017, members of several different regulatory agencies (in Norway, Finland, the Netherlands, and Germany) argued biosimilars are interchangeable. Following an exploration of the theoretical risks of switching, the regulators concluded, "On the basis of current knowledge, it is unlikely and very difficult to substantiate that two products, comparable on a population level, would have different safety or efficacy in individual patients upon a switch. Our conclusion is that biosimilars licensed in the EU are interchangeable "(Kurki et al. 2017)."

In order to arrive at (hopefully) similar conclusions about interchangeability in the U.S., the FDA has released a draft guidance, "Considerations in Demonstrating Interchangeability With a Reference Product." In this guidance, the FDA spells out the additional data necessary to receive an interchangeability designation. For one, the FDA requires switching studies (at least three switches) with primary endpoints measuring PK/PD, as these assessments are more sensitive to changes in immunogenicity than efficacy endpoints (FDA 2017b). [The switching study also must take into account immunogenicity and safety (FDA 2017b).] The FDA also specifies a U.S.-licensed reference product be used in the switching study (FDA 2017b).

In addition, a large portion of this guidance homes in on the biosimilar's presentation. After all, should a biosimilar become substitutable at the pharmacy level, patients will be faced with a slightly different device. This could lead to increased risk of human error and, worst-case-scenario, adherence issues. Sponsors are expected to evaluate the differences in their presentation using a threshold study, and should any major differences in design pose concerns, a company may need to glean data from comparative-use human factors studies (FDA 2017b).

Lastly, the agency also addressed the use of post-marketing data for a biosimilar—particularly those with more complex molecular structures. Though real-world evidence (RWE) may not provide the more sensitive PK/PD information, this data, on a case-by-case basis, could be provided along with switching studies to demonstrate interchangeability (FDA 2017b).

The release of the FDA's guidance was a highly anticipated and celebrated achievement amongst a number of stakeholders, especially patients and physicians. However, it's also crucially important to note that an interchangeability designation does not mean an interchangeable biosimilar is of higher quality, more similar, or superior to a non-interchangeable biosimilar. Similarly, it's also easy to assume a biosimilar approved as interchangeable will be immediately available for substitution at any pharmacy in the U.S. However, in order for this to be the case, each individual state must have approved legislation permitting the substitution

of interchangeable biosimilars. According to the National Conference of State Legislatures (2018) 45 states (more current) and Puerto Rico have signed laws permitting substitution. Many of these laws stipulate substitution may not occur if the physician has specified the brand is medically necessary. In addition, depending on the state, patients and their doctors must be notified of the substitution.

The Issue of Harmonization

This discussion of interchangeability, and the different ways it has been addressed in the U.S. and across the EU, highlights a particular challenge facing the biosimilar space. Though the concept of biosimilarity is well-understood by all major regulatory agencies, occasionally differences may arise in scientific interpretation of a biosimilar development program or application. Regulatory agencies may have different ideas about the adequacy of biosimilarity margins (Welch 2017), one-vs.-two-assay approaches (Welch 2017), the need for animal studies, local patient inclusion in clinical trials, or the use of locally-sourced comparator products (Welch 2016b). If there is not enough communication between the company and the regulator(s), a company could find it needs to obtain additional data in order to receive approval in different countries.

But the industry has begun to see movement toward greater alignment among major regulatory agencies. In a keynote address at the DIA Biosimilars event in October 2016, the FDA coined the term "scientific alignment" to describe the move toward scientific unity amongst regulators. The FDA avoids using the term "harmonization" because this would imply guidelines and regulatory documents are identical amongst all nations, and this is highly unlikely to occur.

One strategy established to bolster consensus among agencies is The Biosimilars Cluster. Launched in 2011, the cluster now includes the FDA, EMA, Health Canada, and the PMDA (EMA n.d.-a). This group meets several times a year to discuss development challenges and the scientific and regulatory issues that may arise as agencies begin receiving applications for the same candidates.

In addition to the cluster, the FDA and EMA have established a program to provide parallel scientific advice to sponsors. Through this program, a sponsor is allowed to request a parallel review between the agencies in order to address a specific question or issue that may arise within a development program (EMA and FDA 2017). Many of these questions may arise because there are no existing guidelines or there are differences between the two agencies' guidelines (EMA & FDA). The aim of this program is to encourage communication between the FDA and EMA and promote sharing information and perspectives. It's also expected to provide the sponsor with a clearer sense of the requirements and an understanding of any differences in opinion about moving the development process forward.

The Future of Biosimilars

Why the Biologics Market Needs Biosimilars

The advent of innovative biologic treatments has changed and continues to drastically change the treatment landscape for patients suffering from chronic illnesses, cancers, and rare diseases. But this innovation has placed immense cost pressure on healthcare systems worldwide, in turn limiting patient access to these important medications—especially in emerging and developing markets. According to the QuintilesIMS Institute's *Outlook for Global Medicines Through 2021* report, global spending on medicine is forecasted to reach $1.5 trillion by 2021, with a majority of that growth being incurred by oncology, autoimmune, and diabetes medicines (2016a).

Many of the first biosimilars on the market are for biologics that have been consistently ranked within the top 10 in prescription or specialty drug spending. Two reports from pharmacy benefits manager Express Scripts and research company Evaluate Pharma (EP) include Humira, Enbrel, Rituxan, Remicade, Avastin, Herceptin, and Lantus in their respective lists of top drugs reimbursed through health insurance exchanges (Express Scripts 2016) and in 2016 worldwide sales (Urquhart et al. 2015).

According to the EMA's European public assessment reports, 36 biosimilars have been approved in the EU in several therapeutic areas as of September 2017: erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), human growth hormone (HGH), anti-tumor necrosis factor (anti-TNF), fertility (follitropin alfa), insulin, and oncology (EMA n.d.-b). The launch of biosimilars in several of these therapeutic categories has led to reductions in prices of reference products, in turn lowering the costs of daily treatment. Portugal, for instance, has seen price-per-treatment-day decrease by as much as 66%, thanks to the arrival of EPO biosimilars, while Romania saw about 62% decreases in price following the arrival of G-CSF biosimilars (QuintilesIMS 2017). As can be expected, the decrease in costs and establishment of market competition has led to increased access for patients. Between 2006 and 2014, biosimilars bolstered patient access to biologic treatments by 44% in the five EU nations, France, Germany, Italy, Spain, and the U.K. (Medicines for EU 2016).

There are a wide range of well-known and bullish predictions about the impact biosimilars will have on overall healthcare spending, both in the EU and the U.S. In the European market, if biosimilars were to reduce eight originators' prices-per-treatment-day by 30%, the EU would realize savings of roughly EUR15 billion within the next five years (QuintilesIMS 2016b). One of the most oft-cited predictions is from the RAND Corporation, which forecasted biosimilars will promote a $54 billion reduction in U.S. biologics spending from 2017 to 2026 (Mulcahy et al. 2017). Express Scripts came forward with even higher expectations for the period 2014–2024, predicting savings of $250 billion following the entrance of 11 biosimilars onto the market (2013).

Biosimilar Uptake Successes

However, realizing these savings will come down to the establishment of reimbursement schemes by governments, as well as methods of incentivizing physicians and patients to use biosimilars. In fact, though many countries may be aligned from a regulatory standpoint with the EMA's guidelines, each country has a different approach to pricing, distributing, and, in many cases, promoting switching of biosimilars. As such, biosimilar commercialization strategies often cannot be carried over and applied from country to country. But as Chap. 3 emphasizes, taking a deeper look into the pricing and market access policies implemented across the EU can lead to some valuable takeaways.

One of the most notable examples of uptake in the EU occurred in Denmark, Finland, and Norway, thanks to their winner-takes-all tender systems. For instance, in 2015, Norway stunned the biosimilar industry when Orion Pharma won the national infliximab tender for Remsima—Celltrion's infliximab biosimilar—by offering a 69% discount. In a presentation at the 2017 Global Biosimilars Congress in London, the Norwegian Medicines Agency's Dr. Steiner Madsen revealed that the country has continued to see increasing uptake of biosimilars, with infliximab garnering roughly 95% of market share and etanercept (biosimilar to Amgen's Enbrel) around 85% of market share (as of September 2017). Similar rates of uptake have been seen in Denmark and Finland, where biosimilar infliximab has captured 96% and 88% of these markets, respectively (as of May 2016) (Welch 2016c).

It's also important to note the role physicians have played in Norway's procurement process. The Norwegian Drug Procurement Cooperation (LIS) selects medicines for the tenders that state hospitals use for procurement (QuintilesIMS 2016b). Physicians and key opinion leaders (KOLs) from the country's four healthcare regions review clinical and cost-related aspects of a drug and are responsible for determining which drugs are awarded the tender (QuintilesIMS). Because of its physician-centricity, this system is well-trusted and, in turn, has encouraged hospital physicians to drive the biosimilar push within their own organizations (QuintilesIMS).

The National Health Service (NHS) of England has also been an interesting case study because of its efforts to bolster biosimilar usage in the country. In September 2017, the NHS released a new commissioning framework for biologics and biosimilars highlighting the savings potentials biosimilars could offer the health system. According to the framework, the NHS could realize savings of 200 million to 300 million pounds per year by 2020/2021 should patients be prescribed the "best value biological medicine" (NHS 2017). The ultimate goal is to ensure 90% of new patients receive the best value biologic (whether it be the reference product or the biosimilar) within 3 months of the launch of a biosimilar (NHS). After 12 months, the NHS hopes to see 80% of existing biologics patients placed on the best value medicine (NHS).

So far, the NHS has seen 44% uptake of both infliximab and etanercept biosimilars (NHS 2017). In fact, 80% of infliximab patients and 58% of etanercept

patients are now being treated with these biosimilars, which, in turn, has saved the NHS 160 million pounds per year (NHS). This percentage of uptake, however, is highly varied by region. For instance, while one central London trust saw 25% uptake of infliximab, a trust 16 miles away had 99% uptake (NHS). As such, there is still work to be done throughout the country to ensure greater savings and biologics access.

Reimbursement Strategies and the Importance of Incentives

The U.K. has also been recognized for its use of a system known as gainsharing. This strategy has been a particularly impactful tool for the clinical commissioning groups (CCGs) providing primary care throughout the U.K. Gainsharing incentivizes prescribers to choose the best-value medicines, the savings from which are then shared amongst stakeholders and reinvested in patient care.

It's impossible to overlook two other key strategies being explored abroad to bolster biosimilar uptake. These efforts include the quota system in Germany, for example, which requires physicians to place a specific percentage of patients on biosimilars (QuintilesIMS 2016b), and the new multi-winner tender system launched in Italy, which is to be used when drugs have more than three competitors on the market (Aideed 2017).

One of the biggest takeaways from the reimbursement landscape abroad is the importance of government involvement in promoting biosimilar use. The government must put forth the necessary policies, guidelines, and systems for incentivizing stakeholders. The implementation of the NOR-SWITCH study by the Norwegian government is a good example of how a government can invest in the generation of real-world data that ultimately will reinforce the safety and efficacy of biosimilar medicines in different patient populations.

There has been more positive movement in EU countries that have been less active in the biosimilar arena. For instance, Ireland, with only 11 biosimilars reimbursable by the state, has low biosimilar uptake compared to other EU nations (Department of Health 2017). However, the country's Department of Health released a National Biosimilar Medicines Policy Consultation Paper in August 2017. This paper outlines the goals of building a framework to increase confidence in and use of biosimilars in Ireland and asks for respondents' recommendations in determining the necessary "mix of policy levers" to promote biosimilar use (Department of Health).

Belgium is another country which has demonstrated slow progress growing its biosimilar market. As one publication from 2014 highlighted, market shares in Belgium were essentially 0% (Dylst et al. 2014). However, the authors concluded biosimilar uptake could be improved through more education, prescription quotas, hospital finance reform, patient registries, and, ultimately, more government and biosimilar company measures to reduce biosimilar uncertainty (Dylst et al.). In fact, since the release of this publication, The Federal Agency for Medicines and Health

Products in Belgium has announced its intentions to launch a biosimilar information campaign and has acknowledged the safety of switching (Medicines for EU 2017). The country also expects the use of gainsharing and tendering will be the key to bolstering uptake (Medicines for EU). Though it's still too early to determine the success of these measures, the latest communication from the Minister of Health encouraged hospitals to quickly begin the tender process following originator patent expirations (Medicines for EU).

As the experiences in Europe have demonstrated, there are many ways a country can promote the use of biosimilars. Though many of these experiences cannot be extrapolated to the U.S. healthcare system, Chap. 3 provides a comprehensive discussion of the important work being done to build a sustainable reimbursement system for biosimilars in the U.S. Ultimately, the most successful efforts to promote biosimilar uptake will involve ongoing discussions amongst all key stakeholders; payer, patient, and physician education; the appropriate discounts, reimbursement strategies, and incentives; and the use of government muscle to put the necessary biosimilar-friendly systems in place.

References

Adcock A, Homhuan A. Biosimilars in Thailand. 2016. http://www.tilleke.com/resources/biosimilars-thailand.

Aideed H. Highlights from the 15th Biosimilars Medicines Conference: how biosimilar procurement strategies could impact the sustainability of healthcare systems and of the biosimilar industry. Blog, Decision Resources Group [DRG]. 2017.

Amgen Biosimilars. The use of biosimilars: switching. n.d.

Association for Accessible Medicines. Generic drug access & savings in the U.S. 2017.

Bayer Pharmaceuticals. Small and large molecules. n.d.

Celltrion Healthcare. Celltrion Healthcare: Norwegian government's independent NOR-SWITCH study supports switching to biosimilar infliximab from originator. 2016.

Chocolis OM. Trends in the biosimilars market. White paper. Decision Resources Group. 2017.

Christl L, Woodcock J, Kozlowski S. Biosimilars: the U.S. regulatory framework. Annu Rev Med. 2017;68:243–54.

Cohen H, Beydoun D, Chien D, Lessor T, McCabe D, Muenzberg M, et al. Awareness, knowledge, and perceptions of biosimilars among specialty physicians. Adv Ther. 2016;33(12):2160–72.

Danese S, Fiorino G, Michetti P. Changes in biosimilar knowledge among European Crohn's Colitis Organization [ECCO] members: updated survey. J Crohn's Colitis. 2016;11(1):1362–5.

Danese S, Gomollon F. ECCO position statement: the use of biosimilar medicines in the treatment of inflammatory bowel disease (IBD). J Crohn's Colitis. 2013;7(7):586–9.

Deloitte. 2016 Global life sciences outlook. 2016.

Department of Health (Ireland). National biosimilar medicines policy: consultation paper 2017. 2017.

DiMasi J, Grabowski H, Hansen R. Innovation in the pharmaceutical industry: new estimates of R&D costs. J Health Econ. 2016;47:20–33.

Dylst P, Vulto A, Simoens S. Barriers to the uptake of biosimilars and possible solutions: a Belgian case study. PharmacoEconomics. 2014;32(7):681–91.

European Commission [EC]. What you need to know about biosimilar medicinal products: a consensus information document. 2014.

European Medicines Agency. Biosimilar medicinal products, brochure. 2013.

European Medicines Agency. Guidelines on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. 2014.

European Medicines Agency [EMA]. Cluster activities. n.d.-a.

European Medicines Agency [EMA]. European public assessment reports. n.d.-b.

European Medicines Agency [EMA] & European Commission [EC]. Biosimilars in the EU: information guide for healthcare professionals. 2017

European Medicines Agency [EMA] & U.S. Food and Drug Association [FDA]. General principles EMA-FDA parallel scientific advice (human medicinal products. 2017.

Evaluate Pharma. World preview 2014: outlook to 2020. 2014.

Express Scripts. The $250 billion potential of biosimilars. 2013.

Express Scripts. 2016 drug trend report. 2016.

Federal Trade Commission. Emerging healthcare issues: follow-on biologic drug competition. 2009.

Food and Drug Association. Scientific considerations in demonstrating biosimilarity to a reference product: guidance for industry. 2015.

Food and Drug Association. Clinical pharmacology data to support a demonstration of biosimilarity to a reference product: guidance for industry. 2016.

Food and Drug Association. Referencing approved drug products in ANDA submissions: guidance for industry. 2017a.

Food and Drug Association. Considerations for demonstrating interchangeability of with a reference product: guidance for industry. 2017b.

Food and Drug Association. Product-specific guidances for generic drug development. n.d.-a.

Food and Drug Association. Biologics License Applications (BLA) process (CBER). n.d.-b.

Food and Drug Association. Information for industry (biosimilars). n.d.-c.

Generics and Biosimilars Initiative Journal (GaBI Journal). Brazilian guidelines for biosimilars. 2012.

Generics and Biosimilars Initiative Journal (GaBI Journal). China releases draft biosimilars guidance. 2014.

Generics and Biosimilars Initiative Online [GaBI]. Small molecule versus biological drugs. 2012a.

Generics and Biosimilars Initiative Online [GaBI]. Global guidelines for biosimilars. 2012b.

Gerrard T, Johnston G, Gaugh D. Biosimilars: extrapolation of clinical use to other indications. Generics Biosimilars Initiat J. 2015;4(3):118–24.

Government of India, Department of Biotechnology, & Central Drugs Standard Control Organization [CDSCO]. Guidelines on similar biologics: regulatory requirements for marketing authorization in India. 2016.

Jeewon J. Korean regulations for biosimilars. Generics Biosimilars Initiat J. 2015;4(2):93–4.

Kurki P, van Aerts L, Wolff-Holtz E, Giezen T, Skibeli V, Weise M. Interchangeability of biosimilars: a European perspective. BioDrugs. 2017;31(2):83–91.

Leng H, Mutoti K, Mbelle N. Regulatory requirements for the development and registration of biosimilars in South Africa. Generics Biosimilars Initiat J. 2015;4(3):150–2.

Loeb J. India's new biosimilar guidelines and their relationship to the rest of the world. BioLawGics Blog. Dechert LLP. 2012.

Medicines for Europe. 10 years of biosimilar medicines: transforming healthcare. 2016.

Medicines for Europe. Market review: biosimilar medicines markets policy overview. 2017.

Milne C, Cairns C. Generic drug regulation in the U.S. under the Hatch Waxman Act: a balance between competition and innovation or confusion and intrigue? Pharm Dev Regul. 2003;1(1):11–27.

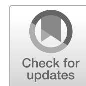
Ministry of Health, Labour and Welfare. Guideline for the quality, safety, and efficacy of follow-on biological medicinal products (interim translation). 2009.

Mulcahy A, Hlavka J, Case S. Biosimilar cost savings in the United States: initial experience and future potential. White paper. RAND Corporation. 2017.

National Conference of State Legislature (NCSL). State laws and legislation related to biologic medications and substitution of biosimilars. 2018.

National Health System England [NHS]. Commissioning framework for biological medicines (including biosimilar medicines), Medicines, Diagnostics, and Personalized Medicine Policy Team, National Medical Directorate, NHS England. 2017.
Power D. Licensing and prescribing biosimilars in Australia. Generics Biosimilars Initiat J. 2013;2(3):152–4.
QuintilesIMS. Outlook for global medicines through 2021: balancing cost and value. 2016a.
QuintilesIMS. Delivering on the potential of biosimilar medicines: the role of functioning competitive markets. 2016b.
QuintilesIMS. The impact of biosimilar competition in Europe. 2017.
Tesser J, Furst D, Jacobs I. Biosimilars and the extrapolation of indications for inflammatory conditions. Biologics. 2017;11:5–11.
Tsuruta LR, Lopes dos Santos M, Moro AM. Biosimilars advancements: moving on to the future. Biotechnol Prog. 2015;31(5):1139–49.
Tufts Center for the Study of Drug Development. Cost to develop and win marketing approval for a new drug is $2.6 billion. 2014.
Urquhart L, Gardner J, Elmhirst E. EP Vantage 2016 preview. 2015.
Wadhwa M, Thorpe R. Terminology for biosimilars—a confusing minefield. Generics Biosimilars Initiat J. 2012;1(3-4):132–4.
Welch A. Health Canada's unique approach to biosimilar regulation. Biosimilar Development. 2016a.
Welch A. What to know about emerging market biosimilar pathways. Biosimilar Development. 2016b.
Welch A. The Norwegian biosimilar phenomenon: from biosimilar to 'biogeneric'. Biosimilar Development. 2016c.
Welch A. IQPC biosimilars event emphasizes need for regulatory alignment. Biosimilar Development. 2017.

Chapter 2
Innovation, Patents and Biologics: The Road to Biosimilar Competition: Factors Influencing Investment, Business Decisions and Marketing of Biosimilars



Erwin A. Blackstone and Joseph P. Fuhr

Abstract This chapter provides an overview of the biosimilar market in the U.S. Biosimilars are defined as highly similar to the originator biologic. Biologics are some of the highest priced drugs that treat some of the most serious diseases. The R&D for biologics can be in the billions of dollars. Recently, patents on various biologics have expired, opening up the market to competition from biosimilars. An abbreviated pathway for biosimilars was created through the 2009 Biologic Price Competition and Innovation Act. Seven biosimilars have been approved with only 3 on the market. Exclusivity and patents and the various legal issues that biosimilars face are examined. We examine the potential gains from biosimilars and the future market structure of the biosimilar market. This analysis will help potential biosimilars producers take a holistic view towards the challenges and prepare for the upcoming competition in making biosimilars a reality.

Keywords Barriers to entry · Biologics · Biosimilars · Competition · Exclusivity · Innovation · Patents · Research and development

Introduction

The characteristics and background for biosimilars was described in Chap. 1. This chapter will provide an overview of the biosimilar market, examining the history of the market, the various barriers to entry, the market opportunities, and reasons

E. A. Blackstone (✉)
Temple University, Philadelphia, PA, USA
e-mail: erwin.blackstone@temple.edu

J. P. Fuhr
Widener University, Chester, PA, USA

Thomas Jefferson University College of Population Health, Philadelphia, PA, USA
e-mail: jpfuhr@widener.edu

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_2

why biosimilars may achieve greater uptake in the U.S. than in the EU. It will particularly concentrate on the issues of exclusivity and patents and the various legal and regulatory issues that biosimilars face. We also examine the potential gains to consumers and society from biosimilars and how the market will evolve, including the future structure of the biosimilar market.

The biosimilar market, especially in the U.S., is in its infancy with only seven biosimilars approved, the first in 2015, and only 3 are currently on the market. There is the possibility for even greater innovations as originator firms have to develop new drugs to replace the biologics that are now subject to competition.

This chapter analyzes the issues involved with biosimilars and suggests how to balance the dual objectives of competition and innovation. The cost to develop a biologic can be in the billions of dollars and for biosimilars in the hundreds of millions of dollars. There are differences and some similarities between the generic and biosimilar markets.

The experience of other countries will be examined to gain insights into the probable implications for the U.S. compared to the EU with over 10 years of experience and over 20 biosimilar approvals. However, there are various factors that make the U.S. market unique.

Biologics R&D Costs

The cost of developing biologics can be quite high and has considerable risk. The R&D costs for one originator biologic have been estimated to be between $1.3 billion and $2.6 billion (Blackstone and Fuhr 2015) and when taking failures into account, the costs could be as much as $5 billion each (Blackstone and Fuhr 2015). There are also post-approval R&D costs amounting to about $312 million (Loo 2016). These costs include FDA requirements to monitor and conduct research on possible side effects which become apparent after more widespread use than in the limited clinical trials. The FDA also requires research on possible long term effects. These studies are more common now and are referred to as phase IV (Loo 2017).

Spending on R&D by biotech firms has been traditionally among the highest of any industry. Standard and Poor's (S&P) estimates that spending on R&D ranges from 15% to 25% of total revenue (Loo 2016). Another report indicated that R&D spending as a percent of sales in 2015 for pharmaceutical firms was 31.4% (Loo 2016). The number of approved biologics declined to 22 in 2016, reflective of the uncertainty and volatility within the industry (Loo 2017).

R&D spending has produced important new drugs which has greatly benefited society. Besides improving quality of life, many drugs have reduced medical expenditures and increased productivity. In terms of new molecular entities approved, the average between 2006 and 2015 was 29.3 per year, increasing to 45 in 2015 from 41 in 2014. The 45 approved in 2014 was the highest since the record approval of 53 in 1996 (Loo 2016).

The great uncertainty involved in drug development can be seen in the statistics that over the past 18 years 96 attempts to develop a melanoma drug were

unsuccessful, 67 failed for lung cancer and brain cancer had 75 unsuccessful efforts. In contrast over the same 18-year span, 7 drugs were approved for melanoma, 10 for lung cancer and 3 for brain cancer (Loo 2016). Another indication of the risk involved in drug development is the fact that only 5 out of a 1000 compounds initially investigated became approved drugs (Loo 2016). Furthermore, only 5–10% of drugs that reach clinical trials get FDA approval (Blackstone and Fuhr 2017) and only 30% of those approved are commercial successes (Loo 2015). Merck has estimated that 75% of its R&D is spent on failures (Blackstone and Fuhr 2012). Interestingly, most drugs fail in phase II clinical trials because of lack of efficacy. If drugs could be tested earlier for efficacy R&D costs could be reduced (Loo 2017).

The biopharma industry is replete with entry and exit of firms, reflective of its dynamic and early stage. For example, 394 biotech firms were listed on the U.S. stock exchange in 2008, declining to 300 in 2013 and expanding to 348 in 2016 (Loo 2016). In February 2017, the number of public companies reached 409 reflecting substantial entry (Loo 2017). Eighty began operations as a result of spinoffs or IPOs (Loo 2016). Moreover, between 2008 and 2013, 180 companies were acquired or went out of business. These exits suggest the riskiness of the industry.

Since R&D costs are high, gross margins tend to be high for a specific drug during the period of exclusivity. For example, the gross margins between 2011 and 2015 for the 8 biotech firms in the S&P 500 was 86.6% (Loo 2016). This was among the highest gross margins of any industry in the S&P 500 Index (Loo 2016). However, these are the most successful firms and hence the data do not take into account the many firms that lose money in the pharmaceutical industry.

Brand name or reference biologic producers that develop new drugs spend substantial sums on R&D. They may experience considerable volatility in returns. In general, the industry is characterized by a kind of lottery in the sense that a few drugs like Humira will be exceedingly profitable but most will barely cover their costs. For example, in 2016 Humira's sales were approximately $18 billion. However, only 30% of drugs are profitable (Loo 2017).

In general, biologics have performed well. For example, between 2011 and 2015 the S&P biotechnology index increased 309.2% compared to 130.8% increase for S&P healthcare index and 61.9% for the S&P 1500 general index (Loo 2017). Nevertheless, the industry's volatility is illustrated by its performance in 2016 when the S&P biotechnology index declined by 14.3% while overall S&P 1500 index gained 10.6%. Biotechnology was the worst performing industry in 2016 (Loo 2017). Further indication of volatility in the pharmaceutical industry is illustrated by Gilead Sciences whose sales increased 122.2% in 2014 and 361% in 2015 but declined by 6.5% in 2016 as a result of heightened competition in the hepatitis C market (Loo 2017). In the case of chemical generics, patent losses often translate into substantial financial losses. For example, Lipitor had sales of $9.57 billion in 2011, its last year of patent protection falling to $3.9 billion in 2012 and yielding only $1.75 billion in 2016 (Sagonowsky 2017b). Interestingly, sales were substantial 5 years after patent expiration. Reference products of biologics should probably retain significant sales and market share given their competitive response. An exception may be if they develop a second generation biologic which is really a continuation and a substitute of the original.

Another issue is the high percentage of newly approved drugs that are orphan drugs, defined as those diseases that occur in 200,000 or fewer people in the U.S. Prior to the Orphan Drug Act of 1983 only 38 such drugs were approved but since then more than 400 orphan drugs have gained approval (Loo 2017). The Orphan Drug Act encouraged orphan drug research by providing 7 years of market exclusivity, tax credits for research and grants for clinical trials. Orphan drugs generally command higher prices and face little competition. In 2016, 9 out of the 22 biologics drug approvals were for orphan drugs and in 2015 21 out of 45 drugs were orphan drugs.

Finally, return on equity for the 8 companies in the S&P 500 Index in 2016 was high but was lower than otherwise because of negative returns for several years for Vertex Pharmaceutical (Loo 2015) which is also suggestive of the high risk associated with the industry. Also, important in terms of risk is the fact that a company's revenues and profits are often dependent on one or a few drugs. For example, Humira contributed more than $16 billion to AbbVie's total revenue (or 63% in 2016 (Sagonowsky 2017c). In conclusion biologics involve substantial risk, many fail, but those that are successful are quite profitable.

Biosimilar Barriers to Entry

There are numerous barriers to entry that biosimilars face. The R&D for biosimilars is considerably higher than that of generics. Biosimilar development is expected to cost between $100 million and $200 million and take between 8 and 10 years (Blackstone and Fuhr 2013). For generics the figures are $1 million to $5 million and 3–5 years (Loo 2015). Celltrion, for example, has invested $112 million in the development of Remsima, a biosimilar for Remicade (Celltrion 2014).

Beside the R&D costs there are various other barriers to entry. Entry into the biosimilar market also requires establishing manufacturing facilities that meet FDA requirements regarding "good manufacturing practices (GMP)." The cost of establishing a manufacturing facility has been estimated to be around $250 million (Blackstone and Fuhr 2015). However, firms might contract out manufacturing to avoid these high initial costs. Whether this is possible depends, among other things, on the utilization rate of existing manufacturing capacity.

The complexity of the process makes expertise in manufacturing quite important. Companies' experienced in biologic manufacturing will have a learning curve advantage, which translates into a cost advantage. Some drug approvals are delayed because of concerns over manufacturing, indicating the importance of manufacturing (Loo 2017). For example, Pfizer's biosimilar for Amgen's Epogen received strong support from the FDA advisory panel in 2017 but the FDA rejected the license application because of concerns over the manufacturing plant (Palmer 2017c). Also, Mylan's EU application for biosimilars was delayed because of deficiencies in its planned manufacturing in Biocon's manufacturing facility in Bangalore India (Stanton 2017b).

There is the possibility of having difficulty in develop a biosimilar for a reference product. The PEGfilgrastim biosimilar is one such unique molecule where several biosimilar players have hit hurdles in their development and commercialization plan. Neulasta's filgrastin protein is bound to a polyethylene glazed molecule which is thought to make biosimilar development difficult (Royzman 2017). For example, Sandoz and Gedeon Richter encountered difficulties and each withdrew its biosimilar application from the European Medicines Agency in 2016 and November 2016, respectively (Royzman 2017). Also both Sandoz and Coherus have received rejection letters from FDA and are expected to resubmit (Mehr 2017a). Also Mylan (partner with Biocon) got a completed response letter from the FDA in October 2017 for their PEGfilgrastim biosimilar (Palmer 2017a). Apotex (partner with Intas) has a BLA submission for PEGfilgrastim biosimilar being reviewed by FDA.

Biosimilars entrants are often large, biologic originators for other reference products. For example, Amgen, Novartis and Pfizer have entered the biosimilar market. In particular, Amgen has developed a biosimilar for Roche's blockbuster drug Avastin. In mid-July 2017 an FDA advisory committee voted 17–0 to recommend approval of Amgen's ABP 215 as a biosimilar for Avastin. Generic firms are also potential competitors as evidenced by an FDA advisory committee 16–0 vote on the same day to recommend approval of Mylan's biosimilar for Roche's Herceptin (Sagonowsky 2017d). There will be many small firms and start-ups in the market. These companies if successful may eventually partner or merge with a larger firms. Celltrion, a South Korean firm, has been successful in developing and manufacturing and then partnering with other firms to sell its biosimilars.

Currently, the FDA requires clinical trials for biosimilar approval. These trials can be quite expensive, exceeding a million dollars (Cohen 2015). Also 85% of clinical trials are delayed because of difficulties in obtaining sufficient patient recruitment (Loo 2015). With so many biosimilars for the same reference product being developed, in some cases over 30, (Stanton 2016) it will be even more difficult to get volunteers which can delay the approval process. The price of obtaining the reference product can be high and Risk Evaluation and Mitigation Strategies (REMS) can make it difficult to obtain samples of the reference product (Brill 2014). This occurs typically when a specialty pharmacy is used and restricts quantities that may be purchased (U.S. Senate, Special Committee on Aging 2016). Also, biosimilars must do post marketing monitoring which adds to the costs.

The originators have actively responded to potential competition from biosimilars in a variety of ways. They have decreased prices, developed second generation biologics of the reference biologic (biobetters), pursued patent extension and litigation, developed better devices for administering the drug, and reduced the frequency of dosages. In some cases, the oddities of patents play a role. For example, Amgen has patent protection on Enbrel in the U.S. until 2029. These patents on Enbrel were granted in 2011 and 2012 even though the innovation occurred in the early 1990s (Pollock 2016). In the case of Enbrel, patent protection could well last over 31 years (Pollock 2016). AbbVie has obtained 70 new patents (most since 2013) on various formulations of Humira, its manufacturing methods, and as treatments for specific diseases which could protect it from competition through

at least 2022 even though the basic patent expired in 2016 (Pollock 2016). Further, the company has indicated that it will vigorously defend patents against potential biosimilars. Indeed, AbbVie in 2017 sued Boehringer Ingelheim concerning its biosimilar for Humira. The suit claimed that Boehringer's biosimilar would infringe on AbbVie's patents. AbbVie noted that once Boehringer gave the 6-month notice of entry it would sue on other patents (Sagonowsky 2017c). This increases the risk and cost of entry.

Some originators are producing second generation biologics which offer various benefits over the first generation biologics. Amgen's second generation of epo has a weekly injection compared to the first generation which requires multiple doses weekly. This could increase adherence which would benefit consumers. Roche has developed Gazyva which is a second generation of Rituxan/MabThera. In clinical trials Gazyva results in a median of 11.5 month extension of life over the first generation (Staton 2013).

So will physicians stop prescribing Rituxan and switch to Gazyva? Roche is switching to subcutaneous reformulations which can deter biosimilar competition. For Mabthera (Rituxan in the U.S.) treatment time has been reduced from 2.5 h to 5 min (IMS 2013). Neulasta, the second generation of Neupogen is administered once a week compared to Neupogen which is given daily (Sierra et al. 2008). Also, a single treatment cycle cost of Neulasta is $3400 compared with Neupogen's cost of $6000 (Dinwoodie 2011).

J&J contends that it does not expect to lose much market share to Remicade biosimilars since it has both its own discount program and "innovative contracts." J&J has stated that "We feel well prepared to face the biosimilar" (Stanton 2017). Pfizer has sued J&J claiming that some of its conduct violates antitrust laws (Crow 2017).

Besides developing the biosimilar, firms must often develop the devices to inject the biosimilar. Often the devices of the reference product are still under patent and thus the biosimilar must develop an alternative injectable device. Also, firms must educate medical providers and patients on the proper use of these devices. Another factor involves the precise form of the biosimilar. For example, the biosimilar Zarxio is not available in a vial and children need to be administered the drug in a vial (Barlas 2016). Thus, it cannot compete in the children's market against Neupogen. In addition, entrants might provide a number of samples to physicians. Unlike generics, biosimilars need to provide wrap around services to their patients, all of these factors increase cost and raise entry barriers.

Also, given the relative newness of biosimilars, firms will need to expend resources on educating stakeholders on what a biosimilar is. For example, both Pfizer and Sandoz are initially using sales forces to make physicians aware of biosimilars. One would expect biosimilar competition to be more similar to brand to brand than brand to generic. CVS states "That means the price differential between a biologic and its biosimilar is more likely to approximate the competition in a multi-brand category of drugs rather than between a reference drug and its generic" (Singh 2016). This is consistent with the early experience of biosimilars where price

differences are in the 15% range. Given more experience with biosimilars, future price competition could lead to even greater discounts.

The difficulty and risk of entry into the biosimilar industry is indicated by some notable exits from the industry. Merck KGaA has decided to get out of biosimilar development and sell its biosimilar unit (Helfand 2017b). Also, Allegran has decided to leave the biosimilar market after it finishes the 4 cancer biosimilars that it is developing with its partner Amgen (Helfand 2017b). Also, several firms stopped development of the Rituxan biosimilar.

FDA Lag in Approvals

The FDA was given no additional funding when the BPCIA was approved. However, the FDA has been permitted by Congress to collect user fees from biosimilar applicants. The FDA has set a goal of 10 months for a review of biosimilars. For 3 of the first 4 approvals this goal has basically been met. However, in the case of Celltrion's Inflectra application of August 2014, the FDA asked for more data and the biosimilar was finally approved in January 2016. However, other biosimilar applicants have filed as early as December 2014 and there is still no word on the applications. The complexity of biosimilars, the case-by-case approach and the increase in applications is leading to a backlog which may get even longer as the number of applications increase. In 2016 the FDA reported that there were 66 biosimilars for 20 reference products in its biosimilar development program (Crespi-Lofton and Skelton 2017). Such increases in approval time effectively increase the cost of entry through the additional capital required and the longer time to get the product on the market.

Economic Justification for Patents and Exclusivity

The economic rationale for exclusivity and patents is to allow innovators the ability to obtain a return on investment provided they supply a commercially feasible product. Patents are important to encourage investment in R&D in pharmaceuticals overall and especially in biologics where R&D costs are even greater than in chemical or small molecule drugs. Patents provide a temporary monopoly of 20 years for the company's innovative product. However, the firm may face competition from other products which are close substitutes for its product, for example, Enbrel and Humira. The patent period begins on the date of the patent application. Also, the pharmaceutical market is unique in that pharmaceutical firms must also receive FDA approval before they can market their drug. This can take many years since the drug will go through various phases before FDA approval. It has been estimated that on average the effective patent protection for a pharmaceutical is around 10–12 years (Blackstone and Fuhr 2017). Furthermore, patents can be challenged and thus there

is no guarantee that a patent will last for 20 years. For example, Amgen's Neulasta was approved for marketing in 2002 and in 2014 Apotex applied for biosimilar approval, suggesting a possible effective patent period of 12 years for the originator product Neulasta.

Data and Market Exclusivity

To provide added incentive to innovate given the uncertainty of patents in the U.S., biologics get a 12-year market exclusivity from the date of FDA approval and this is not subject to litigation. There has been much debate surrounding the optimal length of the market exclusivity. During the debate about exclusivity, proposed periods ranged from 7 to 12 years (Blackstone and Fuhr 2017). If the length is too short, it could lead to less innovation and some potentially life-saving biologics will not be developed. If the length is too long it could lead to higher prices for a longer length of time and access issues. We believe that it is better to have a drug available which can benefit society rather than not have the drug developed at all or much later. Thus if an error is to be made, it should be on the long side.

Under the BPCIA, the originator is entitled to four years of data exclusivity, which means that a biosimilar cannot use the reference product's data during this time period. In fact, the FDA will not accept an application for a biosimilar until 4 years have elapsed from the originator's approval for marketing. In addition, the 12-year market exclusivity means that no biosimilar can be marketed during that period.

The question is whether data and especially market exclusivity are necessary given the existence of patent protection. Industry groups argue that patents on biologics are not as strong as those on small molecule or chemical drugs because biologics are more complicated so that greater ability exists to "work-around" such patents. Further, they argue that it often takes 10–12 years to obtain FDA approval for marketing so that only 8–10 years may remain on the patent. In addition, research by Henry Grabowski contends that a minimum of 12 years is required to recoup investment in drug development (Grabowski et al. 2011). Moreover, the BPCIA does not provide the possibility of patent extensions allowed under Hatch-Waxman when FDA approval takes an inordinate time. In any event, market exclusivity serves as an insurance policy to promote investment in biologics since it cannot be legally challenged. The market exclusivity is also desirable, given the fact that many small start-up firms have invested in biologic development. These start-ups need some protection from competition to be able to attract investment to develop their biologics.

On the other hand, some suggest that patent protection is adequate and market exclusivity of 12 years provides excessive monopoly power and restricts access to needed drugs. One critic points out that Canada, for example, only provides 8 years of market exclusivity for biologics (Shin 2016). However, the U.S. has been burdened with subsidizing the rest of the world in R&D for pharmaceuticals.

Further, the critic contends that entry barriers and the nature of brand-to-brand competition, where prices decline 18–27% from the originator's price, will still allow sufficient profits to encourage drug innovation (Shin 2016). However, we have seen prices decreases greater than that in the EU. In any event, market exclusivity of 12 years is desirable to ensure adequate incentives for innovation.

Patents as Barriers to Entry

It was difficult to develop the law that opened the pathway for the FDA approval process for biosimilars which is still evolving as is the interpretation of the law. There are still various legal and regulatory issues, especially involving patents. Unfortunately, in the U.S. it is often not clear when biologic patents are invalid or no longer in effect. This can lead to much uncertainty and higher risk for biosimilars. There is a high probability of a patent dispute. For example, from 2011 to 2014 over 90% of initial generic entrants in the chemical drug market faced patent disputes (Frois et al. 2016). This pattern is continuing in the biosimilar market where there are already many patent disputes. In fact, the cases discussed below as well as others involve patent issues that confirm this. In some cases, the originator is able to block entry by being granted a preliminary injunction. Even without the injunction, if a biosimilar company decides to enter the market before the patent issues are resolved it risks losing the patent case which could result in substantial payments to the originator for patent infringement. This often leads to delayed entry. Pfizer and Merck have decided to enter the Remicade market at risk whereas Sandoz has currently decided not to enter the Enbrel market until the patent issues are resolved. This is especially problematic for small firms that lack adequate financial resources to surmount such difficulties. The BPCIA has tried to reduce some of the risk with the "patent dance." The specifics of the patent dance are discussed in Chap. 4.

Among the types of patents are those that cover the biological properties of the product or what is called a "composition of matter" patent. Another type of patent covers the manufacturing and selling of the biologic for a specific illness or condition or what is called a "use" patent. There is also a formulation patent which is often a mixture of more than one drug or an application for a particular illness. Finally, there is a patent that covers the manufacturing of a biologic or what is called a "process" patent (Loo 2015). In general, it is thought that the composition of matter patent provides the strongest protection (Loo 2015). The number of patents and their complexity are greater for biologics than for more simple chemical drugs. For example, Abbvie's biologic Humira has "more than 60 patents, including two product patents, 11 use patents, 26 formulation patents, nine process patents and eight patents on delivery devices" (Serebrov 2016). This makes patent litigation more complex and uncertainty greater for biosimilars. Moreover, even though Humira's composition of matter patent expired in 2016, additional patents are thought to protect it from biosimilar competition through 2022 (Loo 2017).

Chapter 4 of this book specifically covers litigation related issues under the BPCIA. However, for general inquisitiveness and education of the reader we discuss below the specifics of the patent dance procedure. We also discuss what has transpired in the litigations cases related to biosimilars e.g. Sandoz vs Amgen, Amgen vs Apotex, Janseen vs Celltrion and the Mylan litigation case.

Patent Litigation Before Application

In 2013 and 2014 the first legal issue concerning patents and biosimilars was contested. It revolved around the ability of a potential biosimilar applicant to legally challenge the patents of the originator (reference product) before it files a biosimilar application. Several biosimilar makers filed declaratory judgement actions concerning this issue and they were not successful. “At the district court level, the pre-application declaratory judgment actions filed by biosimilar makers were dismissed on two separate grounds. First, the courts held that before the acceptance of an application there was no justiciable case or controversy under the Declaratory Judgment Act and Article III of the Constitution. Second, they reasoned that the BPCIA itself prohibited applicants from bypassing the patent dance by filing early declaratory judgment actions” (Fischer 2016).

One case that made it to the Federal Circuit was Sandoz v. Amgen, 773 F.3d 1274 (Fed Cir.2014). The Court “affirmed dismissal of the declaratory judgment action on justiciability grounds without reaching the statutory question” (Sandoz v. Amgen, 773 F.3d 1274 (Fed Cir.2014). There have been no recent cases concerning pre-application declaratory judgement and this issue seems to have been resolved. The next issue and litigation revolved around the patent dance itself.

Litigation Concerning the Patent Dance

Sandoz v. Amgen

Sandoz which submitted the first application for a biosimilar questioned whether the patent dance was mandatory. It claimed that it was voluntary and refused to dance which resulted in litigation. The case also concerns when the biosimilar can provide the required 180-day notice of its intent to commercially market its product.

The specifics of the case are as follows. Amgen has been marketing the biologic filgrastim with the brand name of Neupogen since 1991. Sandoz filed an aBLA in May of 2014 to market its biosimilar to Neupogen called Zarxio which the FDA accepted for review on July 7, 2014 (Amicus Brief 2017). On July 8, 2014 Sandoz notified Amgen of its intent to market its biosimilar filgrastim once it received FDA approval. However, it did not provide Amgen with a copy of its application nor a

description of its manufacturing process. Amgen filed an (artificial) infringement action of its filgrastim method of use patent. Amgen also claimed a violation of California's unfair competition law because Sandoz did not provide advance notice of commercial marketing.

The FDA approved Sandoz's application for Zarxio on March 6, 2015 for the same indications as Neupogen (Amicus Brief 2017). Soon after the FDA's approval, the district court decided that Sandoz did not have to participate in the patent dance, did not have to provide its application to Amgen, and did provide timely notice to Amgen. The district court denied the unfair competition claim and Amgen's request for injunctive relief.

The decision was appealed to the Federal Circuit which enjoined the marketing of Zarxio until it reviewed the lower court decision. Sandoz responded in federal circuit court that Amgen's patent was invalid and not infringed and also wanted a judgement that it did not violate the BPCIA. Amgen obtained Sandoz's biosimilar application through the discovery process.

The appellate court decided that Sandoz did not have to provide Amgen with its application, did not have to participate in the patent dance, and Amgen's remedy was simply to file its (artificial) infringement suit. However, it ruled that notice of first commercial marketing could not be given before FDA approval so that it enjoined Sandoz from marketing Zarxio until September 2, 2015, 180 days after Sandoz provided Amgen with its second notice of intent to market Zarxio (Amicus Brief 2017).

This case presents a number of issues. In particular, there is the interpretation of the BPCIA which revolves around the word "shall." The BPCIA reads that the applicant shall provide the reference producer its application. The Federal Circuit determined that "shall" does not involve a mandatory requirement. Also, involved is whether the notice of commercial intent can be given before FDA approval. Both Sandoz and Amgen appealed the case to the Supreme Court.

These issues were seemingly resolved by the U.S. Supreme Court's June12, 2017 decision where the Court decided that federal injunctive relief for not complying with the requirements of the BPCIA was not available, compliance with the BPCIA was not mandatory and the biosimilar applicant could give the 180-day notice before its application is approved (Sandoz Inc. v. Amgen Inc. et al, No.15-1039). The Supreme Court thus affirmed the Federal Circuit's non-mandatory decision but reversed the 180-day notice portion. The Supreme Court also did not decide whether an injunction was available under California law but remanded that issue to the Federal Circuit.

The patent dance, notification requirements, and provision of the biosimilar's application have merit. The patent dance is supposed to encourage early patent resolution (perhaps within the 12-year market exclusivity) so that entry of the biosimilar could occur soon after the patent protection and market exclusivity periods end. The provision of the biosimilar application and manufacturing information allows the reference producer to determine any infringement issues. The 180-day notice allows the reference producer time to file any infringement action before biosimilar marketing occurs. Finally, requiring that notice of commercial entry be given only

after FDA approval could extend the period of market exclusivity beyond the 12-year statutory period. If the FDA approval comes after the 12-year exclusivity period, the reference producers would have at least 12.5 years of protection from competition instead of the statutory period of 12 years.

It appears that the Supreme Court's decision favored biosimilar applicants by making the "patent dance" optional including providing the application to the reference producer or the sponsor, requiring enforcement action to be done by the reference producer through patent infringement suits, and allowing notice of commercial marketing before FDA approval. The biosimilar generally gains greater discretion and choice in resolving patent disputes (Gordon 2017).

Critics of the decision suggest that the enforcement mechanism is weak so that now greater uncertainty exists. Biosimilars may choose to enter at risk instead of having patent disputes resolved through the procedures of the "patent dance". On the other hand, the threat of substantial damages could lead biosimilars to delay entry. If biosimilars have reason to believe that they are infringing on a valid patent, the court could impose triple damages. The decision thus did not clarify the landscape (Quinn 2017).

The BPCIA provisions involving disclosure of information in the patent dance have some troublesome aspects. The reference producer gains competitively sensitive information about the biosimilar's manufacturing operation which could be used by the reference producer to inhibit competition. Discussions between competitors could also lead to antitrust-type issues depending on the number and timelines of entry by other biosimilars. Understandings could develop that could be harmful to buyers of the biosimilar and reference biologic.

The "patent dance" provisions provide a good opportunity for collusion between the reference producer and the first biosimilar entrant. The negotiations are secret, non-transparent and provide an opportunity for gain to both parties at the expense of other potential biosimilar entrants and the public. The parties could agree to not list all the patents and the reference producer could agree not to challenge the entrance of the biosimilar. By keeping some patents from litigation the reference producer could benefit from reduced competition (Carrier and Minnit 2018). The advantage of having only one biosimilar in the market is clear. For example, an executive of J&J indicated that competition is much greater when more than one biosimilar is competing (Stanton 2017b). Further, research on generic drugs shows that when 7 or more firms compete, prices approximate the cost of manufacturing, namely, the competitive level (Reiffen and Word 2005).

Amgen v. Apotex

In *Amgen v. Apotex* 827 F. 3d 1052 (2016) the Federal Circuit Court of Appeals also considered issues related to the 180-day notice. Apotex claimed that since it went through the patent dance, the 180-days could start before approval. The specifics of the case are as follow. Amgen had been marketing Neulasta since its FDA approval

in 2002. The biologic is a human engineered protein that increases white blood cells to help fight infections for those undergoing chemotherapy (Amgen v. Apotex 827 F. 3d 1052 (2016). In October 2014, Apotex filed an aBLA to market a biosimilar to Neulasta. The FDA accepted the application on December 15, 2014. On December 31, 2014 Apotex provided Amgen a copy of its aBLA and a description of its manufacturing process for pegfilgratism. On February 2, 2015, Amgen provided Apotex its list of three patents to which Apotex responded on April 17, 2015. Apotex stated that it did not intend to market its biosimilar until two of the three patents expired and claimed that the third was either not infringed or was invalid. On that same day, April 17, 2015, Apotex notified Amgen that it intended to market its biosimilar when it received FDA approval. On June 16, 2015, Amgen responded concerning patent infringement. After Apotex and Amgen negotiated about the patents involved, on August 6, 2015, Amgen sued for infringement of its two existing patents. One of the patents expired in October 2015 so the infringement action then claimed only one patent, no. 8,952,138(827 F. 3d 1052 (2016). The Federal Circuit Court had just decided *Amgen v. Sandoz.* The district court rejected Apotex's claim that it gave the appropriate180 days' notice, noting that the notice has to be given after FDA approval. Apotex argued that the notice requirement could extend the 12-year exclusivity to as much as 12.5 years if the notice were given at the end of the 12-year exclusivity period. The Federal Circuit in affirming the district court noted that the 180-day notice would usually be given within the 12-year exclusivity period and thus should not be expected to extend the 12-year exclusivity (827 F. 3d 1052 (2016). The court added that the purpose of the 180-day requirement is to provide the reference product producer the opportunity to challenge any alleged infringement once the biosimilar is licensed to enter and the specifics of its license are known. The Federal Circuit is probably correct in its opinion that most often the 180 days' requirement notice should not extend the 12-year exclusivity and that the originator should have 180 days to review the specifics of the licenses. As an added note the FDA still has not ruled on Apotex's application. In any event, the U.S. Supreme Court's decision makes clear that the notice of commercial marketing can be given during the time that the application is pending at the FDA. The Supreme Court noted the good policy arguments on both sides of the issue.

Janssen v. Celltrion

This case involves the issue of whether failure to complete the "patent dance" meant that the patent holder was not limited to reasonable royalties even if an infringement suit was filed more than 30 days after the completion of the exchange of information. The specifics of the case are as follows. Janssen filed a patent infringement suit after the FDA accepted Celltrion's aBLA for its biosimilar for Janssen's Remicade (Janssen 2017a). Celltrion engaged in the "patent dance" by exchanging information with Janssen but stated that it would not negotiate any disputed patents as is required by the BPCIA. Accordingly, Janssen argued that the information exchange process

was not completed and therefore it did not have to file an infringement suit within 30 days in order to avoid being limited to reasonable royalties.

On March 2, 2017, the district court in Massachusetts held that Celltrion's non-compliance with the 'patent dance' rules meant that Janssen was not limited to reasonable royalties, and if it won it could collect lost profits. The court noted that only those patents that have gone through the entire prescribed process of the BPCIA are subject to the limitation of reasonable royalties in case of infringement (Janssen 2017b). The court also interpreted "shall" as a mandatory requirement. The ruling could greatly increase the riskiness of entry until patent issues are resolved for those who do not complete the 'patent dance'.

With respect to the patent issue itself the court held that Janssen's patent '471 on Remicade was invalid, a decision Janssen appealed to the Federal Circuit and to consideration within the U.S Patent and Trademark Office. Celltrion also claimed that it did not infringe on Janssen's '083 patent (Staton 2013). The patent is scheduled to expire in September 2018, "Celltrion withdrew its motion to dismiss the '083 patent (i.e., the sole remaining patent in the case) and instead filed a motion for partial summary judgment of non-infringement of that patent" (Big Molecule 2016b). The litigation will probably not be settled before the patent expiration. In the meantime, the biosimilar Inflectra has been launched at risk.

Other Patent Litigation

Mylan

Mylan and Biocon, partners in the development of a biosimilar for Roche's Herceptin, reached an out of court settlement of their patent dispute with Roche. Roche has agreed to give the partners a global license and in turn the partners agree to no longer pursue the two challenges pending before the Patent Office. The biosimilar has not yet been approved and the negotiated biosimilar launch date has not been revealed (Sagonowsky 2017f). However, Mylan is facing a 3 month delay from the FDA and is now set for a December 2017 review (Palmer 2017b).

In addition, by settling out of court the patent issues will not be resolved, and the next entrant will presumably need to go through the patent dispute. Thus, Mylan/Biocon could be the only biosimilar entrant in the market until the patent issues are resolved unless the next approved biosimilar wants to enter at risk. The settlement means that society gains a biosimilar and Roche maintains its patent. However, the settlement may result in the launch delay of other biosimilars, since it prolongs the patent dispute issues.

Other Patent Issues

Illustrative of the likelihood of ongoing patent litigation is the situation with AbbVie. The company in 2017 had its "135" Humira patent for a method for treating rheumatoid arthritis struck down in a case before the U.S. Patent and Trademark Office's Patent and Trial Appeal Board (Helfand 2017a). AbbVie indicated that it would appeal the internal patent review decision. The company has won some challenges of its patents (Helfand 2017a). Indicative of AbbVie success is its settlement with Amgen under which Amgen agrees not to enter with its Humira biosimilar until 2023 by which time Abbvie's key U.S. patent will have expired. The settlement suggests the strength of Abbvie's patent (Sagonowsky 2017a). Another company Janssen, a subsidiary of Johnson and Johnson, sued Bioepis claiming that its biosimilar to Remicade infringed on three J&J patents (Sagonowsky 2017e). These cases indicate the ongoing litigation costs, uncertainty, and risk involved with the industry.

Interchangeability

As discussed in Chap. 1 of this book and later in a chapter specifically covering it, interchangeability continues to be a controversial and interesting aspect of biosimilar competition and approval process.

Under the BPCIA a biosimilar can apply for interchangeability status and if successful can get 1-year interchangeability exclusivity. It would also allow pharmacists, where state law permits, to substitute the interchangeable biologic without obtaining a physician's authorization. The physician has the option to write do not substitute. To achieve interchangeability approval it would seem to require substantial switching studies between the biosimilar that wants such status and the reference product. Moreover, the 1-year exclusivity applies only to interchangeability. The interchangeable biologic could and probably would face competition from other biosimilars for the reference product.

Since many biologics are infused and thus physician administered they would not reach the pharmacy level. Besides the additional costs to achieve interchangeability there is the added risk of being denied, which may be a marketing disadvantage. Also, at issue is whether payers are willing to pay a price premium for interchangeable biologics which seems unlikely. For such reasons there is little market advantage to being interchangeable (Blackstone and Fuhr 2012).

Indicative of the modest market advantage of interchangeability is the fact that Zarxio, the first biosimilar in the U.S. market could have made a strong case for interchangeability based on its switching studies but did not attempt to be granted that status (Mehr 2017b). There are some rumors that the FDA had recommended against such an attempt.

Pricing and Market Opportunities

As has been discussed there are considerable barriers to entry in the biosimilar market but there are also market opportunities. Biologics comprise only 1% of the prescriptions filled in the U.S. but account for around 28% of prescription-drug spending (Sarpatwari et al. 2015). Revenues for biologics are growing at twice the rate of global drug revenues overall (Blackstone Fuhr 2015). For example, the compounded growth rate of sales between 2012 and 2015 for companies in the S&P 1500 Biotechnology Index was 20.3% (Loo 2016). Further, biotech drugs have increased from 18.9% of worldwide prescriptions and over the counter sales in 2010 to 23% in 2014 (Loo 2016). It is expected that the biologics market will increase at a rate of more than 20% per year, and that by 2025 more than 70% of new drug approvals would be biological products (Erickson 2010).

IMS projects that the global biologics market will exceed $390 billion by 2020 and comprise 28% of the global value of pharmaceuticals (IMS 2016). U.S. sales in 2014 were around $200 billion and grew over 10%. In fact, the U.S. constitutes around 50% of the biologics market.

Blockbusters generate very large profits which gives firms the incentive to enter. Many biologics have annual sales of over a billion dollars with some having sales over $10 billion. Suppose that a biosimilar obtains 10% of that blockbuster's sales of $10 billion or a billion dollars, even if discounted at 20% the drug would generate $800 million annually in revenue. Over 30 biologics have lost or will soon lose patent protection which represents $80 billion in revenues (Blackstone and Fuhr 2017). The reference products for the 5 biosimilars that Sandoz has in its 2020 development portfolio alone generated nearly $44 billion in 2015 global sales (Helfand 2016).

Given the potential market opportunity, there is expected to be an influx of biosimilars into the market and there are many firms developing biosimilars. Noteworthy, the high potential profits in biologics have attracted pharmaceutical firms. Indeed, by 2020 it is projected that 8 of the top 10 worldwide sellers of prescription biotech drugs will be pharmaceutical companies (Loo 2016).

Biologics are among the highest priced drugs. For example, in 2016 Soliris, the most expensive biologic, had an annual patient expense of $600,000 (Loo 2016) and Amgen is developing a biosimilar for it (MacDonald 2016). The next highest annual expense is for Naglazyme at $485,747 (Loo 2016). Such drugs are increasing insurance plan expenses. For example, specialty drugs, a category that includes biologics, accounted for 27.7% of insurance plan costs in 2013 and grew to 37.7% in 2015 (Loo 2016). The high prices for some biologics and other drugs have generated much controversy.

In the EU biosimilar competition initially resulted in price decreases of around 20–30% and the first 2 biosimilars in the U.S. entered with a 15% discount. However, Pfizer negotiated additional discounts (Stanton 2015). Further, Merck 2nd Samsung introduced their biosimilar to Remicade at a 35% discount versus the 15% offered earlier by Pfizer (Sagonowsky 2017b).

The low initial discount is not surprising. In the generics market large discounts do not occur with only one competitor is in the market, especially since the first generic entrant often has 180-day exclusivity. However, once the generic exclusivity period ends, the entry of additional firms drives down the price. Given the higher costs of developing biosimilars one would not expect prices to decrease as much as 80–90% as we have seen in the generic market. However, as the market matures and competition increases, discounts could reach at least 50% in most markets.

Many biosimilars are being produced by brand name companies, which, because of their reputation, should be at less of a competitive disadvantage than early entrants into the generic market. One would expect biosimilar competition initially to be similar to brand to brand than brand to generic.

U.S. Market

Chapter 3 of this book addresses and explains in detail the biosimilar reimbursement landscape in US. The authors also compare and contrast the US biosimilar reimbursement scenario with the experience thus far in EU.

Biosimilar uptake in the EU has been successful when stakeholders have the right incentives. Most EU countries have a single payer system. The U.S. market is more complex with various private and public payers. The government payers include Medicare, Medicaid, and the Veterans Administration. Each has its own unique reimbursement rules. Medicaid and the Veterans Administration are able to negotiate discounts through rebates. For example, under the Affordable Care Act of 2010, the Medicaid rebate increased from 15.1 to 23.1% of the average manufacturer price (Loo 2017). However, Medicare Part B drug reimbursement is based on average sales price. When Medicare Part D was enacted for prescription drugs the government was forbidden to use its market power to negotiate drug prices. This provision of the law has been very controversial and there has been constant debate concerning revoking it. There are also many private payers with various degrees of negotiating power and different reimbursement rules.

High biologic prices are leading to pressure by payers to switch to lower priced biosimilars. In the U.S. market, third party private payers have the ability to negotiate the best deal for their clients. Private payers can use various strategies to encourage the use of biosimilars. They can use tier formularies with lower co-pays for biosimilars or the lower priced biologic, step therapy where a lower priced product is used first and if it does not work the patient can be switched to a different medicine, prior authorization, exclusion of higher priced medication from the formulary, higher buy-and-bill margins for less expensive drugs, and exclusive arrangements (Singh 2016).

In the U.S. exclusive arrangements could be utilized to encourage more competition which could lead to more rapid expansion of the use of biosimilars. An example of how the exclusive arrangement could develop in the U.S. biosimilar market can be drawn from the U.S. hepatitis C drug market. Gilead entered the U.S. market

with Sovaldi at a list of price of $84,000. This was for a 12-week treatment which cured the patient. AbbVie entered with Viekira Pak and negotiated an exclusive arrangement with a private payer at a discount of around 46% (Stanton 2015). Gilead responding with a price decrease to get a different exclusive arrangement.

Thus, private payers can act similar to the tender system used in various countries in the EU which is essentially an exclusive arrangement with the manufacturer. This has led to considerable discounts, as much as 72% in Norway (Generics and Biosimilar Initiatives 2015). Prices for biologics are generally higher in the U.S. than in the EU so discounts can be greater. Moreover, private payers are responding to biosimilar entry in various ways. Express Scripts has reported that it is eager to adopt biosimilars to cut costs. The Ohio Public Retirement System has a lower co-pay for biosimilars than for reference products. These lower prices could lead to greater uptake. CVS has put the biosimilar Zarxio on its formulary and has excluded the reference product Neupogen. CVS noted: "For biosimilars, we plan to do this by either utilizing the biosimilar as a more cost-effective alternative to the reference brand, or by utilizing the competition presented by a biosimilar to seek deeper discounts from the reference brand manufacturer" (Singh 2016).

A factor that can discourage the use of biosimilars is the "buy and bill" system of reimbursement commonly used by commercial insurers. The system involves providers buying the drug and then billing the insurer for the drug including a fee for the transaction costs of buying, storing and dispersing the drug. Suppose the reference product is priced at $1000 and the biosimilar is priced at a 15% discount or $850. If the fee is 6% of the products price the biosimilar provider receives $9 per dose less clearly discouraging use of the biosimilar. Medicare provides a fee based on the average selling price of the reference product so that there is no discouragement for biosimilars (Lagaso 2017).

Bundling of payments where a provider receives a fixed price that includes all components of treatment would encourage the use of less expensive inputs, including biosimilars. Many biologics are physician administered so that bundling could be easily adopted for these biologics. Accountable Care Organizations (ACOs), encourage the use of lower priced biologics since providers earn more when costs are reduced. Reference pricing, where patients pay out of pocket for prices above the insurance reimbursement rate, can result in patients using biosimilars (Blackstone and Fuhr 2015). Also, private payers can educate patients and doctors about biosimilars to decrease the possible resistance from these stakeholders who can gain from lower prices (Serebrov 2016).

Another factor that can lead to greater uptake is switching. The countries in the EU that have seen the greatest uptake are those that have allowed switching. Switching occurs when patients that were originally on the reference product are switched to a biosimilar. Originally in many EU countries only new patients were prescribed biosimilars. Established patients were kept on originator biologics. This made the available market for biosimilars much smaller and hindered uptake. This changed with the Norway tender. In fact, the NOR-SWITCH study concluded switching from Remicade to Remsima is well tolerated and feasible (Buer et al. 2017).

Switching will be key to uptake and is occurring in the U.S. According to CVS "Because biosimilars are therapeutically equivalent to reference biologics, we expect minimal "grandfathering of patients" (Singh 2016).

Potential Gains from Biosimilars

It is important to note that the primary policy objective from biosimilar competition is to increase consumer welfare. This can be measured by combining the decrease in price and increase in access. The market share of biosimilars alone may not be a fully informative metric. Unlike the generic market where the originator in some cases actually increased price (Kanavos 2008) the reference product producer has often responded to biosimilar competition by lowering its price. Consumers will benefit from this lower price even though the market share of biosimilars will be lower because of this response. Thus, the relevant welfare benchmark is not price of the biosimilar relative to the reference product but the price of the reference product before competition. Prices of biosimilars will be about 25–30% less than the original price of the reference products. Also, without competition one would have expected that the prices of biologics would increase each year. Thus, a 25% decrease in the original price of the reference product will underestimate consumer welfare gain. An added benefit to consumers is the increase in quantity due to lower prices which reflects an increase in access.

The savings to consumers and society could be much greater in the case of biosimilars than for generics because of their higher prices. For example, Revlimid which treats multiple myeloma had an annual cost in 2015 of $128,666. A 30% decrease in price would result in savings on this drug of around $38,600 per patient. In contrast, Lipitor, one of the world's blockbuster drugs, lost patent protection in 2011. The annual cost for a 20 mg regimen of treatment with the statin Lipitor in 2011 was $1939. Even if the generic price were 90% below that of Lipitor, annual per patient savings would only be $1745 (Blackstone Fuhr 2017). Biosimilar competition is thus expected to result in substantial benefits even with modest price reductions. However, in markets for certain indications where there is considerable competition among originator biologics, prices will probably not decrease as much as in markets where monopolies exist.

A RAND study estimated that savings from biosimilar competition could amount to $44.2 billion in the U.S. over 10 years (Mulcahy et al. 2014). Generic drugs have saved nearly $1.5 trillion in healthcare costs between 2004 and 2013 (Generic Pharmaceutical Association 2014). Given experience with biosimilars and increasing competition, the benefit could well be much greater over time. According to IMS "greater acceptance of biosimilar medicines in a growing number of therapy areas and an active pipeline of 56 new products in clinical development are expected to deliver total savings of as much as $110 billion to health systems across Europe and the U.S. through 2020" (Constantino 2016).

Another less obvious benefit from biosimilar competition may be a heightened interest in developing new drugs as existing drugs lose market share and profits.as a result of biosimilar competition. For example, Roche in 2017 was beginning to face biosimilar competition for its Rituxan, Herceptin and Avastin biologics. Its CEO indicated that Roche's strategy is based on innovation to counter the loss from biosimilar. He stated, "We stand now at 16 breakthrough therapy designations by the FDA" (Stanton 2017a). One could conclude that biosimilar competition is likely to increase pressure for developing new drugs to maintain profits and satisfy investors.

Biosimilar Risks and Return on Investment

An important issue that potential entrants into the biosimilar market must consider is whether they can get a return on their investment? Unlike the generic market where the first entrant can often get a return on investment in the first six months due to exclusivity, the same is not true for biosimilars. The R&D cost of biosimilars are much larger than that of generics. Thus, it should take much longer to recoup their R&D costs. However, given the higher prices of biologics even if the margins are lower, the profit generated can be higher. This along with the large markets, many exceeding a billion dollars, can result in the potential for substantial profits.

The bigger the blockbuster the greater the profits and the greater incentive for biosimilars to enter, which will result in greater competition and greater discounts. On the other hand, the larger the market the more the reference product firm will defend its patents and employ other competitive tactics. The smaller the market revenues, the less is the incentive to enter, which could result in lower discounts and potentially higher profits. Biosimilar firms must decide which markets to enter. Complicating the decision is the possibility that reference producers may on occasion choose to price lower than the short-run profit maximizing price to deter or slow down entry.

Each company must perform due diligence to determine what its biosimilar portfolio should be. Each reference product is a unique market with many variables that must be taken into consideration. Many firms are making huge investments to develop biosimilars. Firms have considerable investments in the form of R&D, manufacturing, clinical trials and other costs. They must consider the different investment needs for each reference product as well as how many competitors they will likely face and the potential price decrease in each market as well as price competition from the reference product.

PNS Pharma claims that there are at least 33 versions of Humira biosimilars in the clinical pipeline and more than 10 in phase III (Stanton 2016). Also, at least 21 Herceptin, 13 Avastin, 35 Rituxan and 27 Enbrel biosimilars are being developed (Blackstone and Fuhr 2017). The FDA has 66 biosimilars programs enrolled for 20 different reference products (Welch 2016).

Let's look at the Humira market. If all 33 succeed, which is highly unlikely, and each gets relatively the same 3% share of the market, a $10 billion market would

result in around $300 million in revenue. If the price decreases by 50% the market revenue for each firm would be $150 million. Increased access could result in each firm achieving $200 million in revenue. In any event this figure is probably too low to achieve a positive return given the substantial investment needed to enter. So 33 firms cannot survive in the market and a majority of the firms will lose money. In most markets only 5 or 6 will survive and potentially get a return on their investment.

Biosimilar firms must also consider that many originators are developing second generation biobetters which will affect the demand for the biosimilar in that particular market. If the second generation is so much better, it can get a price premium and the market for the originator biologic and biosimiars could greatly diminish. Under such a scenario, the biosimilar will find it difficult to make a profit. However, with value based pricing if the second generation is only marginally better, the biosimilar will put pricing pressure on second generation biologics.

There is also uncertainty as to whether a biosimilar will obtain FDA approval. The initial Sandoz application for a Neulasta biosimilar received a complete response letter from the FDA rejecting it. However, Sandoz expects to address the issues and submit the study requested by the FDA in 2018 (Big Molecule Watch 2016a). Also, there is currently the 180-day required notice to the reference product producer before a biosimilar can enter the market but this should not pose a problem given the U.S. Supreme Court decision. Even if a biosimilar is approved it cannot enter until the 12-year exclusivity expires. Even more uncertain is the issue of when patents expire and their validity. Thus, a $200 million (or greater) investment could be made 10 years before any revenue is generated. Taking into account the time value of money, firms may need to recoup at least twice the amount invested in R&D. This also does not take into account the other costs of the biosimilar firms. As many have stated biosimilars are not for the faint of heart.

The Future of Biosimilars

Biosimilars are likely to become an increasingly important part of the biologics industry, much like generics in the pharmaceutical industry. As consumers and physicians become more aware and accepting of them, they are likely to yield greater price reductions and improve patient access. However, unlike the generic industry the originator will be an important competitor in the market. The pathway for biosimilars will become more standard as the FDA gains more experience. Entry barriers will remain greater than for generics and patent disputes will continue. Many firms will fail but a relative few will succeed. In certain markets second generation biologics will control the market, leaving little opportunities for biosimilars. Most mature markets will result in the originator and around 5 biosimilar firms.

Conclusion

The biologic market is now in the process of having biosimilar competition. The BPCIA interpretation and the issue of patents, market exclusivity and other barriers to entry need to be considered in the context of balancing both the incentive for innovation and stimulating competition. As is true with new and developing markets, there are more competitors than can survive long term. As in all pharmaceutical market there will be a few winners and many losers.

The patent dance is an interesting attempt to resolve patent disputes so as to facilitate protection for the reference producer when appropriate but also to permit entry by biosimilar by clarifying the patent situation. Permitting only reasonable royalties in cases of infringement of patents undergoing the entire patent dance procedures could at least reduce the damages from biosimilar entry in the case of disputed patents. The 180-day notice before biosimilar entry has merit and normally would be expected to occur within the 12-year market exclusively. The market exclusivity itself seems appropriate given the time required to obtain marketing approval.

We expect that after physicians and patients become aware of and more knowledgeable about biosimilars, price reductions should be greater than the current 15%. Since established pharmaceutical companies will often be producing biosimilars, this should increase the likelihood of more substantial price reductions attributable to biosimilars. Reducing the time to approval of biosimilars could enable greater price reductions. Establishing clear guidelines for entry of biosimilars could also be helpful. Patent disputes are likely to be part of the industry but limiting damages for infringement in some cases seems appropriate. However, it is important to note the great risk that biologic reference producers take in developing biologics. Drawing the appropriate balance between innovation and competition is difficult.

References

Amicus Brief of United States Solicitor General in Sandoz v. Amgen, March 10, 2017.

Barlas S. Early biosimilars face hurdles to acceptance, the FDA has approved few, so lack of competition is keeping prices high. Pharm Ther. 2016;41(6):362–5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894512/.

Big Molecule Watch. Approval of Sandoz's pegfilgrastim may be delayed until at least 2018, October 28, 2016a. http://www.bigmoleculewatch.com/2016/10/28/sandozs-pegfilgrastim-approval-delayed/.

Big Molecule Watch. Janssen v. Celltrion: Janssen appeals judgement invalidating the '471 patent, October 24, 2016b. http://www.bigmoleculewatch.com/2016/10/24/janssen-v-celltrion-janssen-appeals-judgment-invalidating-the-471-patent/.

Blackstone E, Fuhr J. The future of competition in the biologics market. Temple J Sci Tech Environ Law. 2012;31:1–30.

Blackstone E, Fuhr J. Economics of biosimilars. Am Health Drug Benefits. 2013;6(8):469–77.
Blackstone E, Fuhr J. Biologics and biosimilars: the possibility of encouraging innovation and competition. The SciTech Lawyer. 2015;11(3):4–7.
Blackstone E, Fuhr J. Biosimilars and biologics; the prospect for competition. In: Endreny L, Declerck P, Chow S, editors. Development of biosimilar drug products. Boca Raton: CRC Press/Taylor & Francis; 2017. p. 413–38.
Brill A. Lost prescription drug savings from use of REMS program to delay generic market entry. Matrix Global Advisors, July 2014. http://www.gphaonline.org/media/cms/REMS_Studyfinal_July2014.pdf.
Buer L, Moum B, Cvancarova M, Warren D, Medhus A, Høivik M. Sitching from Remicade® to Remsima® is well tolerated and feasible: a prospective, open-label study. J Crohns Colitis. 2017;11(3):297–304.
Carrier M, Minnit C. Biologics: the new antitrust frontier. Univ Ill Law Rev. 2018;1.
Celltrion Healthcare Co., Ltd. and Celltrion, Inc. Plaintiffs, -against- Kennedy Trust for Rheumatology Research (formerly known as The Mathilda and Terence Kennedy Institute of Rheumatology Trust), United States District Court Southern District of New York 2014 Defendant. Case No.: 1:14-cv-02256 (PAC) *Oral Argument Requested* Plaintiffs' memorandum of Lawin opposition to defendant Kennedy trust's motion to dismiss the complaint or to stay the action. http://www.fdalawblog.net/REMICADE%20-%20Celltrion%20Opp%20re%20Kennedy%20Trust%20MTD%20Celltrion%20DJ%20Complaint.pdf.
Cohen M. Sourcing innovator products in the age of biosimilar research. Generic and Biosimilars Initiative, June 1, 2015. http://www.gabionline.net/Sponsored-Articles/Sourcing-innovator-products-in-the-age-of-biosimilar-research.
Constantino T. IMS Health: surge in biosimilars to drive significant change in health system costs, patient access and competition by 2020. IMS Health, March 28, 2016. http://www.imshealth.com/en/thought-leadership/quintilesims-institute/news-and-press/ims-health-surge-in-biosimilars-to-drive-change-in-health-system-costs.
Crespi-Lofton J, Skelton J. The growing role of biologics and biosimilars in the United States: perspective from the APhA biologics and biosimilars stakeholders conference. J Am Pharm Assoc. 2017;57:e15–27.
Crow D. Pfizer takes on J&J with antitrust lawsuit. Financial Times, September 2017.
Dinwoodie N. Biobetters and the future biologics market. BioPharm. November 1, 2011. www.biopharminternational.com/biopharm/Upstream+Processing/Biobetters-and-the-Future-Biologics-Market/ArticleStandard/Article/detail/746281.
Erickson B. Untangling biosimilars. Chem Eng News. 2010;88:25–7. https://www.uspharmacist.com/article/biosimilars-current-approvals-and-pipeline-agents#sthash.eTMQdnaS.dpuf.
Fischer A. U.S. biosimilar litigation so far. Biologics Blog, February 1, 2016. https://www.biologicsblog.com/u-s-biosimilar-litigation-so-far/.
Frois C, Mortimer R, White A. The potential for litigation in new era of biosimilars. Law360, September 20, 2016. https://www.law360.com/articles/842318/the-potential-for-litigation-in-new-era-of-biosimilars.
Generic Pharmaceutical Association. Generic drug savings in the U.S., Sixth Annual Edition. 2014. http://www.gphaonline.org/media/cms/GPhA_Savings_Report.9.10.14_FINAL.pdf.
Generics and Biosimilar Initiatives. Huge discount on biosimilar infliximab in Norway March 13, 2015. http://www.gabionline.net/Biosimilars/General/Huge-discount-on-biosimilar-infliximab-in-Norway.
Gordon L. The Supreme Court delivers a win for biosimilar manufacturers in Sandoz v. Amgen IP Intelligence, June 28, 2017.
Grabowski H, Long G, Mortimer R. Data exclusivity for biologics. Nat Rev Drug Discov. 2011;10:15.
Helfand C. Novartis' Sandoz aims for 5 biosim launches by 2020. FiercePharma, June 20, 2016. http://www.fiercepharma.com/pharma/novartis-sandoz-aims-for-5-biosim-launches-by-2020.

Helfand C. Key Humira patent goes down via IPR-but biosim makers shouldn't celebrate yet, analysts warn. FiercePharma, May 17, 2017a.

Helfand C. Merck KGaA in advanced talks to sell off biosims unit—and others could follow. FiercePharma, March 9, 2017b. http://www.fiercepharma.com/pharma/merck-kgaa-advanced-talks-to-sell-off-biosims-unit-and-others-could-follow?utm_medium=nl&utm_source=internal&mrkid=1008892&mkt_tok=eyJpIjoiTVRCbE5HVTVOekprTVRVNCIsInQiOiJVaEtmTngzbm1LXC9ZNWpKa2lZMys1bG10QXNSMU9zdzdnMWZHaVJRYm5IXC9MUW5WQXZtRmptOEhrXC9sb3lcL0tleDlUemx1ZUY0b25VSWZ2WTdlMTB3R3c5bGtTSnBCSmExTnd0Z2JqekpLQ2hha3hkNzVxT1VaRExwYnFmQTF1bE4ifQ%3D%3D.

IMS. Delivering on the potential of biosimilar medicines: the role of functioning competitive markets. March 2016. http://www.imshealth.com/files/web/IMSH%20Institute/Healthcare%20Briefs/Documents/IMS_Institute_Biosimilar_Brief_March_2016.pdf.

IMS Institute for Healthcare Informatics. Declining medicine use and costs: for better or worse? A review of the use of medicines in the United States in 2012. May 2013. www.imshealth.com/deployedfiles/ims/Global/Content/Insights/IMS%20Institute%20for%20Healthcare%20Informatics/2012%20U.S.%20Medicines%20Report/2012_U.S.Medicines_Report.pdf.

Janssen v. Celltrion, Damages: 'Patent Dance' May Determine Availability of Lost Profits. Lexcology, March 29, 2017a.

Janssen v. Celltrion: A Misstep in the BPCIA Patent Dance. Intellectual Property Alert, March 8, 2017b.

Kanavos P. Generic policies: rhetoric vs. reality. Euro Observer. 2008;10(2):1–6.

Lagaso J. Biosimilar drugs often translate to lower provider profits, says Navigant analysis. Healthcare Finance News, July 12, 2017.

Loo J. Standard and Poor's industry survey: biotechnology. S&P Capital 1Q, McGraw Hill Financial March 2015.

Loo J. Standard and Poor's industry survey: biotechnology. S&P Capital 1Q, McGraw Hill Financial August 2016.

Loo J, Biotechnology, Standard and Poor's industry survey: biotechnology. S&P Capital 1Q, McGraw Hill Financial, March 2017.

MacDonald G. Amgen developing Soliris biosimilar. Biopharma-Reporter.com, June 21, 2016. http://www.biopharma-reporter.com/Markets-Regulations/Amgen-developing-Soliris-biosimilar.

Mehr S. Pegfilgratim: 0 for 3 on biosimilars at FDA. Biosimilars Review and Report June 13, 2017a.

Mehr S. Pharmaceutical patent litigation and the emerging biosimilars: a conversation with Kevin M. Nelson. Am Health Drug Benefits. 2017b;10(1):23.

Mulcahy AW Predmore Z Mattke S. The cost savings potential of biosimilar drugs in the United States RAND. 2014. https://www.rand.org/content/dam/rand/pubs/perspectives/PE100/PE127/RAND_PE127.pdf.

Palmer E. Mylan and Biocon biosimilar of Amgen's Neulasta hit with manufacturing–related CRL. FiercePharma, October 10, 2017a.

Palmer E. Mylan and Biocon Herceptin biosimilar delayed by FDA for 3 Months. FiercePharma, September 1, 2017b. http://www.fiercepharma.com/pharma/mylan-and-biocon-herceptin-biosim-delayed-by-fda-for-three-months.

Palmer E. FDA rejects Pfizer's Epogen biosimilar for the second time. FiercePharma, June 22, 2017c.

Pollock A. Makers of Humira and Enbrel using new drug patents to delay generic versions. New York Times, July 15, 2016. https://www.nytimes.com/2016/07/16/business/makers-of-humira-and-enbrel-using-new-drug-patents-to-delay-generic-versions.html?_r=0.

Quinn G. Industry reaction to SCOTUS decision in Sandoz v. Amgen. IP Watchdog, June 14, 2017.

Reiffen D, Word M. Generic drug industry dynamics. Rev Econ Stat. 2005;87:37–49.

Royzman I. Hurdles for Neulasta biosimilars, Biologics Blog, February 13, 2017.

Sagonowsky E. Thanks to Amgen's biosimilar settlement AbbVie's $20 billion forecast looks real. FiercePharma, September 28, 2017a.

Sagonowsky E. The decade's top 10 patent losses, worth a whopping $915 billion in lifetime sales. FiercePharma, August 17, 2017b.

Sagonowsky E, AbbVie files another Humira patent suit, this time against Boehringer Ingelheim. FiercePharma, August 3, 2017c.

Sagonowsky E, Roche's hold on Avastin, Herceptin sales loosens as FDA panel backs Amgen and Mylan biosims. FierceBiotech, Drug Development, July 13, 2017d.

Sagonowsky E. J&J fires off legal salvo to block Remicade biosim launch. FiercePharma, May 19, 2017e.

Sagonowsky E. Roche's Herceptin one step closer to biosimiliar battle with Mylan licensing deal. FiercePharma, March 13, 2017f. http://www.fiercepharma.com/pharma/mylan-roche-make-nice-herceptin-biosim-settlement?utm_medium=nl&utm_source=internal&mrkid=1008892&mkt_tok=eyJpIjoiTXpnNVl6Wm1aak01WkRBeCIsInQiOiJFY0pXTnhqUnFJTmkrc1pWV2E5Z3A0QU9KM0MxNWlHdmpxMjJ0MkNodkhwd3FVMGtXaWUza1hJVllWN2RKdzlzVE5Cb0dDK1JmOUoyK21WeVAzc3RXVW5YXC9tQXY4TjFpTjFJXC8zQUpUM1p3NHB3QU4yTFNQOWJsSkNDYThlNkpzIn0%3D.

Sarpatwari A, Avorn J, Kesselheim A. Progress and hurdles for follow-on biologics. N Engl J Med. 2015;372:2380–2.

Serebov M. Biosimilar uptake in the US could outpace adoption in the EU. Thomas Reuter. 2016. http://stateofinnovation.thomsonreuters.com/biosimilar-uptake-in-the-us-could-outpace-adoption-in-the-eu.

Serebrov M, A bioworld report biosimilars: U.S. market opportunities and critical strategies. Thomas Reuter. 2016. http://images.info.science.thomsonreuters.biz/Web/ThomsonReutersScience/%7Bacfdc5b8-c04e-46ec-b1ea-a34db8bc411a%7D_biosimilars_premium.pdf.

Shin F. Leaping from the 'patent cliff' into the global drug gap: overcoming exclusivity to provide affordable biosimilars. Loyola Los Angel Int Comp Law Rev. 2016;37:419.

Sierra J, Szer J, Kassis J, Herrmann R, Lazzarino M, Thomas X, Noga J, Baker N, Dansey R, Bosi A. A single dose of pegfilgrastim compared with daily filgrastim for supporting neutrophil recovery in patients treated for low-to-intermediate risk acute myeloid leukemia: results from a randomized, double-blind, phase 2 trial. BMC Cancer. 2008;8:195. https://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-8-195.

Singh S. Basics about biosimilars. CVS Health, May 3, 2016. http://insights.cvshealth.com/basics-about-biosimilars.

Stanton D, Humira biosimilar developers: can you find who's taking on AbbVie?. BioPharmaReporter.com, December 23, 2016. http://www.biopharma-reporter.com/Markets-Regulations/Humira-biosimilar-developers-can-you-find-who-s-taking-on-AbbVie?utm_source=newsletter_daily&utm_medium=email&utm_campaign=23-Dec-2016&c=1rbVdiOlqvWRNV239QvUfUAf06V120%2Fo&p2.

Stanton D. Roche offsetting biosimilar competition with 16 breakthrough therapies. Biopharma-Reporter.com, July 28, 2017a.

Stanton D. J&J confident of retaining Remicade rewards as biosimilar clouds gather. Biopharma Reports, May 31, 2017b.

Staton T, Roche's new Gazyva beats predecessor Rituxan at staving off CLL. FiercePharma, November 7, 2013. http://www.fiercepharma.com/pharma/roche-s-new-gazyva-beats-predecessor-rituxan-at-staving-off-cll.

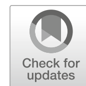
Staton T, Hep C meds cost-effective, but affordable? With new 46% discounts, maybe. FiercePharma, March 17, 2015. http://www.fiercepharma.com/sales-and-marketing/hep-c-meds-cost-effective-but-affordable-new-46-discounts-maybe.

Staton T, Pfizer's new Inflectra biosim data puts J&J's Remicade on notice in Crohn's. FiercePharma, February 21, 2017. http://www.fiercepharma.com/marketing/pfizer-s-new-inflectra-biosim-data-puts-j-j-s-remicade-notice-crohn-s.

U.S. Senate, Special Committee on Aging. Sudden price spikes in off-patent prescription drugs: the monopoly business model that harms patient, taxpayers, and the U.S. Health Care System, December 2016.

Welch AR. How to make the most of your FDA biosimilar meetings. Biosimilar Development, November 14, 2016. https://www.biosimilardevelopment.com/doc/how-to-make-the-most-of-your-fda-biosimilar-meetings-0001.

Chapter 3
The Changing US Reimbursement Landscape and Biosimilars



Molly Burich

Abstract The continued rise of prescription drug spend, in particular specialty drugs has created the need for high-quality, low-cost options for biologics. The passage of the Affordable Care Act (ACA) created an FDA approval pathway for biosimilars, but that is the tip of the iceberg as it pertains to how a biosimilar will actually get to the patient. As a general rule, legislation contains very little detail and requires much regulatory guidance by various government agencies to be developed to facilitate a product coming to market. The Centers for Medicare and Medicaid Services (CMS) is the primary federal agency to provide guidance on the Medicare coverage, coding and payment for biosimilars in the US. CMS will also provide some guidance on the use of biosimilars in state Medicaid programs; however, much of that detail will be determined by the state program. Commercial health plans will also weigh in with their own coverage decisions, but typically follow the coding of biosimilars as set forth by CMS. This chapter will assess the existing reimbursement landscape for biosimilars and will address key considerations and implications around current decisions that will impact biosimilars' market entry in the US.

In many ways, payers, governmental and commercial, will determine the ultimate success of biosimilars in the US by developing benefit designs and provider and patient incentives to drive biosimilar use. This chapter will explore how lessons on biosimilar incentive structures can be gleaned from the European experience and highlight what payer activity we have seen to date with marketed biosimilars in the US.

Disclaimer: The views, opinions and statements made in this publication are solely those of the author and may not reflect the views of Boehringer Ingelheim Pharmaceuticals, Inc. or its affiliates.

M. Burich (✉)
Boehringer Ingelheim Pharmaceuticals Inc., Danbury, CT, USA
e-mail: molly.burich@boehringer-ingelheim.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_3

Keywords Medicare · Medicaid · Reimbursement · Biosimilars · Commercial payers · Payment · Coding · Physicians · Patients · Manufacturers · Interchangeability · J-code · Coverage gap

In the US, specialty drug spend continues to increase; a trend which many stakeholders believe is unsustainable to support future innovation. With both the rise in available specialty medications as well as conditions like cancer becoming more chronic, long-term conditions, spending has exceeded targets for the past several years. Additionally, specialty drug approvals by the Food and Drug Administration (FDA) have risen significantly over the past several years, with 2010 beginning the first of several years where the number of approved specialty drugs exceeded traditional small molecule drugs (PwC Health Research Institute 2016). First, let's begin by defining a specialty drug. A specialty drug is a prescription product that has one or more of the following: special conditions for storage, handling or administration and/or monitoring. Typically, specialty drugs are biologics, injectable or infused, though increasingly they may also be oral (Health Affairs–Health Policy Briefs 2013). In part because of their handling requirements, but also their development, specialty drugs are often more costly than traditional small-molecule drugs. Why is this trend significant to biosimilars? Biosimilars represent the most significant way to increase competition and bring lower cost alternatives to market.

Reimbursement Overview

Reimbursement is a broad term that typically encompasses the coverage, coding and payment for a specific product. It is a critical issue for both specialty drugs broadly and specifically biosimilars. Let's walk through each component as a foundation of understanding.

Coverage

Typically means any requirements put in place by a health plan or payer that the patient must meet in order to access the drug. Coverage will be primarily determined by the benefit in which it is covered—either the medical or pharmacy benefit (Table 3.1).

Coding

Drug codes can come in a variety of different forms and are the primary communication between the healthcare provider (HCP) and the payer. Codes are used to communicate which drug the patient is receiving.

Table 3.1 Overview of medical and pharmacy benefits

	Medical benefit	Pharmacy benefit
Route of administration	Usually physician-administered injectable and infused drugs	Usually all self-administered injectable and oral drugs
Acquisition	Buy and bill model is most prevalent; physician buys the drug and bills the payer after administering the drugs	Specialty pharmacy (SP) is the most common method, though some products may be available via special arrangements between an SP and a retail location
Site of care	Physician office Hospital outpatient department Stand-alone infusion center	Patient's home
Reimbursement	Primarily average sales price (ASP)	Typically based on a negotiated price between the payer and manufacturer or wholesale acquisition cost (WAC)
Patient responsibility	Typically a set percentage (e.g., 80/20 split) determined by the patients' plan	Many small molecule drugs pay a flat copayment; increasingly specialty drugs have a coinsurance (e.g., a certain percentage) on them
Payer restrictions	Payers typically manage medical benefit drugs less than pharmacy benefit but may implement pre-certification requirements; Physician incentives may exist to drive utilization of specific products	Payers have a series of utilization management tools: Prior authorization (PA), step-edits, quantity limits, tier placement, cost-share and SP use
Required codes	J-codes	National drug codes (NDCs)

Note: generalizations about both the pharmacy and medical benefit

Healthcare Common Procedure Coding System (HCPCS) Level II codes are used for medical benefit drugs. J-code is the most common of the HCPCS codes used for drugs. J-codes are alpha-numeric, five digit codes that being with "J" and are followed by four numbers. The Centers for Medicare and Medicaid Services (CMS) own the HCPCS coding process, meaning CMS controls the development and implementation of drug-specific coding. All payers—including commercial and Medicaid programs typically follow CMS' lead on use of J-codes. Typically a product only has one J-code, and it is often designated at the appropriate strength, for example, "1 mg" or "10 mg"level. For example, the J-code for Remicade® is J1745—"Injection, infliximab, excludes biosimilar, 10 mg" (Janssen Pharmaceuticals 2016). Physicians will utilize the J-code and bill the appropriate number of units, utilizing the Remicade example, if the doctor dosed 100 mg, they would bill 10 units.

Pharmacy benefit codes for drugs are different; pharmacy claims typically utilize National Drug Codes (NDCs). NDCs are generated and maintained by the FDA upon approval and/or introduction of new formulations/dosing. NDCs are very specific and a single drug can have multiple NDCs underneath them based on the number of available products.

Payment

Payment is perhaps the most critical component of reimbursement; however, the first two components ultimately drive the accuracy and timeliness of a payment.

Payment for medical benefit drugs encompasses two components: (1) what the HCP is paid and (2) what the patient owes. As outlined in the table above, typically payment for medical benefit drugs is based on ASP. ASP is based on a series of data points including rebates and discounts and is updated quarterly by CMS. Manufacturers of medical benefit drugs submit quarterly information to CMS and CMS releases a file of all medical benefit drugs with ASPs. Similar to the coding process, CMS owns this calculation process; however, other payers including commercial and Medicaid programs utilize ASP to pay for products.

Patients cost share will be highly dependent on their specific plan. However, typically medical benefit cost-share is a percentage—example, 20% is owed by the patient for in-network drugs and services.

Payment for pharmacy benefit drugs is similar with two main components: (1) what the pharmacy is paid and (2) what the patient owes. Pharmacy benefit drugs are typically based on a negotiated price (e.g., rebates and discounts) between the payer (or pharmacy benefit manager—PBM) and adjudicated upon the pharmacy submitting the claim for payment. This is the general process whether it is retail or SP submitting the claim.

Patient cost-share depends on a variety of factors—including tier placement and benefit design. For small molecule drugs—the patient will typically pay a flat copayment amount determined by tier (e.g., tier 1: $5; tier 2: $20). Increasingly, high-cost, specialty drugs covered on the pharmacy benefit require a coinsurance from the patient. Coinsurance is often based on the list-price (or WAC), which is pre-discounts. So a patient could have a 20% coinsurance monthly on a specialty medication until they hit their annual out-of-pocket (OOP) limit for the year. Coinsurance can be quite costly for patients who are also juggling coinsurance on any medical benefit services or drugs as well.

In summary, reimbursement is a complex set of considerations and highly dependent on the payer and their requirements.

Medicare Reimbursement and Biosimilars

Now that we have a basic foundation for reimbursement in the US system, it is time to dive into biosimilar reimbursement considerations by payer. In many ways, CMS is a top decision-maker in terms of reimbursement policy for biosimilars.

Medicare Part B

Medicare Part B is the benefit that covers services within the physician office and hospital outpatient setting. Additionally—as discussed, the Part B benefit includes payment for physician-administered drugs that are administered in either the physician office or hospital outpatient department.

Coverage

As with other parts of Medicare, CMS national provides regulatory guidance to its' Medicare Administrative Contractors (MACs) who are responsible for actually managing, processing and paying claims for both Medicare Parts A and B. There are 12 MACs who are multi-state and regionally based. MACs however, do not act as a traditional payer in the sense that they have very limited ability to restrict or manage access to drugs or services, but in particular to drugs. As such, as long as a drug meets the requirements outlined below, it is eligible for reimbursement under Part B (Fig. 3.1).

Coding

As discussed previously, CMS controls the assignment of HCPCS codes for physician-administered drugs. Through a combination of legislation (e.g., the Social

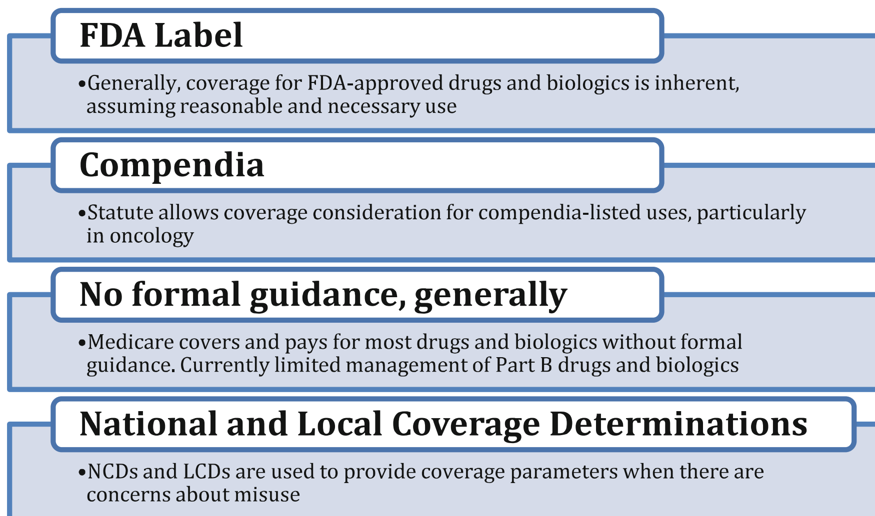


Fig. 3.1 ADVI Health, 2017

Security Act (SSA) section 1847a) and regulation, CMS has developed a series of interpretations which guide their coding process. For new brand products, CMS generally considers them to be "single-source" products and therefore assigns a unique J-code and payment rate to those products. The HCPCS process for assigning J-codes is an annual process with CMS assigning codes mid-year with the codes taking effect the following January. Notably, outside of the formal HCPCS process, CMS can assign a "Q-code" which is situated like a J-code, but is typically designed to be a temporary code for a drug or service.

In the CY2016 Medicare Physician Fee Schedule (MPFS) proposed rule, released in July 2015, CMS surprised many stakeholders by establishing a coding and payment policy for biosimilars. At this point, there was only one approved, unlaunched biosimilar in the market, Zarxio® (filgrastim-sndz) by Sandoz. In the proposed rule, CMS outlined that for the purposes of coding and payment, CMS would consider biosimilars utilizing a generic-drug framework, as "multi-source" drugs and therefore, all biosimilars for the same reference product would share a J-code and a blended ASP-based payment rate. At the same time, reference products would retain their own separate J-code and payment rate. In the proposed rule, CMS sought public comment on this proposed policy. CMS received many comments on this policy, the vast majority of which stood in opposition of this policy. Opposition was heard from physician groups, patient advocacy groups, Congress (Biosimilar Development 2017a), commercial payers and manufacturers. In November 2016 in the final CY2016 MPFS, CMS opted to finalize this policy and it took effect January 1, 2016. Interestingly, CMS has actually yet to assign any J-codes for biosimilars, only Q-codes have been assigned.

The launch of Renflexis™ (infliximab-abda) the second biosimilar to reference product Remicade®, put this policy into play in July 2017. Since Inflectra® (infliximab-dyyb) the first biosimilar to Remicade launched in October 2016, there was an established Q-code. Presumably, now that Renflexis™ has launched, all claims for both Remicade® biosimilars will be submitted utilizing Q5102, the Q-code first assigned to Inflectra®.

In conjunction with finalizing its' policy for blended J-codes, CMS noted that the tracking concerns (described below) of a blended J-code warranted some type of a solution. CMS decided to establish a series of manufacturer-specific modifiers that will be required to be appended to the claim form when a biosimilar claim is submitted (Centers for Medicare and Medicaid Services 2018a) (Table 3.2).

Stakeholder Considerations on the Blended J-Code Policy

While CMS did finalize the policy in November 2015, several key groups of stakeholders continued to express concern over the policy.

Table 3.2 Existing approved and launched biosimilars and their codes prior to January 1, 2018 (Centers for Medicare and Medicaid Services 2017a)

Biosimilar:	Reference product:	Code	Descriptor	Manufacturer and modifier
Zarxio® (filgrastim-sndz)	Neupogen®	Q5101	Injection, Filgrastim (G-CSF), biosimilar, 1 mg	Sandoz—ZA
Inflectra® (infliximab-dyyb)	Remicade®	Q5102	Injection, infliximab, biosimilar, 10 mg	Pfizer/Hospira—ZB
Renflexis™ (infliximab-abda)				Merck/Samsung—ZC

Physicians Physicians who bill Medicare Part B are required to "buy and bill" the products—meaning they purchase the product and bill Medicare upon administering it. This creates some interesting dynamics if the ASP of a product changes between the time they bought it and the time they bill Medicare (described in more detail below). This dynamic was likely to be a concern to physicians and could have resulted in driving them away from the instability of a blended reimbursement in favor of the reference product, which under existing policy is protected from these dynamics.

Patients The concern for patients, who are typically immune from any issues related to coding of drugs, is that under the policy, CMS is treating all biosimilar products equal. This ignores the potential variability in clinical profiles between biosimilars. The absence of product-specific J-code for each biosimilar could potentially hinder pharmacovigilance efforts designed to support patient safety.

Other Payers As discussed, the assignment of HCPCS codes is managed by CMS. As such, commercial (including the Exchanges) and Medicaid payers are beholden to CMS coding decisions. A primary concern for payers is that they are unable to effectively manage and/or implement product preferencing across biosimilars, because they all shared a code. This could limit commercial and Medicaid payers from being able to drive biosimilar uptake at the product level, they could only push biosimilars at the class level. This is likely not considered an ideal situation for payers, particularly as they continue to try to bend the cost curve associated with costs. Additionally, physician and patient concerns outlined above are relevant across all payers, not just Medicare.

Manufacturers Manufacturers, particularly those investing in biosimilars came out very strongly against the CMS coding and payment policy for biosimilars. That is primarily related to the investment required to make biosimilars. A generic drug typically takes 3–5 years to develop at a cost of \$1–\$5 M. However, a biosimilar takes 7–10 years to develop at a cost upwards of \$200 M (FTC 2009). The differences in time and cost reflect the complexity and increased analytical and clinical data requirements for a biosimilar compared to a generic drug. As such,

it is concerning that CMS applied the same coding and payment methodology to generics and biosimilar products when they are vastly different.

2018: A New Coding and Payment Policy Will Emerge

In July 2017, in the CY2018 MPFS proposed rule, CMS re-opened its' controversial coding and payment policy for biosimilars by seeking public comment. CMS received over 200 public comments, the majority of which asked the Agency to overturn the blended coding policy. As outlined above, given that number of stakeholders who opposed this policy, the outpouring of opposition was not surprising. In November 2017, CMS reversed the coding policy and stated (Centers for Medicare and Medicaid Services 2017b):

> Thus, in this final rule, we finalizing the policy to separately code and pay for biosimilar products under Medicare Part B ... Effective January 1, 2018, newly approved biosimilar biological products with a common reference product will no longer be grouped into the same HCPCS code.

In the final rule, CMS reviewed many of the arguments presented by commenters and concluded that the concerns around the long-term sustainability of the market were the most compelling to drive a policy change (Centers for Medicare and Medicaid Services 2017b).

> ... We have also considered how the payment policy could affect market entry of new biosimilar manufacturers. If payment amounts limit manufacturers' willingness to invest in the development of new biosimilars, it could in the long term, decrease the number of biosimilar biological products that are available to prescribe and thus impair price competition. Given that the United States' biosimilar biological product marketplace is still relatively new, we believe that it is important to maintain a payment policy innovation as well as reasonable pricing for consumers ... We believe that this policy change will encourage greater manufacturer participation in the marketplace and the introduction of more biosimilar products, thus creating a stable and robust market, driving competition and decreasing uncertainty about access and payment ...

In Q1 2018, CMS released guidance that included new instructions to MACs on appropriate coding for biosimilars and provided new, separate codes for Inflectra(R) (Q5103) and Renflexis(R) (Q5104). Source: https://www.cms.gov/apps/ama/license.asp?file=/Medicare/Medicare-Fee-for-Service-Part-B-Drugs/McrPartBDrugAvgSalesPrice/Downloads/2018-Oct-ASP-Pricing-File.zip

Payment

As outlined above, the payment dynamic in Medicare is such that it can create perverse financial incentives due to the financial realities of the buy and bill landscape. The Medicare Modernization Act (MMA) of 2003 ushered in the ASP payment system into Medicare. The primary reason for changing to ASP was because of concerns around abuse of the average wholesale price (AWP) methodology. The

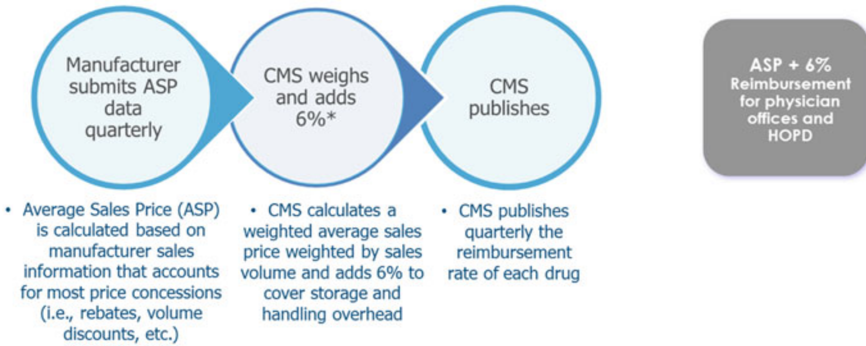


Fig. 3.2 ASP calculation and timing (ADVI Health 2017)

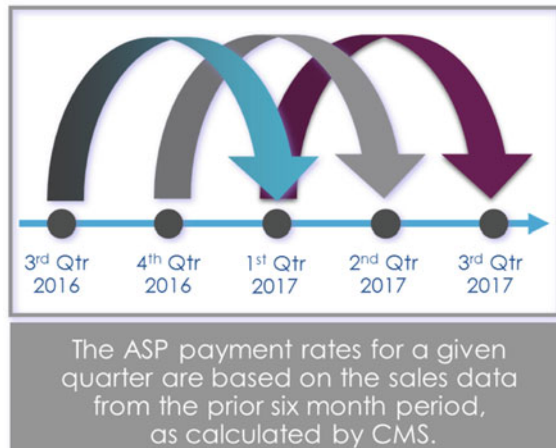


Fig. 3.3 Quarterly payment for drugs reflects 2 quarter lag (ADVI Health 2017)

goal of the ASP system was to develop a transparent, public system and take into account other rebates and discounts so the ability for physicians to make money on drugs via buy and bill was limited. Currently, in both the physician office and hospital outpatient department, Medicare pays for physician-administered drugs at ASP +6%. For biosimilars, Congress mandated that they be paid at the same rate, except that the 6% comes from the reference product ASP, this was designed to be a physician incentive to utilize biosimilars (Figs. 3.2 and 3.3).

The ASP system comes with its' own challenges for physicians, the most significant of which is the 2 quarter lag that is associated with payment. While manufacturers submit their data quarterly, the published ASP is always based on 2 quarters prior, to allow CMS time to actually manage the data as evidenced above.

It is possible that by the time a drug is administered, the ASP of the product has changed and the physician may be financially "underwater" on the drug, meaning they will get reimbursed less than what they paid for. This dynamic can occur with any drug, but the risk was set to be further multiplied when there were multiple

drugs blended into one code. With the revision to the coding policy, biosimilars will be assigned their own payment rate and thus mitigate some of the concerns around being financially underwater.

In summary, biosimilars could be a very promising development to curb costs for Medicare Part B. The changes to the coding and payment policy are anticipated to generate a more sustainable and competitive long-term biosimilar market in the US which will help curb costs.

Medicare Part D

Medicare Part D is the newest benefit within Medicare. The MMA of 2003 which ushered in ASP payment in Part B, also included the development of a prescription drug benefit in Medicare, known as Medicare Part D. Prior to this, there was no coverage for pharmacy benefit drugs in Medicare. Part D is administered by private health plan companies that form their own formularies but CMS national sets certain rules. Part D plans must cover two drugs per USP class/category and cover substantially all drugs in the six protected classes.

CMS has review and oversight of the formularies, but plans can make positive formulary changes (e.g., removing a utilization management tool, decreasing cost-sharing and/or lowering tier placement) any time of year without CMS permission. However, plans are not permitted to make "negative" changes during the benefit year without CMS approval. Additionally, CMS sets forth specific review period, timelines and an appeals process that all Part D plans have to follow.

When Part D was established, it included a provision called the "coverage gap" also referred to as the "donut hole." The coverage gap is a period of time where patient cost-share is significantly higher than other phases of coverage. The original intent of the coverage gap was (1) to keep the cost of implementing Part D low and (2) to ensure patients have "skin in the game" when it comes to paying for their medicines. Annually, CMS sets the monetary threshold for the coverage gap, referred to as the patients true-out-of-pocket or TrOOP. The figure below illustrates the dollar thresholds for the coverage gap in 2017 (Fig. 3.4).

Low-income subsidy patients do not ever hit the coverage gap, the coverage gap is applicable only to those patients who have not received the LIS subsidy or are dual-eligible (e.g., eligible for Medicare because of age/disability and Medicaid due to low income).

The last important facet of the coverage gap is the ACA-mandated creation of the Coverage Gap Discount Program (CGDP). As a result of the coverage gap thresholds increasing every year and the concern of the financial strain it puts on patients, most of whom are on a fixed income, the ACA included the CGDP provision to help patients through this phase of coverage. The CGDP included a provision that requires the manufacturer of brand products to pay a 50% rebate on their drug while the patient is in the coverage gap. Thus, in 2017, during the coverage gap for a brand product, the manufacturer pays 50%, the patient pays 40% and the plan pays 10%. Additionally, the 50% rebate paid by the manufacturer

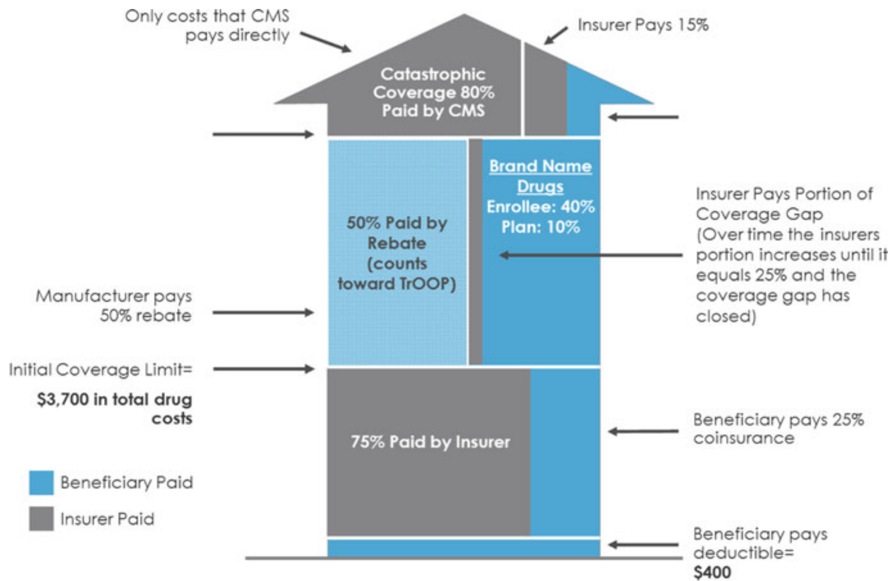


Fig. 3.4 Overview of Part D coverage gap (ADVI Health 2017)

counts towards the patients TrOOP. By 2020, the coverage gap "closes" meaning it equalizes patient cost-share to a consistent level of 25% throughout the year until the patient hits catastrophic coverage (in 2017 that dollar amount is $4950).

The reason why the coverage gap and the CGDP are so critical to biosimilars lies within the text of the ACA. The text of the CGDP excludes both generic and biosimilar (any products approved under the 351k pathway) products. For generic products, which are typically very low cost, this is a logical exclusion—a manufacturer wouldn't need to pay a 50% rebate on a very low cost drug. However, it is an entirely different consideration for biosimilars. Even if biosimilars come in at a significant discount, that may still be a discount on a relatively high priced product to begin with, so it is entirely possible a discounted biosimilar could be several thousand dollars. Therefore, the exclusion of biosimilars from the CGDP is very concerning for both patients and Part D plans.

What Does this Mean for Patients? (Biosimilar Development 2017b)

- The impact of biosimilars being excluded from the CGDP is twofold for patients:
  – They face a higher cost share for a biosimilar
  – They remain in the donut hole longer

- While the patient is in the donut hole, the 50% manufacturer rebate on branded drugs counts toward the patient's TrOOP costs, thus moving the patient through the donut hole faster and helping to reduce the cost to patients.
- Since biosimilars are excluded from the CGDP, any discount offered to the plan on the biosimilar would not count toward TrOOP and therefore, the patient has both higher cost share and is in the donut hole longer than if they were on a branded drug.
- Notably, the coverage gap closes for patients in 2020; patients will pay 25% of the cost of a branded or generic/biosimilar drug once they reach the coverage gap. However, the impact to TrOOP will still result in the patient being in the coverage gap longer if they take a biosimilar.

What Does this Mean for Part D Plans?

- The impact of biosimilars being excluded from the CGDP is:
  - As evidenced in Fig. 3.1 below, the cost of the biosimilar will be higher to the plan through 2020 and beyond:

    Plan responsibility for a brand product in the donut hole in 2020: 25%
    Plan responsibility for a generic product in the donut hole in 2020: 75%

- The result of this policy makes it potentially very unlikely for a Part D sponsor to push the use of biosimilars under Medicare Part D

Figure 3.5 illustrates how the exclusion of biosimilars from the CGDP is an immediate concern for patients through 2020 and remains a concern for health plans:

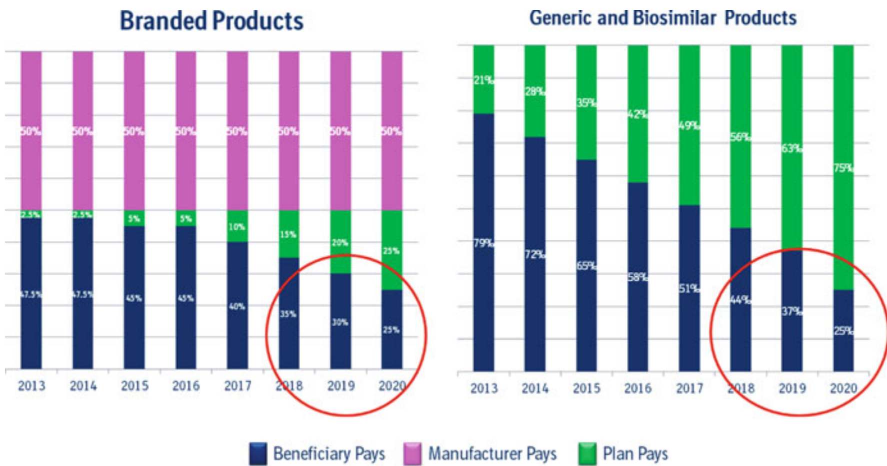


Fig. 3.5 Stakeholder responsibility for coverage gap costs

Solutions have been suggested though not implemented. In April 2016, Avalere Health conducted an independent analysis on this topic and outlined two potential solutions (Avalere Health 2018):

- Allow biosimilars to participate in the CGDP and pay the 50% rebate
- Allow CMS to create a new biosimilar tier (and necessary program rules) to give Part D plans flexibility on how they manage (e.g., coverage, specify cost share, etc.) biosimilars

Note that either change requires legislative action. While both represent potential solutions, the first is the more likely and timely solution as it can be utilized for approved, but not-yet-marketed, biosimilars under Medicare Part D. As the U.S. biosimilars market is in its infancy, reimbursement policies to appropriately support their use are essential, and CMS is positioned to play a pivotal role in the development of these policies. Addressing the coverage gap in Part D is a critical place to start.

In November 2017, CMS released an off-cycle proposed rule pertaining to CY2019 Part D and Medicare Advantage. In the rule, CMS proposed changes to cost-sharing levels for biosimilar (Centers for Medicare and Medicaid Services 2017c). CMS proposed increased flexibility for plan sponsors to categorize biosimilars as generics for the purposes of cost-sharing:

- For Low-Income Subsidy (LIS) patients—throughout the plan year
- For non-LIS beneficiaries—during the catastrophic phase of coverage

CMS finalized this policy in early 2018, these provisions provide Part D plans the flexibility to drive biosimilar utilization.

In February 2018, a provision was included in the Bipartisan Budget Act of 2018 to make biosimilars “applicable drugs” for the purposes of the Coverage Gap Discount Program effective January 1, 2019. This change is significant for biosimilars to be on a level playing field with the reference product.

Medicaid Reimbursement and Biosimilars

With the expansion afforded by the ACA, Medicaid programs have expanded significantly since 2014. As of May 2017, approximately 68.8 M individuals were enrolled in Medicaid (another 5.7 M were enrolled in the Children’s Health Insurance Program [CHIP]), this is an increase of approximately 17 M since the implementation date of the ACA-included expansion (Centers for Medicare and Medicaid Services 2018b). To-date 32 states including Washington DC expanded their programs under the ACA. Expansion allowed states to expand to both a higher income threshold as well as a new population of childless adults.

Medicaid is a jointly-funded program between the federal government and states. The Federal Medical Assistant Percentage (FMAP) determines the share of each state’s Medicaid expenditures that are funded by the federal government. FMAP

is determined based on the states average income to the national personal income average. FMAP varies based on state income and reevaluated every 3 years. States have the flexibility to determine where non-federal funding for Medicaid comes from, but generally it comes from state appropriations. As it relates to ACA expansion, states who opted to expand were able to capture 100% of the costs of the new population via FMAP funds. By 2020, the FMAP for this population will decrease to 90%.

As it relates to prescription drugs, all Medicaid programs offer prescription drug benefits despite that not being a required benefit in statute. An important component of prescription drugs in Medicaid is the mandated Medicaid Drug Rebate Program (MDRP). The MDRP requires manufacturers to pay statutorily mandated rebates on all prescription drugs. The ACA increased the mandated rebate for drugs to 23.1% as compared to 13% for generic drugs (Centers for Medicare and Medicaid Services 2017d). Over 600 manufacturers currently participate in the MDRP and participation in the MDRP also extends to two other federal pricing programs: the 340B drug pricing program and the Federal Supply Schedule operated by the Department of Veterans Affairs. The MDRP also requires drug manufacturers to provide the Medicaid program with the lowest price they offer across the marketplace, otherwise known as “Medicaid Best Price.” Prior to this being codified in law, health maintenance organizations (HMOs), hospital systems and group purchasing organizations (GPOs) were able to negotiate deep discounts in the 1990s, often deeper than Medicaid programs. Therefore, Congress established the Best Price mandate in the Omnibus Reconciliation Act of 1990 (OBRA ’90).

In addition to the mandated rebate, brand drugs also face a Consumer Price Index-Urban (CPI-U) adjustment. For products on the market for several years that have taken price increases regularly, can essentially be at a very high rebate level once the 23.1% plus the CPI-U adjustment is taken. Thus, Medicaid programs have a variety of brand drugs they receive at a very high rebate. The other dynamic relevant to prescription drugs in Medicaid, though not mandated in statute, is the inclusion of supplemental rebates. States can negotiate with drug manufacturers to generate additional rebates and further reduce expenditures. Supplemental rebates are reported back to the Federal government on the basis of the FMAP amounts. States can enter into supplemental rebates alone or in conjunction with other states (Department of Health and Human Services 2014).

Medicaid programs establish Preferred Drug Lists (PDLs) which are their version of formularies that spell out products that are covered and utilization management tools applied. Since Medicaid eligibility is based on an individual being low-income, there are very low copayments assigned to drugs—and many Medicaid programs do not even collect these due to the burden on patients. As such, the PDL is the primary way Medicaid programs control access and spending.

The last and relevant component of Medicaid is the influx of Managed Care Organizations (MCOs) infiltrating what was traditionally a state-run program. Traditional MCOs typically have more sophisticated tools and approaches to manage spending, so increasingly MCOs have entered into contractual relationships with state programs to control escalating expenditures. Currently, 39 states controls

with MCOs and in 2014, nearly 55 M individuals had some aspect of their care touched by an MCO (Centers for Medicare and Medicaid Services 2018c). With the continued proliferation of MCOs in Medicaid, referred to as Medicaid Managed Care, the ACA included provisions to allow MCOs to keep supplemental rebates they negotiate with drug manufacturers.

As it pertains directly to biosimilar reimbursement, these considerations are critical. In March 2015 and reaffirmed in 2016, CMS notified manufacturers that for the purposes of the MDRP, biosimilars will be considered single-source, brand drugs, and therefore, responsible for the 23.1% mandated rebate (Centers for Medicare and Medicaid Services 2016) as well as any supplemental rebates. This of course is direct opposition to CMS' treatment of biosimilars within Medicare Part B, where biosimilars are considered "multi-source" and in Part D, where biosimilars are "non-applicable" drugs.

Outside of CMS' interpretation of biosimilars within Medicaid, is the reality of how the reference products are currently rebated and covered under Medicaid. Since many products that are likely to have a biosimilar competitor have taken price increases and been on the market for several years, it is entirely possible that products will be highly rebated and/or at 100% rebate. Therefore, biosimilar penetration could be very challenging in state Medicaid programs as Medicaid programs may not want to give up that rebate for the reference product, and/or it does not make financial sense for them to do so.

Coverage and payment for physician-administered drugs (or Part B drugs) in Medicaid varies significantly from program to program. But effective in 2006, states were required to collect specific information on physician-administered drugs to collect rebates. While coverage and reimbursement methodology varies across programs, since physician-administered products are required to pay the same rebates as self-administered drugs, the dynamics of products that are at a high level of rebate with biosimilar competition will face the same complexities in getting traction long-term. Currently, many Medicaid programs also utilize the ASP methodology established by Medicaid, though it is up to the individual program to determine the additional percentage they pay above the 6%, if any. As discussed earlier, state Medicaid programs are also beholden to the CMS coding decision for biosimilars, thus there will be limited ability for a state Medicaid program to do any product preferencing since all products are blended into one code currently.

In summary, the dynamics of the MDRP may mean biosimilar utilization in Medicaid programs is low compared to other government programs; quite ironic given that Medicaid is for low-income individuals who could benefit from a lower cost option.

Commercial Reimbursement and Biosimilars

Given the continued rising spend on specialty drugs, commercial payers have long supported the proliferation of a robust biosimilar market in the US. And while the potential cost savings associated with biosimilars is undoubtedly attractive to

Table 3.3 PBM biosimilar decisions

PBM	Activity	Years
CVS Health Express scripts	Excluded Neupogen® in favor of Zarxio® Excluded Lantus® in favor of Basaglar®[a] Silent thus far on Inflectra® and Renflexis™	2017 2018
United healthcare (Optum)	Excluded Neupogen® in favor of Zarxio® Excluded Lantus® in favor of Basaglar® Silent thus far on Inflectra® and Renflexis™	2017 No release of 2018 formulary information at time of publication

[a]While Basaglar is not technically a biosimilar right now, based on a provision within the ACA, basal insulins will transition to the biosimilar pathway come March 2020

payers, like most things reimbursement-based, complexities around the clinical and financial considerations loom large.

As discussed in Chap. 5, the purpose of the biosimilar pathway is not to independently prove safety and efficacy, but rather to prove biosimilars are highly-similar to the reference product, with no clinically meaningful differences. Stakeholder lack of familiarity relating to biosimilars and potential concerns will be top of mind for commercial payers as they consider aggressive moves towards biosimilars. Table 3.3 below illustrates the top pharmacy benefit managers (PBMs) moves on biosimilars for 2018 formularies:

To date, payer activity has been slow; however, there remains a limited number of biosimilars launched, so it is not unexpected. Recent payer research illustrates that commercial payers are relatively open to incorporating biosimilars into their formularies. Recent research indicates payers expect biosimilars to be priced at a 20–30% discount (Covance Inc 2014).

The financial considerations for a commercial payer around biosimilars will vary significantly by the benefit where the product is covered, therapeutic area and existing marketplace considerations. Some areas, like immunology, have many brand options and thus are inherently competitive from a payer perspective in terms of rebates and/or discounts being offered. This is particularly prevalent for pharmacy benefit products. The medical benefit of commercial payers still remains less competitive in terms of rebates/discounts and is more driven by physician choice versus payer influence. Commercial payers will most certainly consider the economics of rebates lost from the innovator compared to rebates gained from a biosimilar manufacturer as they evaluate biosimilar inclusion in formulary decisions.

Physician acceptance of biosimilars is likely to also influence commercial payer considerations around biosimilars. For example, although payers have tools in their arsenal to push utilization of biosimilars, physician acceptance will still be critical. Commercial payers rely on physicians to drive their formulary choices and buy into

those choices, as not to create disruption at the provider or patient level. This concept will remain especially critical when it comes to a new market like biosimilars.

Patient acceptance is also critical for commercial payers to consider. While payers can look at a variety of utilization management tools; patient cost-share changes may be the most powerful to drive biosimilar utilization. However, this will depend significantly on the therapeutic area and what the patient responsibility has been for the brand product. Some areas like immunology have had a competitive landscape with many brand products competing for market share. As such, copay cards which insulate patients from payer-mandated coinsurance or higher copayment amounts, have been prevalent. Thus, a payer may have limited means to drive biosimilar utilization due to changing of copay or coinsurance structure of a brand product. Brand companies are unlikely to forgo copay cards in light of biosimilar competition.

In summary, commercial payers will play a critical role in the uptake of biosimilars moving forward in the US. Early activity shows payer interest remains high particularly among the big three PBMs in the US. However, it is not as simple as lower price equals payer utilization, the financial picture is complex and payers do have to consider the level of disruption they are willing to put forth on physicians and patients. Lastly, as discussed earlier, commercial payers are beholden to the CMS coding decision for biosimilars, thus there will be limited ability for a commercial payer program to do any product preferencing since all products are blended into one code currently.

Spotlight on the European Experience

By: Jennifer Cook, Global Market Access lead, Biosimilars, Boehringer Ingelheim GmbH Jennifer.cook@boehringer-ingelheim.com

Relative to Europe, who has had biosimilars since 2006 and over 20 launched; the US market is very young in having biosimilars experience. It begs the question, are there lessons we can learn from the European experience around reimbursement policy?

Biosimilar Experiences in Europe So Far: Lessons Learned

Since 2006, European countries have approved 36 biosimilars, accumulated over 700 million patient days on therapy and realized significant savings (European Medicines Agency 2017). Biosimilars in the US is still in its relative infancy; how they will affect the rising spend of biologics in the US remains to be seen.

While the European and US regulatory environment share many similarities related to biosimilars, including the approach for demonstrating similarity and the extrapolation of indications, the marked differences in how drugs are purchased, priced, and distributed makes it difficult to fully extrapolate Europe's experience to the US. However, there are valuable experiences from the intended and unintended impact of various pricing and market access (P&MA) policies that have been implemented across Europe to date.

Discounting biosimilars is one way to affect the adoption rates, however evidence shows that that tailor made payer policies that address the specific market characteristics and needs of the payer, provider and pharmaceutical companies are essential to drive uptake and foster a more sustainable market. A recent study by Rémuzat et al. (2017) looking at the key drivers for market penetration in Europe has shown that incentive policies to enhance uptake remain an important driver of biosimilar penetration, while biosimilar price discounts have no impact on share.

An overview of the policies that have been introduced across selection European countries can be seen in Table 3.4.

Table 3.4 Overview of key physician and pharmacists incentive policies in select EU countries

Type of Policy	France	Germany	UK	Italy	Spain
Physicians					
Pharmaceutical prescription budgets	x	√	√	√	x
Prescription quotas	x	√	x	√	x
Monitoring of prescription patterns	x	√	√	√	√[c]
Financial incentives or penalties	x	√	√	√	x
Prescription conditions/guidelines	√	√	√	√	x
Switching	√[a]	√[b]	√	√	√[b]
INN prescribing	x	x	x	x	x
Education & Information	√Not specifically targeting physicians	√	√	√Not specifically targeting physicians	√Not specifically targeting physicians
Pharmacists					
Substitution	√	√Only for specific groups of biosimilars	x	x	x
Financial incentives or penalties	√	√	√	√	√
Education and Information	√Not specifically targeting pharmacists	√Not specifically targeting pharmacists	√Not specifically targeting pharmacists	√Not specifically targeting pharmacists	√Not specifically targeting pharmacists

√ Policy potentially impacting biosimilar uptake | √ Policy not expected to impact biosimilar uptake | X Policy not applicable to biosimilars

European Policies: Those Driving and Not Driving a Competitive Biosimilar Market

Germany: Driving

We have seen varying levels of uptake of biosimilars by country and by product across Europe. Let's take a look at Germany; a country where biosimilar adoption has been relatively strong, Filgrastim, a physician-administered product for mostly acute use, and thus largely a payer-driven therapy selection, the uptake of biosimilars has been greater than 50%, while for somatropin, a chronic treatment where therapy selection is largely driven by patients and providers, uptake has been quite limited. Across Europe, Germany has seen the greatest incentives to drive biosimilar use. The statutory health insurers (Gesetzliche Krankenversicherung) officially called "Sickness funds" (Krankenkassen, KKs) and physician's associations (Kassenarztliche Vereinigung, KVs) have been pro-active in reaching out to physicians in order to encourage them to prescribe biosimilars and meet their prescribing targets as well as their regional biosimilar quotas, which can be as high as 50% relative to originator biologics to new patients. This has been supported by education campaigns, highlighting the savings potential of biosimilars (IMS Health, 2016a, b). However, the lack of monitoring has meant that up until now, many of the German KV regions have not met their biosimilar quotas, leaving room for improvement.

From a procurement perspective, the Sickness funds (Krankenkassen) in Germany have 'open house rebate contracts' with pharmaceuticals. The Sickness funds negotiate rebates to reduce net price in return for agreed market shares with the manufacturer. This challenges the possible price advantage of biosimilars vs originator on the net level. In addition, theFederal Joint Committee (GBA)) along with the National Association of Statutory Health Insurance Funds (GKV-Spitzenverband) has also instituted reference pricing for biosimilars further reducing the price advantage of biosimilars vs originator on the list price level.

UK: Driving

Adoption of biosimilars has been amongst the highest in the United Kingdom (UK). For example Benepali, a biosimilar of Etanercept (launched in March 2016), achieved 52% biosimilar volume market share just after 14 months and the uptake was significantly faster than Infliximab biosimilar (39% biosimilar volume share after 14 months).

Biosimilars in the UK are procured through regional tenders, and adoption has been greater in regions where financial incentives in place e.g. gain share agreements that rewards cost effective prescribing. Any savings which are generated through the implementation of the biosimilars are split between Clinical Commissioning Groups (CCGs) who are responsible for funding and the hospital responsible for prescribing.

However, this is not yet a wide spread phenomenon in the UK due to administrative burden and the complexity of splitting the savings generated between CCGs and hopsitals.

In 2015, NICE (National Institute of Health and Care Excellence) imposed guidelines recommending the use of more cost-effective drugs, and as part of their Health Technologies Adoption programme, NICE developed a step by step guide for clinical and non-clinical staff on how to facilitate use of biosimilars to Infliximab (Inflectra® and Remsima®), including practical advice on topics like patient identification, switching, monitoring and project management (NICE 2015).

Over the last few years, with the increase number of biosimilar players coming onto the market, we have seen increasing levels of discounts required on the net price level, reducing the attractiveness of the market for pharmaceutical companies.

France and Belgium: Not Driving

Biosimilar adoption in France and Belgium has been the lowest across Europe. Price is currently the main driver for biosimilar access. Both countries have a national management process i.e. fixed biosimilar price reduction. Furthermore, in the retail setting; France has introduced mandatory list price discounts which are not balanced by P&MA policies to drive the use of more cost effective treatment options impeding the use of biosimilars (Medicines for Europe 2016). In the hospital setting, they have introduced the gainsharing (T2A drugs) to facilitate broader use of biosimilars. However, to date on the prescription side there has been no payer guidance, prescribing incentives or quotas to incentivize physicians to prescribe biosimilars. Physicians typically base their prescription decision on the hospital formulary.

Italy and Spain: Not Driving

Despite the initial cultural resistance, Italy and Spain are now catching up. In Italy, regional and local quotas/usage guidelines are already in place for existing biosimilars (filgrastim, somatropin, epoetin) in Tuscany, Veneto and Campania. However, as in Germany, the quotas are not legally binding, and so far, real-life prescribing is not fully reflecting the regional quotas that have been set.

Biosimilars for somatropin, epoetin, filgrastim, and infliximab are currently purchased in regional or local/hospital tenders in Italy and Spain. However, the single-winner tenders, as well as the mandatory price discounts on the list price, limits the flexibility for pharmaceuticals to negotiate the price (9). This creates an unfavourable procurement environment reducing the commercial attractiveness for manufacturers who are looking to invest longer term.

Norway: Driving Short-Term but Long-Term Remains to Be Seen

Discount levels in Norway is the highest we see in Europe. Earlier this year, Orion won the Norwegian market for infliximab with a 72% price reduction for the biosimilar Remsima®. Other examples can be seen with epoetin and filgrastim that are used for in-hospital treatment (paid by the regional health authority) and by patients at home (paid for by national insurance). Tender prices of biosimilars of these two products are discounted up to 89%, with a high volume of sales to hospitals (Mack 2015). Despite these high levels of discounts, sales of epoetin and filgrastim are still not comparable to sales of biological drugs that do not have biosimilar competition, such as pegfilgrastim and darbepoetin, which dominate the market suggesting there are other factors at play.

Norway has a national healthcare system which is centralized and highly integrated. The Norwegian Drug Procurement Cooperation (LIS) is responsible for procurement and delivery agreements for pharmaceutical manufacturers in cooperation with state-owned hospitals, thereby reducing costs through a national tendering process. The Norwegian authorities pay for the bulk of drugs through national insurance and hospitals (regional health authority budgets).

Infliximab has had a more rapid penetration into the market than other biosimilars available in Norway. Several factors have contributed to this. The national annual tender processes for tumour necrosis factor (TNF) biosimilar drugs (infliximab) has been running at the LIS since 2007, therefore the process is well established. A key factor was that the specialists, clinicians and hospital management all supported the recommendations from a clinical and economic perspective. Furthermore, its low price puts Remsima® at the top of the ranking for all its in-hospital treatment indications. The drug costs for in-hospital Remsima® are much lower than are the drug costs for home treatment with injectable products, which resulted in a larger proportion of treatments being carried out in hospitals where the switching of patients to biosimilars is also widely accepted and common practice. Therefore, is a great example demonstrating factors including healthcare systems specificities in terms of pricing, reimbursement, procurement in addition to payer policies that meet the needs of the physicians and patients are all necessary components to promote uptake.

Interchangeability

The EU regulatory path has no separate development for interchangeables and leaves biosimilar substitution up to the individual member states. As such, each European country adopts biosimilars differently as we have discussed here, most without allowing automatic substitution. The exception is now France, which recently passed a law allowing substitution by pharmacists, but only for treatment-naïve patients. Germany is also expecting the introduction of automatic substitution in the future. Currently, there is convergence across EU countries that biologic

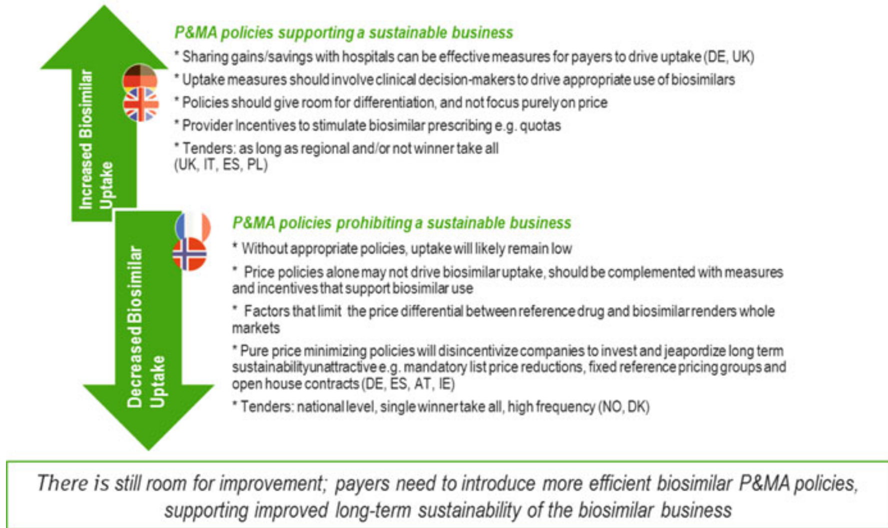


Fig. 3.6 Summary of policies supporting or prohibiting a sustainable, competitive biosimilar business

medicines should not be substituted at the pharmacy level without the involvement of the clinical decision maker. The lack of interchangeability in many EU countries has been a natural barrier to biosimilar adoption, and as a result, biosimilars haven't yet eroded the EU market share of several reference biologics as fast as many experts had expected.

In summary, the experiences so far with Biosimilars in Europe illustrate the heterogeneity between biosimilar products, therapy areas, and countries (both across and within). There is not just one formula that will work to achieve the projected savings potential, but learnings can be taken from all areas as evidenced below (Fig. 3.6).

Reimbursement Landscape Moving Forward

The key question behind all of the considerations laid out in this chapter is—in a time of significant change in the US healthcare system, what does this mean for biosimilar utilization long-term?

The US is quickly trying to move away from the fee-for-service (FFS) model to a more value-based approach to delivering and paying for healthcare. This shift is taking many forms across all payers and is primarily driven by concerns around rising costs of care, FFS incentives driving overuse and continued concern that growing expenditures do not reflect quality of care.

The federal government has actually been an early mover in driving the shift. In January 2015, then HHS Secretary Sylvia Burwell announced very ambitious targets for value-based payment:

By 2016: 85% of Medicare FFS payments tied to quality or value	By 2016: 30% of Medicare payments will be tied to quality or value through Alternative Payment Models (APMs)

While the goals were initially suggested to be very ambitious for a government agency; HHS succeeded by hitting these goals, early, in March 2016; 9 months ahead of schedule. A division within CMS, the Center for Medicare and Medicaid Innovation (CMMI) was established via ACA to develop, test and implement APMs to drive savings and improve quality. To-date, CMMI has over 80 programs being tested and many were instrumental in HHS hitting its' goal to move towards quality and value.

In addition to federal activity, commercial payers are also doing a lot to drive the shift away from FFS. As it pertains to biosimilars, in theory, biosimilars fit very well into the idea of improving value and reducing cost. After all, the value equation in healthcare is fairly straight forward, value = quality divided by cost. The higher the quality the more improved the value. Biosimilars represent the opportunity for sustained quality at a lower cost.

Additionally, as APMs further take shape, biosimilars fit very well into more episodic-based models. Biosimilars aim to provide the same safety and efficacy, at a lower cost, thus they should be a natural fit within evolving payment models. In addition to APMs, the shift away from FFS has also ushered in an influx of value frameworks – or entities evaluating the value of medicines in a more complex and methodological way. The Institute for Clinical Effectiveness and Research (ICER), NCCN, ASCO and Memorial Sloan Kettering Cancer Centers have all created tools to further define the "value" of products. Again, in a biosimilar market where biosimilars come in at a discount to the reference product, the value of biosimilars should be high. Thus, market forces are driving towards reimbursement for high-value products, and it would seem biosimilars fits into this movement quite well.

In final summary, the reimbursement landscape for biosimilars is complex and in some areas requires changes to ensure a long-term robust and sustainable market. However, biosimilars represent an opportunity to help curb drug costs. As such, an appropriate reimbursement framework that encompasses coverage, coding and payment is mission critical for the US to see a sustainable, robust biosimilar market.

References

ADVI Health. July 2017.

Avalere Health. Medicare beneficiaries will pay more for biosimilars than for their biologic reference products in Part D. 2018. http://avalere.com/expertise/life-sciences/insights/medicare-beneficiaries-will-pay-more-for-biosimilars-than-for-their-biologi.

Biosimilar Development. Biosimilars forum applauds senator roberts, reps. barton, eshoo for leadership on medicare policy on biosimilar coding. 2017a. https://www.biosimilardevelopment.com/doc/biosimilars-forum-senator-roberts-reps-barton-eshoo-medicare-biosimilar-coding-0001.

Biosimilar Development. CMS' coverage gap is a significant challenge for biosimilars. 2017b. https://www.biosimilardevelopment.com/doc/cms-coverage-gap-is-a-significant-challenge-for-biosimilars-0001.

Centers for Medicare and Medicaid Services. Medicaid biosimilars website. 2016. https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Prescription-Drugs/Downloads/Rx-Releases/MFR-Releases/mfr-rel-103.pdf.

Centers for Medicare and Medicaid Services. Part B average sales price files. 2017a. https://www.cms.gov/apps/ama/license.asp?file=/Medicare/Medicare-Fee-for-Service-Part-B-Drugs/McrPartBDrugAvgSalesPrice/Downloads/2017-July-ASP-Pricing-File.zip.

Centers for Medicare and Medicaid Services. CY2018 medicare physician fee schedule final rule. Published November 2017b. https://www.gpo.gov/fdsys/pkg/FR-2017-11-15/pdf/2017-23953.pdf.

Centers for Medicare and Medicaid Services. CMS proposes policy changes and updates for medicare advantage and the prescription drug benefit program for contract year 2019 (CMS-4182-P). November 2017c. https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2017-Fact-Sheet-items/2017-11-16.html.

Centers for Medicare and Medicaid Services. Medicaid drug rebate program. 2017d. https://www.medicaid.gov/medicaid/prescription-drugs/medicaid-drug-rebate-program/index.html.

Centers for Medicare and Medicaid Services. Part B payment website. 2018a. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Part-B-Drugs/McrPartBDrugAvgSalesPrice/Part-B-Biosimilar-Biological-Product-Payment.html.

Centers for Medicare and Medicaid Services. Medicaid enrollment website. 2018b. https://www.medicaid.gov/medicaid/program-information/medicaid-and-chip-enrollment-data/report-highlights/index.html.

Centers for Medicare and Medicaid Services. Medicaid managed care website. 2018c. https://www.medicaid.gov/medicaid/managed-care/index.html.

Covance Inc. Biosimilars: the commercial challenge. 2014. https://www.covance.com/content/dam/covance/assetLibrary/whitepapers/Biosimilars-WPCMA001.pdf.

Department of Health and Human Services. States' collection of offset and supplemental rebates – OIG report. December 2014 https://oig.hhs.gov/oei/reports/oei-03-12-00520.pdf.

European Medicines Agency, Biosimilars in the EU: information guide for healthcare professionals. 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2017/05/WC500226648.pdf.

FTC. Federal trade commission report. 2009. https://www.ftc.gov/sites/default/files/documents/reports/emerging-health-care-issues-follow-biologic-drug-competition-federal-trade-commission-report/p083901biologicsreport.pdf.

Health Affairs–Health Policy Briefs. Specialty pharmaceuticals. November 2013. http://www.healthaffairs.org/healthpolicybriefs/brief.php?brief_id=103.

IMS Health. Delivering on the potential of biosimilar medicine: the role of functioning competitive markets. Parsippany, NJ: IMS Health; 2016a.

IMS Health. The impact of biosimilar competition. Parsippany, NJ: IMS Health; 2016b.

Janssen Pharmaceuticals. Remicade billing guide. 2016. https://www.janssencarepath.com/sites/www.janssencarepath.com/files/remicade-billing-guide.pdf.

Mack A. Norway, biosimilars in different funding systems. What works? GaBi. 2015;4(2):90–2.

Medicines for Europe. Payers' P&MA policies supporting a sustainable biosimilar market. Final report. July 2016.

NICE. Introducing biosimilar versions of Infliximab: Inflectra and Remsima. 2015. http://www.biosimilars-nederland.nl/wp-content/uploads/2015/08/NICE-Adoption-Resource-Biosimilar-Infliximab-2015_07_31.pdf.

PwC Health Research Institute. Medical cost trend: behind the numbers. 2016.https://www.pwc.com/us/en/health-industries/behind-the-numbers/assets/pwc-hri-medical-cost-trend-chart-pack-2016.pdf.

Rémuzat C, Dorey J, Cristeau O, Ionescu D, Radière G, Toumi M. Key drivers for market penetration of biosimilars in Europe. J Mark Access Health Policy. 2017;5(1):1272308.

Chapter 4
Litigation-Related Issues Under the Biologics Price Competition and Innovation Act

Brian D. Coggio, Ron Vogel, and Tasha Francis

Abstract The chapter summarizes the pertinent provisions of the Biologics Price Competition and Innovation Act ("the Act") that relate to patent issues, including the so-called patent dance, the two waves of litigation, and due to its significance to biosimilar practice, *inter partes* review.

The chapter first examines the statutory framework of the Act that governs the contemplated exchange of confidential-information between the sponsor and biosimilar applicant resulting in the list of patents to be litigated—the infamous patent dance. It also discusses the two distinct waves of litigation that the Act envisioned, and the 180-day notice of commercialization that triggers the second wave. Also provided are suggestions for both sponsors and applicants on preparing for litigation under the Act. The chapter then reviews the leading cases that have addressed the workings of the Act including the information exchange process, the patent dance, discovery, and remedies. Significantly, the Supreme Court ruled that information exchange and patent dance provisions are optional, and the 180-day notice of commercialization can be given at any time after FDA filing. Certain ramifications of the decision are explored in the cases discussed. The chapter concludes with a discussion of Hatch-Waxman safe harbor, which is applicable to biologics/biosimilars and unchanged under the Act, and *inter partes* review, which will continue to play a significant role in biosimilars' attempts to avoid district court litigation for their proposed products.

Brian Coggio is of Counsel and Ron Vogel and Tasha Francis are associates with Fish & Richardson P.C. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the firm or its clients.

B. D. Coggio (✉) · R. Vogel
Fish & Richardson P.C., New York, NY, USA
e-mail: coggio@fr.com; rvogel@fr.com

T. Francis
Fish & Richardson P.C., Minneapolis, MN, USA
e-mail: tfrancis@fr.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_4

Keywords Hatch-Waxman Act · Infringement · Patent dance · Patents · Litigation · *Inter Partes Review* · BPCIA

Introduction

Over 30 years ago, in an attempt to balance disparate interests in competition and innovation with small molecule pharmaceuticals, the Drug Price Competition and Patent Term Restoration Act was signed into law. Commonly known as the Hatch-Waxman Act, this legislation established an Abbreviated New Drug Application (commonly, an "ANDA") enabling generic drug manufacturers to gain regulatory approval of their products without needing to prove safety and efficacy, merely bioequivalence. The filing of an ANDA allowed a patentee under 35 U.S.C. § 271(e)(2)(A) to institute a patent infringement action asserting patents listed in the so-called Orange Book, which identifies patents (other than process patents) relevant to the approved drug. The Hatch-Waxman Act, however, did not encompass biologic drugs, such as monoclonal antibodies, which are of a molecular size and complexity far greater than small molecule pharmaceuticals. Biologic drugs, such as Humira®, Enbrel®, Rituximab®, Avastin® and Herceptin®, now dominate the best selling drugs globally.[1]

The Biologics Price Competition and Innovation Act ("BPCIA") was enacted in 2009 to provide an abbreviated pathway for regulatory approval of biologics ("biosimilars") by filing an abbreviated Biologic License Application ("aBLA"). Similar to the Hatch-Waxman procedure, the counterweight to this abbreviated (but much more complex) pathway, is the BPCIA's amendment to the Patent Act that created an artificial act of infringement allowing reference product sponsors ("Sponsors") to institute infringement suits under 35 U.S.C. § 271(e)(2)(C) against biosimilar applicants ("Biosimilars") based on the filing of an aBLA. To facilitate these suits, the BPCIA provides, among others, for a complex set of information exchanges, commonly dubbed the "patent dance."[2]

The basics of the BPCIA and related litigation issues, including *inter-partes* reviews ("IPRs"), are discussed in this chapter.

The BPCIA and What It Amends

Listed below is a compilation of the statutory provisions impacted by the BPCIA.

- Public Health Service Act, 42 U.S.C. § 262
  - § 262(k)—Biosimilar application requirements.
  - § 262(l)—Patent exchange ("patent dance") procedures.

[1] *See* https://www.thebalance.com/top-biologic-drugs-2663233

[2] For a comprehensive discussion of the background of the BPCIA, see Carver et al. (2010).

- Patent Statute, 35 U.S.C. § 271(e)(2) and § (e)(1)
  - § 271(e)(2)(C)(i) and § (C)(ii)—New artificial acts of infringement.
    - Similar to Hatch-Waxman Act (§ 271(e)(2)(A)), but includes process patents.
  - § 271(e)(2)(C)(i) Governs instances where the Biosimilar provides its biosimilar application to the Sponsor and a patent exchange list is generated.
  - § 271(e)(2)(C)(ii) Governs instances where the Biosimilar fails to provide its biosimilar application to the Sponsor or does not participate in the patent dance.
  - § 271(e)(1)—Safe Harbor—stays the same.

As of March 23, 2020, all applications for proteins seeking FDA approval under § 505(b)(2), formerly known as "paper NDAs," will no longer be considered "drugs," but will be subject to the BLA requirements because the underlying reference drug product will then be considered a "biologic." At that time, a number of proteins that were approved as drugs will be transitioned to biologics, e.g., insulin and human growth hormone.

Types of Biosimilar Products/Naming Biosimilar Products

The BPCIA created an abbreviated licensure pathway for biological products that are demonstrated to be "biosimilar" to or "interchangeable" with an FDA-licensed biological product. The FDA requires licensed biosimilar or interchangeable products to meet the agency's standard of safety and efficacy. As part of the BPCIA, the FDA created the so-called "Purple Book" that lists biological products including biosimilar and interchangeable products licensed by the FDA, and, importantly, the Sponsors' exclusivity. The book includes the date a biological product was licensed and whether any biosimilars or interchangeable biosimilars have been approved with respect to a licensed biologic.

There are two types of biosimilar products: (1) biosimilars and (2) interchangeable biosimilars.

A "biosimilar" product is "highly similar" to the Sponsor's product notwithstanding minor differences in clinically inactive components, and there are no meaningful differences between the Biosimilar's and Sponsor's products in terms of safety and potency. 42 U.S.C. § 262(i)(2). The FDA, beginning in April 2015, released various guidance documents (including ones directed to scientific and quality control considerations) to assist Biosimilars in the approval process.

An "interchangeable biosimilar" product must be biosimilar and must also be expected to produce the same clinical result as the Sponsor's product in any given patient. Moreover, if administered more than once, the risk in terms of safety or diminished efficiency of switching between the Biosimilar's and Sponsor's products

is no greater than the risk of using the latter without the switch. 42 U.S.C. § 262(i)(3). In January 2017, the FDA released a guidance on interchangeability.[3]

The FDA has also issued a draft guidance on naming biosimilar products.[4] The proper name for all such products will include:

Core name: International Nonproprietary Name (INN) assigned by the World Health Organization (WHO).

Suffix: Four unique letters to demonstrate that the biosimilar product has not been designated interchangeable.

BPCIA Statute

Act of Infringement

Title 35, U.S.C. §§ 271(e)(2)(C)(i) and (ii) create new artificial acts of infringement allowing a Sponsor to institute suit when a Biosimilar files an aBLA.

Section 271(e)(2)(C)(i) governs instances when the Biosimilar provides its aBLA to the Sponsor, and the parties generate a list of patents to be asserted. As discussed *infra,* there are two ways that this list can be generated. The ensuing infringement action is limited to the listed patents.

Section 271(e)(2)(C)(ii) governs instances when the Biosimilar fails to provide its aBLA or participate in the patent dance. The Sponsor can then assert all of its patents against the Biosimilar, including process patents, by filing a declaratory judgment action.

Since the information exchange and patent dance provisions are optional, as discussed further *infra*, actions under the latter section may increase.

Confidentiality Provisions

Before addressing the details of the information exchange process and the patent dance, it is noteworthy that the BPCIA strictly limits the individuals who have access to the biosimilar application and related manufacturing details. 42 U.S.C. § 262(l)(1). All individuals who receive confidential information must agree to be bound by these provisions.

Unless the parties agree, the Biosimilar's confidential information may be disclosed to only one in-house attorney who does not "formally or informally" engage

[3] "Considerations in Demonstrating Interchangeability with a Reference Product," issued by CDER and CBER (January 2017).

[4] "Nonproprietary Naming of Biological Products," issued by CDER and CBER (Aug. 2017).

in patent prosecution related to the Sponsor's product (a so-called "prosecution bar"). Accordingly, this attorney must be walled-off from any communications with individuals prosecuting patents relating to the Sponsor's product. 42 U.S.C. § 262(l)(l)(B)(ii)(ll). Questions often arise as to how far this bar extends, e.g., does it cover IPRs, reissues, etc.?

One or more outside counsel, but only attorneys, may have access to the confidential information. But they are also subject to a prosecution bar preventing them from prosecuting patents relating to the Sponsor's product. 42 U.S.C. § 262(l)(l)(B)(ii)(l). Questions can arise as to whether the bar applies to the entire outside firm or just the individuals involved in the biosimilar litigation.

If any relevant patent has been exclusively licensed to the Sponsor, a representative of the licensor may also have access to the confidential information. 42 U.S.C. § 262(l)(1)(B)(III). A prosecution bar would also apply to this individual.

A Sponsor's proposed experts must be cleared by the Biosimilar before they are allowed access to Biosimilar's confidential information. Thus, a Sponsor may wish to select counsel with the appropriate technical background as it will likely take time to retain and clear experts. Indeed, a Sponsor anticipating an aBLA filing would be prudent to identify and retain potential experts in advance.

The So-Called "Patent Dance"

The "Patent Dance" refers to the complex process by which Sponsors and Biosimilars exchange information resulting in an initial list of patents to be litigated. The process, however, is optional, and the decision to institute the dance is controlled by the Biosimilar.

Biosimilar's Disclosure of Confidential Information

Within 20 days of notification that its application has been "accepted for review," the Biosimilar "shall" provide the Sponsor with confidential access to its aBLA and "such other information" that describes its manufacturing process. 42 U.S.C. § 262(1)(2) (A)-(B). As discussed *infra*, the term "shall" is not mandatory.

This disclosure is subject to the confidentiality provisions discussed *supra*. As discussed *infra,* disputes have arisen regarding the scope of the information the Biosimilar discloses.

If the Biosimilar refuses to disclose its confidential information, the Sponsor can file a declaratory judgment and assert all its relevant patents, even though it most likely lacks critical information pertaining to the biosimilar product and, in particular, the process for its manufacture.

Significantly, the Biosimilar is not required to continue its participation in the patent dance even though it originally chose to participate. In view of the Supreme Court's recent decision in *Amgen*, discussed *infra,* it would appear that a court cannot order the Biosimilar to enter the dance or, if it has begun, to complete it.

The Sponsor's Initial Patent List

Within 60 days of receiving the Biosimilar's confidential information, the Sponsor must list all of its patents, including those exclusively licensed-in, that it could reasonably assert against the Biosimilar's product or process. It must also identify any patents available for license ("Initial Sponsor List"). 42 U.S.C. § 262 (l)(3)(A).

It is unclear whether the Sponsor can later assert any patents (excluding later-issued or licensed-in patents) not included on its initial list. This concept is referred to as "list it–or lose it."[5] Title 35, U.S.C. § 271 (e)(6)(C) seems to indicate that further litigation is foreclosed on the unlisted patents. Until this issue is finally resolved, the Sponsor should list any patent that it could possibly assert against the Biosimilar's product or process.[6] Moreover, even if "list it–or lose it" does not apply and a further suit is possible, the Sponsor may have difficulty obtaining a preliminary injunction preventing marketing of the Biosimilar's product based on any patent not on the Initial Sponsor List.

The Biosimilar's Response

Within 60 days of receiving the Initial Sponsor List, the Biosimilar may list any patents it believes the Sponsor could assert against its product or process ("Initial Biosimilar List"). 42 U.S.C. § 262 (l)(3)(B)(i). However, if "list it–or lose it" applies, the Biosimilar should not add any additional patents to the dispute, since litigation on them may be forever barred.

In its response, the Biosimilar must provide a claim-by-claim analysis for each patent identified in the Initial Sponsor List (and any patents the Biosimilar adds) of the factual and legal basis as to why such claim is invalid, unenforceable or will not be infringed by its product or process or that the Biosimilar will wait until the patent expires before marketing its drug. 42 U.S.C. § 262(l)(B)(ii).

Lastly, the Biosimilar must provide a response regarding each patent that the Sponsor indicated is available for license. 42 U.S.C. § 262 (l)(B)(iii).

One might conclude that the Biosimilar's response would be binding in future litigation. Not so. At least one court held that it was "not controlling" and could be modified later as it was an optional pre-litigation letter. On appeal, the Federal Circuit held that these statements have "some probative weight," but are not binding.[7]

---

[5] *See* Coggio and Vogel (2016).

[6] *See Amgen Inc. v. Hospira, Inc.*, 866 F.3d 1355 (Fed. Cir. 2017).

[7] *Amgen Inc. v. Apotex Inc.*, 2017 U.S. Dist. LEXIS 13919 (S.D. Fla. Jan. 31, 2017) *aff'd*, 712 Fed. Appx. 985 (Fed. Cir. 2017).

The Sponsor's Reply

Within 60 days of receipt of the Initial Biosimilar List, the Sponsor must provide a claim-by-claim analysis of infringement of each patent identified in its initial list and any patents added by the Biosimilar. But it only needs to provide a "response to the statement concerning validity and unenforceability." 42 U.S.C. § 262 (l)(3)(C). The difference in wording between infringement and validity/unenforceability proofs would appear quite significant and should guide the Sponsor in preparing its response.

The Negotiated List

After the Biosimilar's receipt of the Sponsor's Reply, the parties have 15 days to negotiate a list of all patents to be litigated ("Negotiated List"). 42 U.S.C. § 262 (1)(4). The Sponsor must then institute suit on all patents on the list within 30 days. 42 U.S.C. § 262 (1)(6)(A). If suit is not filed within that period (or if filed and dismissed without prejudice, or not prosecuted in good faith), the Sponsor's remedy for infringement of these patents is limited to a reasonable royalty. 35 U.S.C. § 271(e)(6)(B).

This is the so-called "First Wave" litigation. In principal, it was to be followed by a "Second Wave" litigation triggered by the Biosimilar's 180-day Notice of Commercial Marketing, discussed *infra*. The Second Wave would cover all patents not asserted in the First Wave and also include later-issued and licensed-in patents. The Federal Circuit had held that the notice could only be given after FDA approval of the biosimilar product. In its recent *Amgen* decision, however, the Supreme Court ruled that the notice could be provided even before FDA approval. Thus, the Biosimilar fully controls when and even if there is a Second-Wave litigation. It can perhaps give notice immediately upon FDA acceptance of its aBLA and collapse the two waves of litigation into one.

The Exchanged Lists

If the parties cannot agree on the patents to be litigated, a complex exchange process ensues to identify those patents to be asserted in the First Wave litigation.

First, the Biosimilar notifies the Sponsor of the number of patents that can be litigated. At this time, the Biosimilar need not identify any particular patents, but only provide the number. 42 U.S.C. § 262(l)(5)(A).

Within 5 days of receipt of this notice, the parties are required to simultaneously exchange lists of patents that each believes should be included in this initial phase of the litigation. 42 U.S.C. § 262(l)(5)(B). But the number of patents listed by the Sponsor cannot exceed the number chosen by the Biosimilar. However, if the Biosimilar stated that zero patents should be litigated, the Sponsor can identify one.

Within 30 days of this exchange, the Sponsor must institute suit on the patents on both lists. 42 U.S.C. § 262(l)(6)(B). If it does not, potential remedies are limited.

Patent Exchange/Patent Dance Time Line

The chart below (Fig. 4.1) is helpful in understanding the dynamics of the patent dance. It demonstrates the need for both parties to plan for the BPCIA litigation process well in advance.

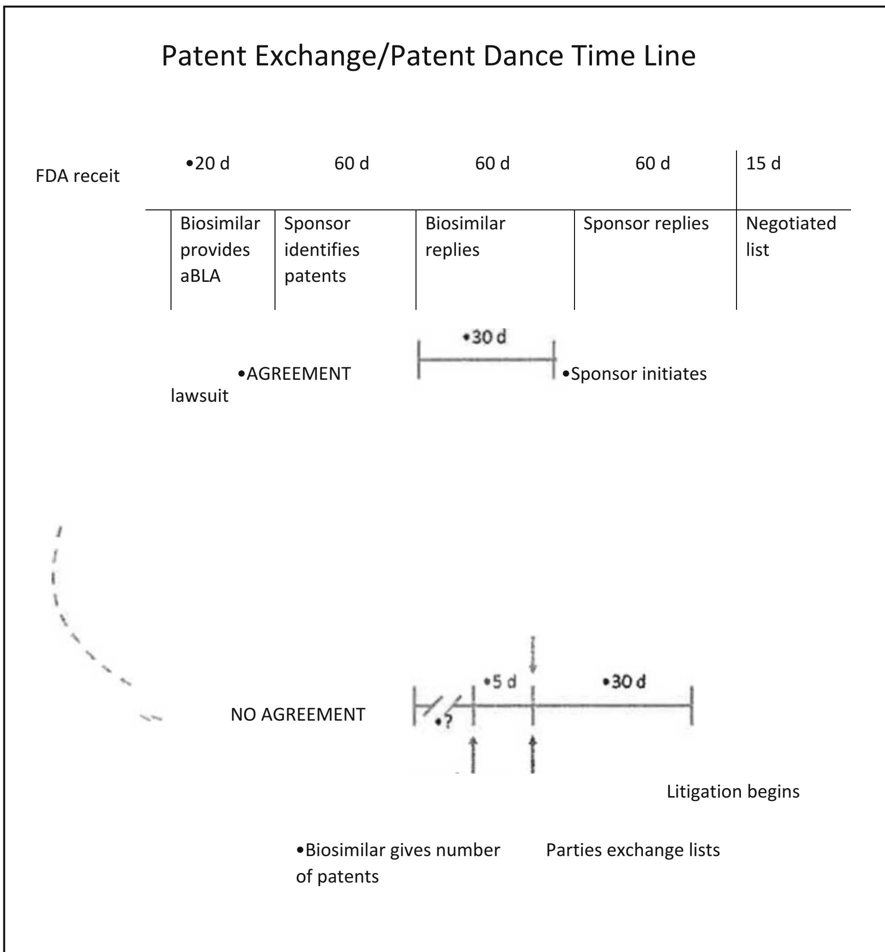


Fig. 4.1 Timeline for the patent exchange/patent dance

The Biosimilar Does Not Dance

In *Amgen*, the Supreme Court held that the information exchange and patent dance procedures are not mandatory. If the Biosimilar fails to participate, the Sponsor cannot force the Biosimilar's participation, but it can institute suit and assert all its relevant patents under 42 U.S.C. § 262(l)(9)(C) and 35 U.S.C. § 271(e)(2)(C)(ii). Although § 262(l)(9)(C) does not specifically mention process patents, 35 U.S.C. § 271(e)(2)(C)(ii) creates an "artificial act of infringement," which allows suit on all patents that could have been listed under § 262(l)(3)(A). This would include process patents. Yet this is somewhat problematic, because the Sponsor lacks information on the biosimilar product, particularly its method of manufacture. This issue is discussed *infra* in conjunction with the *Hospira* litigation.

The Biosimilar's inaction could benefit the Sponsor. First, the Sponsor can assert all its patents and is not limited by the patent dance procedure. Second, since an Initial Sponsor List is not generated, no "list it–or lose it" problem can arise. Third, since the Biosimilar has not provided any information, it would be hard pressed to oppose a motion for preliminary injunction if pertinent information on infringement becomes available later, and the Sponsor asserts additional patents. Indeed, the Supreme Court noted the possibility, but expressed no view on whether a district court should consider a Biosimilar's failure to participate in the dance as relevant to the Sponsor's motion for a preliminary injunction.

The 180-Day Notice of Commercialization

The BPCIA requires that the Biosimilar provide the Sponsor with 180-days notice before marketing of the biosimilar product begins. 42 U.S.C. § 262 (1)(8)(A). The notice is required regardless of whether the Biosimilar engaged in the patent dance. Significantly, the notice can be given before or after FDA has approved the biosimilar application.

It would appear that the BPCIA envisioned two separate waves of litigation—one on the originally listed patents, and a second wave on patents not part of the First Wave litigation. This Second Wave is triggered by the Biosimilar's 180-day notice. Under the Supreme Court ruling in *Amgen*, *infra,* the Biosimilar can provide this notice at any time, and thereby trigger the Second Wave litigation. It is, of course, possible that the Biosimilar may provide notice on the day its application is accepted for review by the FDA. In that instance, there would apparently only be one wave of litigation involving all the Sponsor's patents. The varied possibilities on the timing of the notice and its effect on litigation have not yet been fully addressed by the courts.

BPICA Comparison to the Hatch-Waxman Act

Many readers may be familiar with the provisions of the Hatch-Waxman Act, which was intended to facilitate market entry by generic. drug manufacturers. This Act, which served as a guide for the BPCIA, deals with small molecules and is much less complicated that the BPCIA. A comparisons between the two is still quite useful.

Certain key differences between the two Acts exist. First, under the Hatch-Waxman Act, the Sponsor is required to identify to the FDA all product, formulation and method of treatment patents that relate to the approved drug product. These patents are subsequently listed in the so-called “Orange Book.” Significantly, no process patents can be listed; thus, they are not litigated in typical Hatch-Waxman actions. Rather, Hatch-Waxman actions are usually limited to patents listed in the Orange Book. Under the BPCIA, however, patents are not identified to the FDA, and thus, no list of relevant patents exists. The information exchange process was seemingly designed to fill this gap. Unlike Hatch-Waxman actions, however, process patents can be asserted in BPCIA actions.

Second, when a Hatch-Waxman action is filed, the FDA is precluded from approving the related generic application for up to 30 months (“30-month stay”), unless the patent-in-suit is held invalid and/or not infringed during the related litigation or the parties settle. No automatic stay of FDA approval exists when a BPCIA action is filed.

Third, the first generic who files a complete ANDA challenging a listed patent receives 180-days of market exclusivity over all later-filed generics. No similar 180-day market exclusivity exists under the BPICA. Rather, the BPCIA’s exclusivity provisions are much more complicated. This and other differences between the two Acts are set forth below (Table 4.1).

The Hatch-Waxman Safe Harbor

The BPICA did not alter the Safe Harbor of the Hatch-Waxman Act, 35 U.S.C. § 271(e)(1) (“the Safe Harbor”). Since that provision is important to both Sponsors and Biosimilars, a brief discussion is warranted.[8]

Under the Safe Harbor, “it is not an act of infringement to make, use, offer to sell or sell in the United States or import into the United States a patented product … solely for uses reasonably related to the development and submission of information [to the FDA].” The Supreme Court has held that this provision provides a “wide berth” of protection.[9] Accordingly, it shields from infringement research and development conducted on compounds where the researcher believes

[8] *See generally*, Vogel and Coggio (2016).

[9] *Merck KGaA v. Integra Lifesciences I, Ltd.*, 545 U.S. 193 (2005).

Table 4.1 Comparison of the BPCIA and Hatch-Waxman Act

	Hatch Waxman Act	BPCIA
Patents identified	Orange Book listing of patents (no process patents), certified against by generic applicant (Para. IV certification)	No patent listing, private exchange of patent information (“patent dance”), which is now optional
Application types	ANDA or § 505(b)(2) “paper NDA”	Biosimilar license application/biosimilar interchangeable license application
FDA stay	Automatic 30-month stay of FDA approval upon filing suit	No automatic stay of FDA approval
Sponsor exclusivity	Five-year marketing exclusivity for new active moiety commencing on FDA approval	Twelve-year marketing exclusivity for new biological structures commencing on FDA approval: • But if application is filed by same Sponsor or manufacturer of the Sponsor’s product (or a licensor, predecessor-in-interest or a related party), the changed biological structure must also result in: – A change in indications, route of administration, dosing schedule, dosing form, delivery system, delivery device or strength, or – A change in safety, purity, or potency
Sponsor exclusivity	Three-year marketing exclusivity for new indication or dosage form	No additional exclusivity for same biological structure
Generic exclusivity	ANDA—First to file and to certify under Para. IV (challenging Orange Book patents) receives 180 days of market exclusivity against later-filed ANDAs • Can be forfeited under various conditions • § 505(b)(2)—no 180-day exclusivity	No exclusivity for biosimilar. First interchangeable biosimilar receives exclusivity against any subsequent interchangeable license application for any condition of use in the Sponsor’s product until the earlier of: • One year after commercial marketing by first biosimilar; • Eighteen months after court decision (appellate court, if appealed) on all patents or dismissal of action against first biosimilar; or • Forty-two months after first biosimilar approval if litigation is still pending, or 18 months after first biosimilar approval if no suit is filed (i.e., where first biosimilar fails to market)
Pediatric exclusivity	Pediatric exclusivity adds 6 months to all exclusivities	Same
Filing limitation	ANDA cannot be filed until 5 years after Sponsor’s FDA approval of new active moiety, but can be filed after 4 years if accompanied by a Para. IV certification	Biosimilar application can be filed 4 years after Sponsor’s FDA approval

that a reasonable probability exists that the compound "may" work for its intended purpose. "Basic research," however, which has never been defined, is not protected. One example of this might be high-throughput screening.

The Safe Harbor protects pre-clinical and clinical trials to develop information for the FDA. The information, however, need not be actually submitted to the FDA for this protection to apply. Moreover, conduct both before and after FDA approval may be protected, although post-approval conduct is scrutinized more closely.[10] Routine post-approval testing, even if requested by the FDA, is likely not protected. Third parties that assist a Biosimilar in preparing its FDA submission, e.g., by supplying active ingredients, are protected.[11] Apparently, stockpiling of commercial batches may not be protected by the Safe Harbor.[12]

It is unclear whether research tools, which play a key role in developing and manufacturing all pharmaceutical products, are protected by the Safe Harbor. This same uncertainty exists in many European countries under the Bolar Amendment, the counterpart to the Hatch-Waxman Safe Harbor.[13]

Preparation for BPCIA Litigation

There are several steps Sponsors and Biosimilar applicants should take to prepare for BPCIA litigation. A few approaches are discussed below.

Sponsor's Preparation

Identify Individuals Who Will Receive Confidential Information

Restrictions on the use and dissemination of confidential information from a Biosimilar are strict. Thus, it is imperative to have fully organized the Sponsor's litigation team before being served with an aBLA. This event triggers a full-scale investigation of the Biosimilar's aBLA and related manufacturing information (if provided), followed by a comparison of that information with each individual claim of the Sponsor's patent portfolio. This ideally requires lawyers with the appropriate scientific expertise. Participation of in-house attorneys is strictly limited. Moreover, individuals participating in continued prosecution of patent applications relevant to the Sponsor's product—the most knowledgeable personnel—are excluded from the team.

[10] *Momenta Pharmaceuticals Inc. v. Teva Pharmaceuticals, Inc.*, 809 F.3d 610 (Fed. Cir. 2015).

[11] *Shire LLC v. Amneal Pharmaceuticals LLC*, 802 F.3d 1301 (Fed. Cir. 2015).

[12] *Amgen Inc. v. Hospira, Inc.,* 2018 WL 4080353 (D. Del. Aug. 27, 2018) ($70 million awarded for stockpiling.).

[13] *See* Coggio (2014).

Lastly, patents that have been exclusively licensed by the Sponsor present additional complexities as a representative of the licensor is also given access to the Biosimilar's confidential information. Thus, any relevant licenses must be examined to determine if they are indeed "exclusive," which is often not an easy task. Regardless, the individual designated to receive confidential information should be chosen in advance.

Compile a List of Potentially Relevant Patents

Sponsors should expect challenges to patents relating to successful biologic products. As noted above, once a challenge is instituted, the time to identify relevant patents is brief (60 days). Thus, Sponsors should prepare by identifying all patents that relate to each of its marketed biologic products—product, formulation, method of use, and manufacturing patents. This list will be used to evaluate potential infringement of any biosimilar product or process (if the process is provided). Under recent Federal Circuit decisions, seemingly all potentially relevant patents can be listed when the Biosimilar fails to provide its aBLA and related process information. Indeed, when no information is provided, the Sponsor can file suit on all potentially relevant patents immediately. It is therefore imperative that the Sponsor know which patents and, if possible, which claims are likely candidates for biosimilar litigation.

Evaluate the Sponsor's Patent Portfolio

The patents identified for potential litigation should be evaluated as follows:

- Validity strength—Is a prior art search required?
- Inequitable conduct—Was all relevant art cited?
- Standing to sue (licensed-in patents)
- Inventorship issues
- Formalities (e.g., maintenance fees)

In addition, necessary infringement proofs should be estimated even though the Sponsor lacks details of the Biosimilar's product and process. As one example, will laboratory testing be required to show infringement? If so, by whom? Conflicts may prevent selection of the Sponsor's first choice of testing facilities. In addition, is an independent expert required to assess infringement (validity)? Conflicts may arise there as well. Retain experts promptly—ideally in advance.

Strengthen the Sponsor's Patent Portfolio

These suggestions may seem obvious, but are often overlooked.

- Pursue continuation/divisional applications where possible
- Is reissue or reexamination appropriate? Are corrections necessary?

- Make sure all appropriate PTA and PTE have been secured
- Attempt to predict potential design-arounds, and determine if they can be patented
- Are there patents that could be licensed-in to strengthen the Sponsor's position?

Establish Patent Updating Procedures

In the context of BPCIA litigation, the Sponsor must monitor its patent portfolio, including newly-issued or licensed-in patents, that could be asserted against a biosimilar product. If the process has begun, the failure of the Sponsor to inform the Biosimilar of such patents within 30 days after issuance or acquisition may limit the Sponsor's remedy for infringement.

Review Relevant License Agreements

Since exclusive licensors can play a major role in BPCIA litigation, it is prudent to evaluate any relevant licenses before the BPCIA process begins. Establish essential criteria for each license:

1. Who enforces the license?
2. Is enforcement required?
3. Is consent to sue needed?
4. Who decides which patents to enforce at each stage?
5. What remedy if licensed patent is not enforced?
6. Is licensor's cooperation required?
7. Are sublicenses permitted?
8. Who decides whether to sublicense and on what terms?
9. Who decides which patent's term to extend?
10. Who decides who receives access to Biosimilar applicant's confidential information?
11. What role does licensor play in the patent exchange process?
12. Renegotiate licenses as needed.
13. Identify a procedure for immediate notification of later-issued patents.
14. Eliminate licensor's official role, if possible. Licensor may be prosecuting other patent applications that cover the Sponsor's product. Confidentiality restrictions may be a problem.

The preceding discussion demonstrates that a Sponsor anticipating a biosimilar challenge should begin to prepare in a timely fashion. The following discussion involves steps a Biosimilar can take to prepare for BPCIA litigation.

Biosimilar's Preparation

To Dance or Not To Dance

As it affects the entire process, the Biosimilar must first decide whether to engage in the patent dance and, if so, how much to participate. In other words, should the Biosimilar dance at all? If so, should it provide its aBLA as well as manufacturing details? Or, should the Biosimilar only provide its aBLA? In either event, the Sponsor will not be able to force the Biosimilar to provide the entire aBLA or relevant process information. The Sponsor's only recourse if the Biosimilar does not fully participate in the dance is to institute a declaratory judgment asserting infringement of all potentially relevant patents. By dancing, the Biosimilar may limit the patents asserted against it and force the Sponsor to detail its infringement contentions in advance of litigation. Moreover, if the Biosimilar does not dance, it will be difficult to argue against a preliminary injunction, since it could have expedited and simplified the litigation process by dancing. Indeed, 42 U.S.C. § 262(l)(8)(C) requires the parties to cooperate to expedite the proceeding when the Sponsor seeks a preliminary injunction after receipt of the Biosimilar's 180-day notice.

The 180-Day Notice of Commercialization

The Biosimilar can provide the required 180-day notice of commercialization at any time, even before its aBLA is approved by the FDA. If the Biosimilar has danced and the litigation is focused on listed patents, the notice allows the Sponsor to assert all remaining patents in a Second-Wave litigation. An early notice may be to the Biosimilar's advantage by dealing with all potentially blocking patents at an early stage. If the Biosimilar has not danced, the Sponsor has the option of asserting all potentially relevant patents at the onset of the litigation. In this situation, the timing of the notice may not make much difference since all the Sponsor's patents will most likely have already been asserted.

Identifying Relevant Patents and Developing Defenses

Once the dance (or litigation) begins, the Biosimilar will need to counter the Sponsor's allegations of infringement by showing that the patents-at-issue are invalid, not infringed and/or unenforceable. This undertaking will take considerable time and effort, especially when the Sponsor lists (or asserts) numerous patents, and the Biosimilar decides to dance. Accordingly, the Biosimilar should attempt to identify patents likely to be asserted by the Sponsor and develop its defenses on a claim-by-claim basis, which will be needed if parties engage in the patent dance.

While noninfringement defenses may be developed without much investigation, validity defenses will often require extensive prior art searching and a detailed

review of the prosecution histories of the pertinent patents (also relevant to noninfringement defenses). For either noninfringement or invalidity analyses, it is prudent to review the prosecution histories of related foreign patents (and pending applications). This requires gathering relevant documents and where necessary, translating them into English (unless a team member reads the relevant language(s)). This will also involve extensive time and effort. Accordingly, there is no time to waste.

The above suggestions highlight the need for the Biosimilar to prepare fully before disclosing its confidential information and engaging in the patent dance, or if refusing to dance, facing immediate litigation on numerous patents.

Litigation

Based on the complexity of the provisions of the BPCIA, the statute has been characterized as a "riddle wrapped in a mystery inside an enigma." Since its inception, Sponsors and Biosimilars have struggled to understand the complex provisions of the statute. Numerous questions have been raised, including:

- Is the patent dance mandatory?
- If one begins to dance, does it have to complete the dance or can it waive provisions of the dance?
- What information is required to be exchanged under § 262(l)(2)(A)?

As litigations under the BPCIA have progressed, these and other questions have begun to surface in the courts, and some have been answered. The first question reached the Supreme Court in 2017.

Is the Patent Dance Mandatory?

Sponsors and Biosimilars were hoping the Supreme Court would clarify two aspects of the BPCIA in *Sandoz, Inc. v. Amgen, Inc.*[14]: (1) whether the confidential disclosure and patent dance provisions are mandatory and enforceable; and (2) whether the 180-day notice of commercialization can be given prior to FDA approval. In a unanimous decision, the Court answered only the second question: Notice of commercialization can be given prior to FDA approval; thus, a Biosimilar need not wait for FDA approval to provide the required 180-day notice. As to the first question, the Court stated it was not required to decide whether the patent dance (or, more specifically, § 262(l)(2)(A)) is mandatory or optional. Instead, it framed the question as "whether 262(l)(2)(A)'s requirement that the Biosimilar provide its

[14] 137 S. Ct. 1664 (2017).

application and manufacturing information to the Sponsor is itself enforceable by injunction."[15] The Court decided that an injunction is not available under federal law. Rather, a Sponsor's only remedy is to file a declaratory judgment action under § 262(l)(9)(C) asserting its relevant patents. The Court remanded to the Federal Circuit to determine if an injunction is available via state law.

The Supreme Court agreed with the Federal Circuit's conclusion, but not its reasoning, that an injunction is unavailable to force a Biosimilar to provide the Sponsor with its aBLA and related manufacturing details. The Federal Circuit interpreted 35 U.S.C. § 271(e)(2)(C)(ii) to make failure to provide an aBLA an element of the act of infringement. It then reasoned that the remedies available for this infringement are limited to the remedies specified in § 271(e)(4)—which do not include an injunction requiring the Biosimilar to provide its aBLA to the Sponsor. The Supreme Court expressly disagreed with this reasoning, noting that only the submission of an aBLA constitutes an act infringement. Failing to disclose an aBLA and related manufacturing information under § 262(l)(2)(A) is therefore not an act of infringement. Thus, agreeing with the Federal Circuit that an injunction is unavailable to force a Biosimilar to provide the Sponsor with its aBLA, the Supreme Court held that the only remedy for failure to provide this information allows is for the Sponsor to file a declaratory judgment action under § 262(l)(9)(C) and § 271(e)(2)(C)(ii) and assert its relevant patents.

At the same time, the Court remanded the case to the Federal Circuit to determine whether California law, where the action began, would treat noncompliance with § 262(l)(2)(A) as "unlawful," and, if so, whether the BPCIA pre-empts any additional remedy available under state law for the Biosimilar's failure to comply with this section. The Federal Circuit held that California state law remedies are unavailable based on field and conflict preemption.[16]

Although no injunction is available under federal (or state) law, the Supreme Court expressed confidence that Biosimilars still had significant incentives to participate in the patent dance. The Court noted that by participating in the dance, Biosimilars will have the opportunity to litigate the relevant patents before the biosimilar product is marketed. In contrast, by failing to provide its aBLA and related manufacturing information, a Biosimilar vests in the Sponsor "the control that the applicant would otherwise have exercised over the scope and timing of the patent litigation and depriving the applicant of the certainty it could have obtained by bringing a declaratory-judgment action prior to marketing its product."[17]

In reversing the Federal Circuit on the timing of the 180-day notice of commercial marketing, the Supreme Court based its decision on a straightforward textual interpretation of § 262(l)(8)(A). According to the Federal Circuit, the provision imposed two timing requirements: the Biosimilar must provide notice after the FDA approves the biosimilar product, and this must be done at least

[15] *Id.* at 1674.

[16] *Amgen Inc. v. Sandoz Inc.,* 877 F.3d 1315 (Fed. Cir. 2017); *see also* Coggio and Vogel (2017).

[17] 137 S.Ct. at 1675.

180 days before the Biosimilar markets that product. The Supreme Court rejected this interpretation, holding that § 262(l)(8)(A) contains a single timing requirement: The Biosimilar must provide notice at least 180 days prior to marketing its product. While acknowledging the numerous and weighty practical ramifications of its interpretation, the Court explicitly did not take them into account. These policy considerations "could not overcome the statute's plain language" and, in any event "are appropriately addressed to Congress, not the courts."[18]

In a one paragraph concurring opinion, Justice Breyer left open the door for the FDA to modify the Court's interpretation of the BPCIA. In his view, Congress implicitly delegated authority to the FDA to interpret the terms of that statute. Thus, if the FDA, with more experience administering this statute, determines that a different interpretation would better serve the statute's objectives, it may have authority to depart from or modify the Court's interpretation.

The Supreme Court's opinion answered some questions—generally siding with the Biosimilar's position—but left much uncertainty. How soon can a Biosimilar give notice of commercial marketing? Can it be when filing its aBLA, or is the BPCIA (and its notice provisions) only triggered by the FDA's acceptance of the aBLA?

In addition, the Court expressed no view on whether a district court could take into account a Biosimilar's failure to follow § 262(l)(2)(A) (or any other BPCIA procedural requirement) in deciding whether to grant a preliminary injunction under 35 U.S.C. § 271(e)(4)(B) or § 283 preventing marketing of the Biosimilar's product.[19] Thus, failure to comply with the information exchange provision of § 262(l)(2)(A) may have significant ramifications beyond those at issue in this one case.

Can Parties Only Participate in Some, But Not All of the Patent Dance?

While the Supreme Court held that Biosimilars cannot be forced to provide their aBLA and manufacturing information to the Sponsor under § 262(l)(2)(A), questions remain regarding the enforceability of the other provisions of the patent dance and the ramifications for skipping them. Recently, several Biosimilars have begun testing the boundaries of the patent dance by providing portions of their aBLAs and manufacturing information, but "waiving" additional requirements of the dance.

For example, in *Immunex Corp. v. Sandoz, Inc.*[20] involving a biosimilar of Enbrel®, Sandoz began the patent dance by providing Immunex with information

[18] *Id.* at 1678.

[19] *Id.* at 1675 n.2.

[20] 2:16-CV-01118 (D.N.J. 2016).

under § 262(l)(2)(A). After Immunex provided Sandoz with a list of patents that Immunex believed could be reasonably asserted under § 262(l)(3)(A), Sandoz responded by "agreeing" to immediate litigation on Immunex's list of patents. In other words, Sandoz waived its right to receive a statement by Immunex under § 262(l)(3)(C) and declared negotiations pursuant to §§ 262(l)(4) and (5) unnecessary. Sandoz then insisted Immunex file suit within 30 days or damages would be limited to a reasonable royalty based on 35 U.S.C. § 271(e)(6). This provision limits a Sponsor's damages for infringement of a patent identified under 42 U.S.C. §§ 262(l)(4) and (5) to a reasonable royalty if the infringement suit is filed more than 30 days after the end of the patent dance. Immunex filed suit against Sandoz in the District of New Jersey within 30 days.

Hospira behaved similarly in its suit with Amgen involving a biosimilar of Epogen®. Hospira provided its aBLA to Amgen, and Amgen then provided a list of patents that could be asserted. Instead of providing its own patent list and detailed statement as authorized by § 262(l)(3)(B), Hospira agreed that every patent Amgen listed would be the subject of the First Wave litigation. Amgen filed suit against Hospira in the District of Delaware before expiration of the 30-day window.[21]

In both instances, the Biosimilars avoided filing their claim-by-claim non-infringement and invalidity contentions, which would have been required had they continued the patent dance.

While previous cases failed to shed light on the ramifications for skipping parts of the patent dance, at least one court has addressed the issue. In the dispute between Janssen and Celltrion involving a biosimilar of Remicade®, Celltrion short-circuited the patent dance by skipping the negotiation steps of §§ 262(l)(4) and (5).[22] After Janssen served its patent list, Celltrion provided a detailed statement in response, and agreed that all of the patents identified by Janssen would be the subject of the First Wave litigation. In the course of its litigation with Janssen, Celltrion moved to dismiss the complaint for lack of standing based on Janssen's alleged failure to add a necessary party. To guide settlement negotiations, the court sought to clarify whether Janssen would be prevented from seeking lost profits under the BPCIA pursuant to § 271(e)(6)(B) in the event the action was dismissed, and Janssen was forced to refile the complaint after the 30-day window expired. The court found that the ordinary meaning of the term "shall," as used throughout §§ 262(l)(4) and (5), indicates a mandatory directive. Thus, Celltrion must have either engaged in good faith negotiations or, failing that, in the patent dance itself before it could limit Janssen's recovery to a reasonable royalty. The court held that only the list of patents that emerge from a properly completed patent dance "are potentially subject to the reasonable royalty damages limitation." In other words, the 30-day time limit for Janssen to file suit, or have its recovery limited to a reasonably royalty, was never triggered because Celltrion failed to complete the patent dance.[23]

[21] *Amgen Inc. v. Hospira, Inc.,* 232 F. Supp. 3d 621 (D. Del. 2017).

[22] *Janssen Biotech Inc. v. Celltrion Healthcare Co.,* 239 F. Supp. 3d 328 (D. Mass. 2017).

[23] *Id.* at 332.

The *Janssen* decision suggests that one possible consequence for parties that choose to participate in the patent dance, but skip a few steps of the dance, is that 35 U.S.C. § 271(e)(6)(B) will not limit a Sponsor's recovery to a reasonable royalty. What additional ramifications may result from completing only parts of the patent dance remain unclear. But certainly, a Biosimilar's lack of good faith participation in the patent dance may be considered as part of a court's analysis in determining injunctive relief.

Which Patents Can a Sponsor Assert Under the BPCIA?

The Federal Circuit in *Amgen Inc. v. Hospira, Inc.* recently issued an opinion in an action over a biosimilar of Amgen's Epogen® regarding which patents a Sponsor can assert in BPCIA litigation.[24] In that case, under § 262(l)(2)(A), Hospira produced its aBLA. Amgen, however, requested information regarding some cell culture media components, which were not available from the aBLA so it could assess infringement of certain patents. Hospira, however, refused to provide the information and as a result, Amgen did not include these patents on its § 262(l)(3)(A) list. Amgen reasoned that this list is limited to patents "for which the reference product Sponsor believes a claim of patent infringement could reasonably be asserted." Without access to the cell culture information, Amgen believed it could not determine whether these patents "could reasonably be asserted." Amgen was also concerned it could be subject to sanctions under Fed. R. Civ. P. 11 if it asserted infringement without the required good faith belief that Hospira's cell culture media infringed these patents. The parties continued to engage in the patent dance, which resulted in litigation that did not involve Amgen's cell culture patents.

During discovery, Amgen attempted to obtain the manufacturing information again, but was ultimately unsuccessful. As such, the district court denied Amgen's motion to compel discovery, holding that Amgen was precluded under Fed. R. Civ. P. 26, from seeking discovery of information unrelated to the patents-in-suit. Since Amgen had not asserted its cell culture patents, no discovery on these patents was relevant to the issues in the pending litigation. Amgen appealed.

Initially, the Federal Circuit denied Hospira's motion to dismiss the appeal while leaving open the question of jurisdiction, requesting briefing on the merits of the discovery dispute as well as additional briefing on whether the court had jurisdiction under the collateral order doctrine or the All Writs Act. These provisions, if satisfied, allow an immediate appeal of an otherwise unappealable district court decision, e.g., discovery disputes.[25] The Federal Circuit later held that it lacked jurisdiction because the denial of the discovery motion was not a collateral order. Rather,

[24] *Amgen Inc. v. Hospira, Inc.,* 866 F.3d 1355 (Fed. Cir. 2017).

[25] *Amgen, Inc. v. Hospira, Inc.*, No. 16-2179 (Fed. Cir. Aug. 12, 2016), ECF No. 16.

the court characterized this as a "run-of-the-mill discovery dispute" that could be appealed at the end of the district court litigation.[26]

Significantly, the Federal Circuit explained that there would be no Rule 11 sanction for mistakenly listing a patent on a § 262(l)(3)(A) list. "The statute provides no sanction for holding or asserting a mistaken belief in good faith" when a reference product Sponsor lists patents under § 262(l)(3)(A).[27] Thus, the court concluded that "the reasonableness requirement of § 262(l)(3)(A) does not preclude a Sponsor from listing a patent for which an applicant has not provided information under § 262(l)(2)(A)."[28] Once a patent is listed on a § 262(l)(3)(A) list, the Biosimilar—if it chooses to dance—must come forward with additional disclosures under § 262(l)(3)(B) that inform the Sponsor whether a claim of patent infringement could reasonably be asserted. Also, the court noted that Rule 11 only requires good faith "to the best of the person's" ability after a reasonable inquiry. "Thus, if a Sponsor forms a belief based on an inquiry limited by an applicant's withholding of information, the Sponsor has still satisfied Rule 11."[29]

Amgen v. Hospira thus holds that Sponsors can list all potentially assertable patents on its § 262(l)(3)(A) list, even if all information needed to assess infringement is not available due to the Biosimilar's conduct. If the Biosimilar, however, provides full access to the processes, the Sponsor must establish the bases for including such patents on its § 262(l)(3)(A) list. Significantly, this only applies if the Biosimilar chooses to continue the dance.

How Much and What Type of Information is Needed Under the "Confidential Exchange" Provision?

Biosimilars that choose to participate in the patent dance must provide the Sponsor with certain confidential information. The statute states:

> Not later than 20 days after the Secretary notifies the [Biosimilar applicant] that the application has been accepted for review, the [Biosimilar applicant] shall provide to the [Sponsor] a copy of the [aBLA] submitted to the Secretary ... and such other information that describes the process or processes used to manufacture the biological product that is the subject of the [aBLA].[30]

The Biosimilar may provide the Sponsor with "additional information" upon request.[31] The statute, however, provides no guidance on what constitutes "additional information."

---

[26] 866 F.3d at 1360.

[27] *Id.* at 1362.

[28] *Id.*

[29] *Id.*

[30] 42 U.S.C. § 262(l)(2)(A).

[31] *Id.*, § 262(l)(2)(B).

As can be seen, § 262(l)(2)‘s wording is vague. It could be argued that the statute requires the Biosimilar to disclose both its entire aBLA and corresponding manufacturing information. Alternatively, a Biosimilar could argue that it complies with § 262(l)(2) if it discloses only its aBLA if that document includes sufficient information concerning the Biosimilar's manufacturing process. Further still, a Biosimilar could argue that a partial disclosure of its aBLA suffices if the portion disclosed provides the Sponsor with adequate information to assess infringement. In view of these interpretations, § 262(l)(2)‘s true meaning is unclear.

The pliable language of this provision allows Biosimilars discretion in crafting their confidential information disclosures. Realizing ambiguity resides, Biosimilars have taken different approaches to the type and amount of information disclosed to the Sponsor. In some cases, the Biosimilar has comprehensively disclosed its aBLA and the corresponding manufacturing information.[32] In others, the Biosimilar has disclosed only its aBLA or has limited its disclosure in some way.

For example, Sandoz's disclosure in its dispute with Immunex in a case regarding Etanercept® was challenged as inadequate.[33] There, in response to Sandoz's § 262(l)(2) disclosure, Immunex alleged that Sandoz "tried to reap the commercial benefits provided to biosimilar manufacturers under the BPCIA while seeking to avoid the obligations in the same Act that Congress established to protect innovators such as Immunex."[34] Immunex alleged that Sandoz provided it "with remote access to a Sandoz-hosted database of TIFF images, modified to include added confidentiality designations, that [Sandoz] represented to constitute its aBLA and information relating to the manufacturing process for [Sandoz's] biosimilar product."[35]

Immunex's complaint clearly alleged that Sandoz's disclosure did not satisfy § 262(l)(2) because: (1) review of the disclosed information was supposedly burdensome; and (2) the information disclosed (including the aBLA) was purportedly not in an unaltered state.[36] Despite these alleged deficiencies, Immunex provided Sandoz with its § 262(l)(3)(A) list of potentially infringed patents. Because the parties did not petition the court to rule on the sufficiency of Sandoz's § 262(l)(2) disclosure, the court did not address this provision of the BPCIA.

Some § 262(l)(2) disputes were even more pronounced than the one in Immunex. For example, the parties in *Janssen Biotech Inc. v. Celltrion Healthcare Co.*[37] vigorously disputed whether Celltrion adequately disclosed its confidential information under § 262(l)(2).[38] There, Celltrion provided Janssen with a copy of its entire aBLA, but refused to disclose information relevant to the manufacture of its

[32] See *Amgen Inc. v. Sandoz Inc.*, No. 16–1276 (D.N.J. 2016).

[33] *Immunex Corp. v. Sandoz Inc.*, No. 3:16-cv-01118 (D. N.J.).

[34] *Sandoz,* No. 3:16-cv-01118, D.I. 1 at ¶ 56.

[35] *Id*. at ¶ 57.

[36] *Id*. at ¶¶ 56–57.

[37] *Janssen Biotech Inc. v. Celltrion Healthcare Co.,* 2015 WL 7078048 (D. Mass. Mar. 6, 2015).

[38] *Id.*

biosimilar product. Janssen's complaint noted that Celltrion's refusal to disclose this information rendered its disclosure inadequate for purposes of § 262(l)(2). Celltrion, however, asserted that Janssen could adequately prepare its § 262(l)(3)(A) list of potentially infringed patents without the manufacturing information and that such information would not be disclosed until Janssen's infringement suit was filed.

The dispute between Janssen and Celltrion focused on whether the Biosimilar has the discretion to refuse disclosure of its manufacturing information when it determines that the Sponsor can adequately gauge whether its patents are infringed by examining only the aBLA. The Massachusetts Court was not asked to decide this issue and has accordingly not resolved the BPCIA's application to this factual presentation.

Amgen's dispute with Hospira®[39] over Epoetin® (discussed above with regard to another issue in the case) involved similar facts to those described in *Janssen v. Celltrion*. There, Hospira, the Biosimilar, turned over its entire aBLA. Notably, according to Hospira, this disclosure included over 507 native files as well as 747,000 additional pages of information concerning Hospira's product and the process used to make it.[40] But Amgen was not satisfied with this disclosure and alleged in its complaint that Hospira violated § 262(l)(2) by not disclosing relevant manufacturing information. Though Amgen conceded that this information could potentially be discovered in subsequent litigation, it complained that Hospira's actions prevented it from conducting a full and complete evaluation of its patent portfolio relevant to Hospira's manufacturing processes. Amgen alleged that Hospira's conduct frustrated the BPCIA's statutory purpose, deprived Amgen of the opportunity of seeking redress for potential infringement, and risked preventing Amgen from ever obtaining information regarding Hospira's biosimilar manufacturing processes (asserting that the delay in disclosure of this information potentially prevents Amgen from discovering it until after Hospira begins marketing its biosimilar).[41] Despite the fact that Hospira's motion to dismiss contended that its disclosure satisfied § 262(l)(2), it never squarely asked the court to determine whether it had satisfied the BPCIA's requirements. Thus, the court never ruled on that particular issue, leaving it for another court to interpret this provision at a later date.

On Feb. 15, 2017, Genentech filed suit with the goal of seeking clarity regarding the meaning of § 262(l)(2)'s disclosure requirements. Genentech sued Amgen in the District of Delaware asking the court to award "urgent declaratory and related relief" in a dispute over bevacizumab.[42] Specifically, Genentech sought a declaratory judgment and an accompanying order that Amgen's § 262(l)(2) disclosure was not sufficient to satisfy that provision. Amgen had only provided Genentech with its aBLA, omitting the relevant biosimilar manufacturing infor-

[39] *Amgen Inc. v. Hospira, Inc.*, No. 15-839-RGA (D. Del. Aug. 5, 2016).

[40] *Amgen,* No. 15–839, D.I. 9.

[41] *Amgen,* No. 15–839, D.I. 1 at ¶¶ 51–53.

[42] *Genentech, Inc. v. Amgen, Inc.*, No. 1:17-cv-00165 (D. Del. Mar. 3, 2017).

mation.[43] Genentech disputed the sufficiency of this disclosure, responding with additional disclosure requests to Amgen, specifically tying the requests to its ability to assess infringement of its patents. Amgen did not provide additional information, leaving Genentech without the requested manufacturing details.[44] Interestingly, Genentech noted that Amgen's position in this dispute was at odds with Amgen's position in *Amgen v. Hospira, supra.* In particular, Genentech noted that "Amgen has acknowledged in other BPCIA litigation (where it is the [Sponsor], not the copier), a Patent Owner cannot fully protect itself as Congress intended if the applicant only produces its aBLA, because many important details about the product are normally omitted. Indeed, when [Hospira] produced only its aBLA ... [,] Amgen sued Hospira in [the District of Delaware] for noncompliance with the BPCIA."[45]

On March 1, 2017, Judge Gregory M. Sleet dismissed Genentech's complaint for lack of subject matter jurisdiction, noting that Genentech's only remedy for a violation of § 262 (l)(2) is to sue for patent infringement.[46]

Thus, years after the implementation of the BPCIA, what is required under § 262(l)(2) remains unclear. The pliable language of that section allows Biosimilars some discretion in crafting their confidential information disclosures. As the above examples demonstrate, some Biosimilars have taken the view that § 262(l)(2) does not require disclosure of the full aBLA, whereas Sponsors interpret the statute to require just the opposite. Biosimilars should be aware that less-than-complete disclosures or disclosures made in certain limited formats may spur complaints from Sponsors and may be considered by courts when deciding motions for injunctive relief. Otherwise, it would appear that Biosimilars control what is provided under § 262(l)(2).

IPRs Are Another Avenue to Challenge Biologic Patents[47]

A number of Biosimilars have turned to *inter partes* review ("IPR") proceedings to challenge the validity of patents that may cover their proposed biosimilar products or processes prior to submission of their biosimilar applications to FDA. This is because the Federal Circuit has held that a Biosimilar cannot file a declaratory judgment action challenging a Sponsor's patent(s) before the former files its aBLA.[48] Moreover, after the Biosimilar application is filed, if the Biosimilar engages in the patent dance, the BPCIA prohibits a declaratory judgment by either

[43] *Id.*, D.I. 1 at ¶ 6.

[44] *Id.*

[45] *Id*. at ¶7.

[46] *Genentech, Inc., v. Amgen Inc.*, No. 1:17-cv-00165 (D. Del. Mar. 3, 2017), Doc. 16.

[47] For further information, *see generally* http://fishpostgrant.com/inter-partes-review/

[48] *Sandoz, Inc. v. Amgen Inc.,* 773 F.3d 1274 (Fed. Cir. 2015).

party.[49] If the Biosimilar fails to provide its aBLA or related manufacturing details, the Sponsor may file a declaratory judgment action asserting all its potentially relevant patents. Thus, IPRs are the only way a Biosimilar can challenge a Sponsor's patents before litigation ensues.

IPRs are a faster and often lower cost means to challenge a patent under a lower burden of proof than required in district court litigation. The proceedings are decided by administrative patent judges with significant patent and technical backgrounds, including chemistry and biology. IPRs may allow a petitioner to avoid the uncertainty of the patent dance under the BPCIA.[50] But petitioners must be aware of the limitations—IPR challenges may only be brought under anticipation or obviousness grounds and only based on patents and printed publications. Moreover, the potential for estoppel in a subsequent litigation exists.

How IPRs Work

IPR proceedings are conducted at the Patent Trial and Appeal Board ("PTAB," or "the Board") to review the patentability of one or more claims in a patent on grounds that could be raised under 35 U.S.C. § 102 (lack of novelty) or § 103 (obviousness), and only on the basis of patents or printed publications.[51] A petition for IPR may be filed by anyone other than the owner of the patent.[52] Importantly, a petitioner is barred from filing an IPR more than 1 year after being served with an infringement complaint.[53] The Patent Owner may file a preliminary response to the petition, and the IPR must be instituted or denied within 6 months of the petition.[54]

An IPR trial will be completed within 1 year from institution (extendable for good cause by 6 months), and the decision can be appealed to the Federal Circuit.[55] Direct testimony at the PTAB takes place through written declarations, while cross-

[49] 42 U.S.C. § 262 (l)(9)(A).

[50] Post grant review (PGR) is less applicable to biologic patents because it is limited to "first-to-file" patent within 9 months of issue. Although PGR allows challenges under almost any statutory ground for unpatentability, including § 112, it requires a higher threshold for institution than IPRs under the "more likely than not" standard. Moreover, because the estoppel effect of PGR is very broad (extending to all statutory bases for invalidity), petitioners who wish to preserve invalidity arguments for district court may not prefer the PGR route.

[51] 35 U.S.C. § 311(b).

[52] A petitioner must also identify any "real party-in-interest." 35 U.S.C. § 315(b). A real party in interest has been defined as "a party that funds and directs and controls an IPR or PGR petition or proceeding." Trial Practice Guide (Rule), 77 Fed. Reg. 48759-60 (Aug. 14, 2012).

[53] 35 U.S.C. § 315(b).

[54] 35 U.S.C. § 313.

[55] 35 U.S.C. § 316(a).

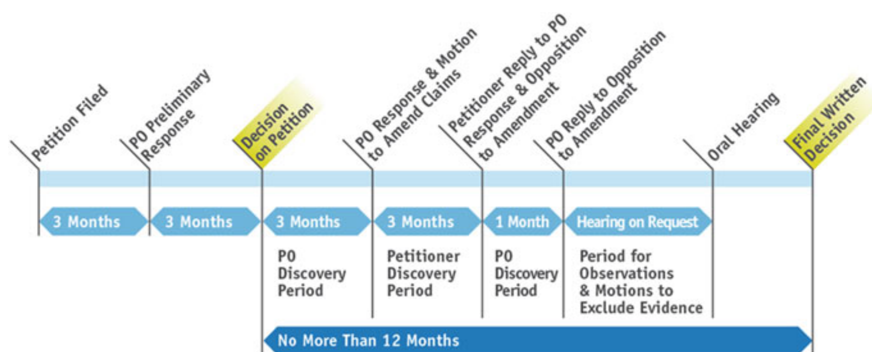


Fig. 4.2 Timeline for an IPR

examination occurs during the deposition of any declarants.[56] Although under PTAB rules a Patent Owner may amend its claims during an IPR, the Board has rarely allowed such amendments.

The hearing is an oral argument before a panel of three administrative patent judges with, as noted *supra,* significant patent experience and technical backgrounds. The limited discovery and lack of live testimony streamlines time and costs, but it can also limit the ability of the Patent Owner to tell an invention story.[57] The IPR Timeline is shown in Fig. 4.2 above.

Legal Standards

The Board will decide to institute based upon a showing that there is a reasonable likelihood that the petitioner would prevail with respect to at least one claim challenged. Importantly, if the Board institutes, it must issue a final written decision on all challenged claims and all grounds raised in the petition. About 48% of biologic-related petitions to date have been instituted, which is lower than the 68% overall institution rate. To date, approximately 90 IPRs have been instituted against more than 45 biologic patents and 1300 claims. The large majority of these proceedings were directed towards the early biologics Rituxan®, Humira®, and

[56]Expert testimony is critical to both the petition and the Patent Owner's response. The Patent Owner can provide expert testimony to rebut a petitioner's expert evidence before the Board makes a decision on institution. IPR institution is, in fact, a preliminary determination on the claim construction and prior art. Engaging an expert early in the process is key.

[57]Discovery is very specific and is typically limited to issues surrounding privity/real party in interest, secondary considerations of non-obviousness, and facts relied upon by an expert that are not readily available.

Herceptin®, which were first, fourth and eighth on the list of best-selling drugs in 2016, with over $31 billion in combined sales.

The PTAB uses a lower burden of proof than applied in district court litigation. Petitioners need only show unpatentability by a preponderance of the evidence (a more likely than not standard), a significantly lower threshold than the "clear and convincing evidence" standard applied in district court actions.

A Supreme Court decision and a proposed PTAB rule change will change PTAB institution and claim construction rules. In April 2018, in *SAS Inst., Inc. v. Iancu*, the Supreme Court held that the Board must decide the validity of every challenged patent claim when it agrees to institute an IPR.[58] The PTO published guidance 2 days after the *SAS* decision stating that "if the PTAB institutes a trial, the PTAB will institute on all challenges raised in the petition," meaning all grounds included in the petition, and not just all claims.[59] Thus IPR proceedings can no longer have non-instituted grounds—if the PTAB institutes trial on one claim, and all claims have been challenged, it must issue a final written decision on all challenged claims and all grounds raised in the petition. The Federal Circuit has endorsed the PTAB's approach, finding that "[e]qual treatment of claims and grounds for institution purposes has pervasive support in *SAS*."[60]

In May 2018, the PTO proposed changing from use of the "broadest reasonable interpretation" claim construction standard in PTAB trials to the "plain and ordinary meaning" standard used by district courts.[61] This change is noteworthy because the Board's "broadest reasonably interpretation,"[62] could encompass more invalidating prior art than the "plain and ordinary meaning" standard.[63] The PTO also proposed allowing the Board to consider any prior claim construction determination concerning a term of the claim in a civil action, or an ITC proceeding, that is timely made of record in an IPR.[64] At the time of writing, the change in claim construction standard had not been implemented.

[58] 138 S. Ct. 1348, 1359–60 (2018). It is not clear whether SAS has retroactive effect on past decisions in which the PTAB instituted and tried cases on fewer than all challenged claims or on only a subset of the grounds requested.

[59] PTAB, *Guidance on the Impact of SAS on AIA Trial Proceedings* (April 26, 2018).

[60] *PGS Geophysical AS v. Iancu*, 891 F.3d 1354, 1360 (Fed. Cir. 2018).

[61] 83 Fed. Reg. 21,221 at 21224 (May 9, 2018).

[62] 37 C.F.R. § 42.100(b); *Cuozzo Speed Techs., LLC v. Lee*, 136 S. Ct. 2131, 2142 (2016) (affirming applicability of broadest reasonable construction standard to *inter partes* review proceedings).

[63] *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005).

[64] 83 Fed. Reg. 21221 at 21222 (May 9, 2018).

Estoppel

A petitioner and its privies are estopped from asserting invalidity in a subsequent proceeding "on any ground that the petitioner raised or reasonably could have raised during that *inter partes* review."[65] This applies to disputes in the PTO and ITC as well as in district court. In *Shaw Industries Group v. Automated Creel Systems*, the Federal Circuit held that estoppel does not apply to grounds denied by the PTAB because an "IPR does not begin until it is instituted."[66] District courts post-*Shaw* have thus held that grounds asserted, but denied in an IPR can still be raised in subsequent litigation. The implications of this decision are clearer after *SAS*.

SAS could broaden IPR estoppel because it eliminated partial institutions—if the PTAB institutes trial on one claim, it must issue a final written decision on all challenged claims and all grounds raised in the petition. Thus, after *SAS*, the rationale underlying the *Shaw* decision may no longer apply. It remains to be seen how district courts will treat estoppel after *SAS*.

The PTAB has taken a broader view of estoppel in denying institution of an IPR based prior art that was not raised, but reasonably could have been raised in the prior proceeding.[67] In doing so, the PTAB interpreted *Shaw* to mean that "estoppel does not apply to any ground of unpatentability that was presented in a petition, but denied institution," and further distinguished between grounds raised, but denied institution from grounds that the petitioner could have raised, but elected not to.[68]

Finally, under 37 C.F.R. § 42.73(d)(3), a Patent Owner is precluded from taking action at the PTO that is inconsistent with an adverse judgment in a district court litigation. The Federal Circuit, however, has stated that "the addition of dependent claims as a hedge against possible invalidity" is a legitimate justification for filing a reissue application in these situations.[69]

Statistics

Biosimilar-related IPR petitions have jumped from 4 in 2014, to 10 in 2015 to over 50 in 2017. This trend will likely continue as parties become more familiar with the procedure and grow their biosimilar programs (as of June 2018, there are 12 approved biosimilars and numerous biosimilars applications in the pipeline). Nearly nine out of ten IPRs in this field have associated district court litigation.

[65] 35 U.S.C. § 315(e).

[66] *Shaw Industries Group v. Automated Creel Systems*, 817 F.3d 1293, 1296, (Fed. Cir. 2016).

[67] *Great West Casualty Co. et al. v. Intellectual Ventures II*, IPR2016-01534, Paper No. 13 (PTAB Feb. 15, 2017)

[68] *Id*. at 12-13.

[69] *Ex Parte Kobelco Research Inst., Inc., Patent Owner & Appellant*, 2010–009563, 2011 WL 3793611, at *2 (Aug. 24, 2011) (citing *In re Yasuhito Tanaka*, 640 F.3d 1246 (Fed. Cir. 2011)).

Analysis has shown that the majority of biologic claims challenged have been formulation and method-of-treatment patents. Composition-of-matter and process patent challenges have been rarer.[70]

Strategies

Petitioners

Petitioners may use the IPR proceedings as part of a "freedom to operate" strategy to clear out patents before a biosimilar application is filed. This is a faster option than district court litigation and offers the possibility of removing key patents from a First or Second Wave litigation under the BPCIA. IPRs are particularly synergistic when defenses under §§ 112 and 101 also exist, as these can be preserved for district court litigation.

Patent Owners

Patent Owners can strengthen patents against IPR petitions and increase the petitioner's burden by including a large number of claims in their applications, including multiple claims of varying scope. This may help limit a petitioner's attack because of the page limits. A Patent Owner should also consider having multiple patents covering aspects of the invention because this makes it more expensive to challenge the patent portfolio. Maintaining pending continuation and divisional applications keeps a patent family alive and allows for amendments, which, in the past, were rarely permitted in IPRs.[71] Patent Owners may have a better shot at amending patent claims under two recent Federal Circuit decisions, which held that a petitioner challenging the validity of the patent has the burden of showing that the new claims are unpatentable.[72]

---

[70] *See, e.g.*, Molenda and Praseuth (2017).

[71] A 2016 PTAB report showed that the Board had denied 112 out of 118 motions to amend.

[72] In *Aqua Products, Inc. v. Matal,* the Federal Circuit discarded the PTAB rule that Patent Owners seeking to amend their patents in IPRs have the burden of proving that the proposed new claims are patentable. 872 F.3d 1290 (Fed. Cir. 2017). The majority held that the Petitioner challenging the validity of the patent has the burden of showing that the new claims are unpatentable. *See also Bosch Automotive Serv. Sols., LLC v. Matal*, 878 F.3d 1027, 1040 (Fed. Cir. 2017), as *amended on reh'g in part* (Mar. 15, 2018) ("the petitioner bears the burden of proving that the proposed amended claims are unpatentable by a preponderance of the evidence"). The PTO issued guidance on motions to amend in view of *Aqua Products* stating that if a Patent Owner files a motion to amend that meets the requirements of 35 U.S.C. § 316(d) (i.e., proposes a reasonable number of substitute claims, and the substitute claims do not enlarge scope of the original claims of the patent or introduce new matter), the Board will proceed to determine whether the substitute claims are unpatentable by a preponderance of the evidence based on the entirety of the record, including

The Patent Owner's preliminary response may provide an opportunity to shape the PTAB's views on claim construction. This is the time to focus on issues amenable to favorable resolution based on the petitioner's evidence or lack thereof, e.g., lack of all necessary parties or untimely filing.

Although not part of the BPCIA, the above discussion has been included because of the key role IPRs have had and will continue to have in conjunction with biosimilar disputes.

Conclusion

The workings of the BPCIA itself and litigation-related issues are being raised and in many instances resolved by the courts. It is hoped that the proceeding discussion sheds some light on many of the issues, provides answers where possible, and outlines suggestions for a Sponsor's and Biosimilar's participation in these proceedings.

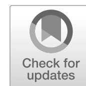
References

Carver KH, Elikan J, Lietzan E. An unofficial legislative history of the biologics price competition and innovation act of 2009. Food Drug Law J. 2010;65(4):671–818, ii.
Coggio BD. Research tools and the Hatch-Waxman safe harbor. Biotechnol Law Rep 2014;22(1).
Coggio BD, Vogel R. Can reference product sponsor forfeit right to sue under BPCIA. IP Law360. 2016 Sept 25.
Coggio BD, Vogel R. Has Amgen already won its BPCIA dispute with Sandoz. IP Law360. 2017 Aug 2.
John Molenda, Richard Praseuth. Current trends in biologics-related inter partes reviews. IP Law360. 2017 July 20.
PTO. Guidance on motions to amend in view of aqua products. Alexandria, VA: USPTO; 2017.
Vogel R, Coggio BD. Safe harbor in the United States and Europe. Pharm Law Ind Rep. 2016;14(47).

any opposition made by the Petitioner. *See*, PTO (2017). Practitioners note that the impact of these changes is unclear because most proposed amended claims have been rejected because the Board has found they are not distinguished from the prior art, and it not clear that shifting the burden of persuasion on petitioners will change that (especially when the burden of coming forward with evidence remains with the Patent Owner).

Part II
Regulatory Aspects of Development and Approval for Biosimilars

Chapter 5
Design and Implementation of Successful Regulatory Strategies in Biosimilar Development



Kamali Chance

Abstract Biological medicines have had a profound impact on the health of patients suffering from many debilitating and life threatening diseases. They have been shown to provide "dramatically reduced disability for patients with inflammatory diseases, such as rheumatoid arthritis, extended the lives of patients with many cancers and also provide lifesaving replacement proteins for patients with rare diseases" (US Food and Drug Administration, http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm048341.htm; Krishnan et al., Ann Rheum Dis 71:213–218, 2012; Cox, Biologics 4:299–313, 2010). Biosimilars are copies of biological medicines and in order to obtain approval they are required to undergo head to head similarity exercises [CMC, nonclinical and clinical (as needed)] against their reference product already marketed in the region/country of interest or countries with stringent regulatory requirements.

Keywords Biosimilar · Biosimilarity · Similarity · Regulatory strategy · Global development · CMC · Nonclinical and clinical

Introduction

Biological medicines have had a profound impact on the health of patients suffering from many debilitating and life threatening diseases. They have been shown to provide "dramatically reduced disability for patients with inflammatory diseases, such as rheumatoid arthritis, extended the lives of patients with many cancers and also provide lifesaving replacement proteins for patients with rare diseases" (US Food and Drug Administration 2012; Krishnan et al. 2012; Cox 2010). Biosimilars are copies of biological medicines and in order to obtain approval they are required to undergo head to head similarity exercises [CMC, nonclinical and clinical (as

K. Chance (✉)
BioSciencesCorp, Graham, NC, USA
e-mail: kamali.chance@biosciencescorp.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_5

needed)] against their reference product already marketed in the region/country of interest or countries with stringent regulatory requirements.

Biosimilars share the same amino acid sequences as their originator biologic but are made up of proteins with potential for posttranslational changes during manufacture, including glycosylation, phosphorylation, among other modifications which may have implications for immunogenicity. The impact of these changes has to be studied in head-to-head similarity studies (analytical/biological, nonclinical and clinical) to ensure that the efficacy and safety profiles of the resulting biosimilar is not negatively impacted (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf; https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf; http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176768.pdf).

In biosimilars drug development, developing a robust global regulatory strategy can help your organization achieve its goals in a more streamlined fashion which could potentially lead to earlier market entries with reduced development costs. For a global regulatory strategy framework (Fig. 5.1) some upfront strategic considerations include identifying countries/markets of interest; competitive intelligence; understanding biosimilar regulations per country/region; Chemistry, Manufacturing and Controls (CMC) requirements, nonclinical and clinical requirements per country/region; and any relevant bridging data requirements. Patents information gathering is also important and is covered in Section I of this book under "Business, Health Economics & Intellectual Property Landscape for Biosimilars". Once all of this information is available, it is matter of pulling all of the information together for meaningful workable regulatory strategy.

It truly requires a teamwork approach from many key staff members who are experts in their relevant fields. For example, the business development staff member will be crucial in identifying countries/markets of interest based on return on investments considerations. Technical CMC, nonclinical and clinical experts will be crucial in product development and testing. The regulatory expert would leverage his/her regulatory knowledge regarding regulatory requirements for countries/regions of interest to help put the overall strategic plan together. Once this information is in hand, then a global strategic roadmap can be constructed to help your organization achieve its short and long term goals.

This chapter examines the key aspects involved in design and implementation of successful regulatory strategies for the development of biosimilars including countries of interest/competitive intelligence; biosimilar regulatory pathways (EMA, FDA and WHO); understanding CMC; nonclinical (in vitro and in vivo); clinical requirements including bridging data; and finally putting the regulatory strategy together.

Countries of Interest/Competitive Intelligence

The business executives in an organization are in a key position to weigh the costs and benefits of investing in biosimilars in various markets of interest as they work

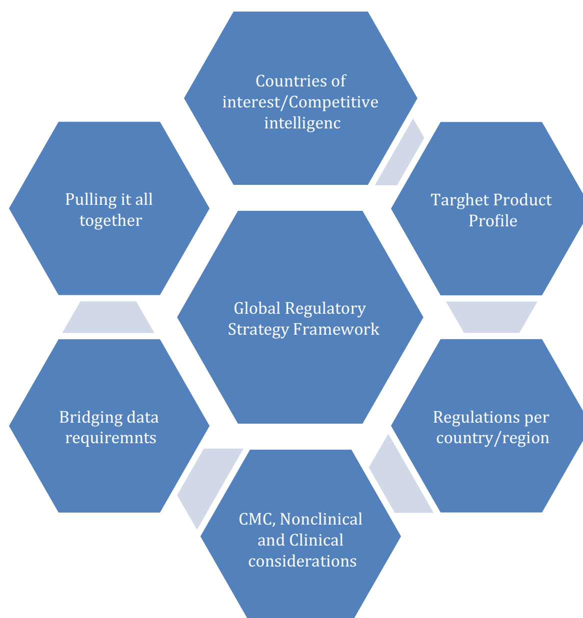


Fig. 5.1 A global regulatory strategy framework

with their strategic team members to understand the development requirements for the various countries/regions relative to return on investment. In order to streamline the product development it is crucial to know upfront the countries/regions of interest for commercialization. This knowledge will help determine what additional bridging studies may be required for a complete CMC package and whether there is a requirement for including patients/healthy subjects from a particular country in a global pivotal PK or efficacy/safety (including immunogenicity) study in order to meet country specific requirements. Multiple studies may be required to satisfy regulators from various countries of interest, if the initial study design does not take this into consideration. This will result in increased costs and longer timelines for bringing the product to market.

If insights into competitive intelligence are not garnered at the outset, this may result in developing a product that is extremely late to market for which return on investment could be minimal. It behoves the organization to gather data on companies which are in the process of developing products of interest and at which stage of development these products are in (Preclinical, Phase I, Phase III, approved and marketed, etc.). An overview of the competitive landscape for the biosimilar of interest is crucial. In order to make a meaning business case, the prevalence of disease information in countries of interest and related global sales information is also needed for making proper projections moving forward.

There are many organizations whose primary function is performing ongoing market research services for the pharmaceutical/biopharmaceutical industry, leveraging such data will help your organization in pulling together actionable insights to maximize return on investment.

Product Target Profile

Product target profile will be limited by the approved indications for the reference product for the countries/regions of interest (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM444661.pdf). The dose, dosing regimen, product administration, etc. will be dictated by what is already approved for the reference product as improvements on the existing product do not fall under the purview of a product being a biosimilar. Based on the country or region of interest, there may be some latitude with regards to the medical device that may be approved for administration of the drug product such as an autoinjector (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM444661.pdf). One needs to be cognizant of patent protections/marketing exclusivities which may be in play (https://www.fda.gov/downloads/drugs/ucm216146.pdf). There may be some options for carving out certain indications such as pediatric indications that may have exclusivity. Discussions with regulators from the countries or regions of interest are highly encouraged.

It should be kept in mind that the overall goal in developing the product target profile is to demonstrate biosimilarity to its reference product. If a potential biosimilar product is developed that shows improved efficacy it will no longer qualify as a biosimilar. A biosimilar product with improved safety profile however, generally would be considered a biosimilar.

Biosimilar Regulatory Pathways

Regulatory authorities such as the European Medicines Agency (EMA) and US Food and Drug Administration (FDA) and other highly regulated markets assess each biosimilar medicinal product development program on a case-by-case basis based on the guidelines currently in effect. Many of the countries around the globe to a large degree have adopted close versions of biosimilar guidelines to those in effect in the US and EU.

Many of the CMC, nonclinical and clinical guidelines are very similar for the US and EU, therefore they will be addressed together, with differences highlighted under CMC, nonclinical and clinical considerations.

EMA Guidelines

EMA was first to issue biosimilar guidelines. Since the European Union approved the first biosimilar—Sandoz' recombinant human growth hormone, Omnitrope, a biosimilar to Pfizer's Genotropin recombinant human growth hormone—in 2006 (http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/000607/WC500043689.pdf; http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/general/general_content_001832.jsp) and this reached the market in 2007, a total of 37 biosimilars have been approved in that region to date (http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000408.jsp&mid=WC0b01ac058002958c).

The EU pathway for biosimilar products approval is established in Article 10(4) of Directive 2001/83/EC. Over the past 10 plus years, EMA has established general and product specific guidelines for the development of biosimilars (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf). These guidance documents lay down the analytical, in-vitro and in-vivo nonclinical and clinical requirements for a biological medicinal product claiming to be similar to another product already marketed in the EU. Over the years, based on experience gained in working with biosimilars and improvements in analytics, EMA has updated many of their guidelines for general and product specific guidances.

EMA encourages stepwise development of biosimilars starting with comprehensive analytical similarity assessment followed by in vitro biological assays and if warranted, in vivo nonclinical study based on need. If non-human primates are the only relevant species, the conduct of *standard repeated dose toxicity studies is usually not recommended*. If appropriately justified, a repeated dose toxicity study with refined design (e.g., using just one dose level of biosimilar and reference product and/or just one gender and/or no recovery animals) or an in-life evaluation of safety parameters is recommended (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf).

Generally, EMA expects a clinical study or studies designed to establish statistical evidence that the proposed product is neither inferior nor superior to the reference product by more than a specified margin. In addition, if the selected PD marker/biomarker is an accepted surrogate marker for clinical outcome, comparative PK/PD studies may be sufficient to demonstrate clinical comparability (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf). Examples include:

- ANC to assess the impact of GCSF
- Early viral load reduction in chronic hepatitis C to assess the effect of alpha interferons
- Euglycaemic clamp test to compare two insulins
- Magnetic resonance imaging of disease lesions can be used to compare two β-interferons in multiple sclerosis

It should be noted that increased immunogenicity as compared to the reference product will question biosimilarity. However, a lower immunogenicity level is acceptable.

With regards to number of clinical studies, if the mechanism of action (MOA) is found to be the same for all indications approved for the reference product, then with adequate scientific justifications, a biosimilar will be granted all indications upon conduct of a successful clinical study in a sensitive indication. If, however, different MOAs are known to exist for the reference product for the various indications, then the likelihood is greater that the sponsor would have to conduct multiple confirmatory trials. Rituximab is one product for which at least two clinical studies would be required, PK in rheumatoid arthritis indication and confirmatory study in an oncology indication (https://www.fda.gov/RegulatoryInformation/LawsEnforcedbyFDA/FederalFoodDrugandCosmeticActFDCAct/FDCActChapterVDrugsandDevices/default.htm#Part_A).

EMA encourages sponsors to seek scientific advice early and often to ensure that their product development plan is robust and will likely meet EMA requirements. A regulatory filing for a proposed biosimilar to a reference biological product has to include complete administrative and quality data, together with appropriate nonclinical and clinical data when the Marketing Authorization Application (MAA) is submitted. The designation of interchangeability status is left to each member state in the EU, it does not fall under the purview of EMA.

FDA Guidelines

The regulatory pathways in the US for the approval of drug products fall under the Federal Food, Drug, and Cosmetic Act (FD&C Act) (https://www.fda.gov/RegulatoryInformation/LawsEnforcedbyFDA/FederalFoodDrugandCosmeticActFDCAct/FDCActChapterVDrugsandDevices/default.htm#Part_A) and the Public Health Service Act (PHS Act) (http://uscode.house.gov/view.xhtml?req=(title:42%20section:262%20edition:prelim)). Section 505 of the *FD&C Act* describes all three types of new drug applications that can be submitted to FDA for approval. See Fig. 5.2 for drug approval pathways for drugs/biologics in the US.

- *New Drug Applications covered under FD&C Act section 505(b)(1)*—This type of application contains full reports of safety and efficacy investigations as well as some of the information from studies not conducted by or for the applicant for which the applicant has obtained a right of reference.
- *505(b)(2) New Drug Applications*—This type of application relies upon "at least some of the information from studies not conducted by or for the applicant and for which the applicant has not obtained a right of reference," although some studies may have been conducted by the sponsor (https://www.fda.gov/downloads/Drugs/Guidances/ucm079345.pdf).

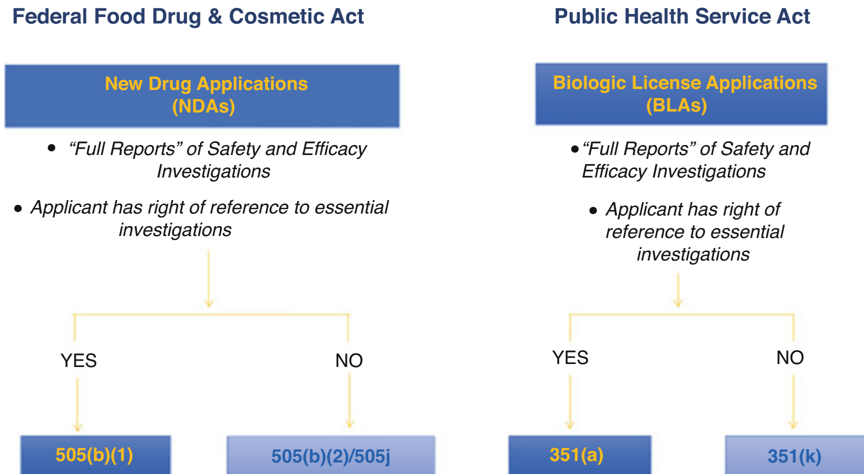


Fig. 5.2 Drug/biologics approval pathways in the US

- *Abbreviated New Drug Applications* covered under *FD&C Act section 505(j)*—This type of application includes information that shows that the proposed product is identical to a previously approved drug product in the following respects: active ingredient, dosage form, strength, route of administration, labeling, quality, performance characteristics and intended use.

In reference to the FD&C Act, most follow on biologics (copies of previously approved biologics via the NDA pathway) will follow the 505(b)(2) pathway (https://www.fda.gov/downloads/Drugs/Guidances/ucm079345.pdf). To date, the FDA has approved several similar biological products via the 505(b)(2) pathway as NDAs. Some of the follow on biologics previously approved by the FDA pursuant to Federal FD&C Act include the following (https://www.accessdata.fda.gov/scripts/cder/daf/):

- Omnitrope (somatropin), a growth hormone; reference product Genotropin®—505(b)(2)—2006
- Valtropin (somatropin), a growth hormone; reference product Humatrope®—505(b)(2)—2007
- Hylenex (Hyaluronidase), family of enzymes that degrade hyaluronic acid to increase tissue permeability; reference product Wydase—505(b)(2)—2005
- Fortical (Calcitonin Salmon), acts to reduce blood calcium (for osteoporosis); reference product Miacalcin—505(b)(2)—2005
- Basaglar (*insulin* glargine injection), for type 1 and type 2 diabetes; reference product Lantus, 505(b)(2)—2016

BLA: 351(a)

- Full analytical development
- Full nonclinical development
- Full Reports" of Safety and Efficacy Investigations for each indication

BLA: 351(k)

- Full comparative analytical development
- Somewhat abbreviated comparative nonclinical development
- Abbreviated Safety and Efficacy Investigations for limited number of indications
- Extrapolation possible,if MOA is same

Fig. 5.3 BLA: 351(a) vs. 351(k)

On March 23, 2020, many of the previously approved NDAs for protein products will fall under the purview of Public Health Service Act and their follow-on protein products will be officially treated as biosimilars (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM444661.pdf). Another exception to the rule has been the FDA approval of Enoxaparin via 505(j) pathway as the sponsor was able to demonstrate finger-print like analytical similarity (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM277709.pdf).

Most biologics, however, are not approved under section 505(b)(1) of the FD&C Act, but fall under section 351(a) of the PHS Act, which originally contained no provision for generic biologics. However, the Biological Price Competition and Innovation Act of 2009 (BPCIA)[1] of H.R. 3590 (the Patient Protection and Affordable Care Act), § 7002, remedied this omission in 2009 and provided the legal mechanism for biosimilars to be regulated under section 351(k) of the PHS Act (https://www.gpo.gov/fdsys/pkg/FR-2011-05-10/pdf/2011-11348.pdf) (see Fig. 5.3).

The biosimilars guidance documents as well as two questions-and-answers guidances regarding Implementation of the BPCIA were first issued on February 9, 2012. FDA's current thinking is reflected in the updates to these guidances in 2015 as well as release of an additional question and answer guidance (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf; https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf; https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM444661.pdf). In 2017, FDA issued two new draft guidances regarding naming of Biologics and Interchangeability (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf; https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/

[1]Biological Price Competition and Innovation Act of 2009 (BPCIA) allowed for the approval of biosimilar products to the previously licensed innovator biologics in the US: https://www.fda.gov/downloads/drugs/ucm216146.pdf

Guidances/UCM273001.pdf). To date, FDA has not issued any product specific biosimilar guidances. As of December 2017, nine biosimilars have been approved in the United States. The FDA has clearly stated that its approach to assessment of demonstration of biosimilarity will include "totality of evidence" as part of the review process of the biologics license application (BLA)[2]. The guidance makes it clear that the FDA does have the authority and the flexibility to determine animal and/or clinical testing requirements on a per-product basis based on an assessment of the comparative analytical and in vitro functional data (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf).

In a global development plan, it is ideal for the pivotal confirmatory clinical studies used to support the MAA in the EU to also be applicable to support the BLA in the U.S. Therefore, in planning the global registration clinical program, global applicability must be a major consideration, with the subject demographics being reflective of the general population of the regions of interest. Also, the reference products must be sourced from the various regions, well characterized, and demonstrated to be similar in terms of CMC (using state-of-the-art analytical testing), nonclinical aspects, and PK. This will provide adequate amount of scientific evidence to support the use of a single active comparator product in pivotal confirmatory clinical study, in geographies where this is allowed.

The BPCIA gives the FDA the authority to determine interchangeability status for a biosimilar product. Details regarding interchangeability requirements are provided below under" Clinical Studies".

The World Health Organization

In 2009, the World Health Organization (WHO) published the "Guidelines on the Evaluation of Similar Biotherapeutic Products (SBPs)," which provide standards for the evaluation these products. Just as with FDA and EMA biosimilar guidelines, these guidelines also require extensive comparability exercises (analytically, non-clinically and clinically) against the reference product. Each country however, is allowed the latitude to designate their own reference product (http://www.who.int/biologicals/areas/biological_therapeutics/BIOTHERAPEUTICS_FOR_WEB_22APRIL2010.pdf). High degree of analytical similarity demonstration can lead to a reduction in non-clinical and clinical data requirements.

In May 2017, the World Health Organization (WHO) announced plans to launch a pilot project for prequalifying biosimilars for cancer (http://www.who.int/mediacentre/news/releases/2017/pilot-prequalification-biosimilars/en/).

[2] Totality of Evidence requirements are detailed in the FDA guidance tiltled "Scientific Considerations in Demonstrating Biosimilarity to a Reference Product" https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf

In September 2017, WHO invited manufacturers to submit applications for prequalification of biosimilar versions of two products in the WHO Essential Medicines List: rituximab and trastuzumab. The WHO will also review its 2009 Guidelines on the evaluation of similar biotherapeutic products (http://www.who.int/biologicals/publications/trs/areas/biological_therapeutics/TRS_977_Annex_2.pdf?ua=1) to ensure that this guidance to national regulatory authorities reflects recent evidence and experience. The WHO has also issued guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (http://www.who.int/biologicals/biotherapeutics/WHO_TRS_1004_web_Annex_2.pdf?ua=1), and has held several workshops on this category of products.

Chemistry, Manufacturing, and Controls

The FDA and EMA (among other regulatory agencies) expect that the expression construct for the proposed biosimilar will "encode the same primary amino acid sequence" as the licensed reference product for the country or region of interest, although the expression system may be different. Minor differences in N- or C-terminal truncations may be acceptable as long as they do not affect product purity, safety, or potency. Ultimately, all differences between the reference product and the proposed biosimilar will require a scientific evaluation of their clinical significance (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf; http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/06/WC500167838.pdf).

The physicochemical assessment of the licensed reference product and the proposed biosimilar should include the following, as applicable: primary, secondary, tertiary, and quaternary structure; post-translational modifications; and functional activities. In the final assessment, the determination of whether a proposed biosimilar can be considered highly similar in quality attributes will depend on factors including the comparative degree of heterogeneity, differences in functional properties, impurity profiles, and degradation profiles (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf; http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/06/WC500167838.pdf).

To streamline the development of biosimilars, all agencies recommend a stepwise approach to product development. They suggest evaluating multiple lots of the reference product throughout its shelf life, as well as proposed biosimilar in development. The acceptance criteria should be based on the differences seen across the various lots of licensed reference product and proposed biosimilar. To avoid the need to conduct further bridging studies late in development, the biosimilar tested in clinical studies should be the one intended for commercialization, and also used for comparative quality and nonclinical testing (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf;

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/06/WC500167838.pdf). FDA highly recommends statistical approach to establishing acceptance criteria (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM576786.pdf).

Nonclinical Studies

The next step in development is to determine what animal studies (including toxicity assessment) may be required based on the differences seen during the comparative analytical testing. This enables the similarity of the proposed biosimilar to be confirmed by non-clinical testing (in vitro and in vivo) against reference products of interest (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf; http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf).

In many countries that have issued biosimilar guidelines—such as Canada, Korea, Brazil, and Mexico—these guidelines are very similar to the EMA guidelines, with minor differences. On this basis, it should be possible for the CMC and non-clinical elements of biosimilar development to be made acceptable to all regulatory agencies in a global development program where similarity of proposed biosimilar to the various reference products has been demonstrated. Having ensured a robust CMC and non-clinical development of the proposed biosimilar against the relevant reference product, it is then important to develop a streamlined clinical development plan (CDP).

Clinical Studies

The guidance documents from relevant countries/regions should be reviewed as a priority, and the recommendations taken into consideration in the clinical development plans. For clinical data, regulatory agencies expect to see comparative PK and/or PD (if a relevant PD marker is available) data in addition to comparative immunogenicity, efficacy and safety data. It should be noted that even if relevant PD measures are not available, sensitive PD endpoints may be assessed if such assessment may help reduce residual uncertainty about biosimilarity. However, the PD measures should be (1) relevant to clinical outcomes; (2) measurable after dosing to ascertain PD response; and (3) have sensitivity to detect clinically meaningful differences. If there is a meaningful correlation between PK and PD results and clinical effectiveness, convincing PK and PD results may make a comparative efficacy study unnecessary (https://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf). It should be noted that more often than not, PK/PD studies are conducted in

healthy volunteers. The exception to the rule is rituximab, for which PK is study is conducted in RA patient population.

If PD marker/endpoint relevant to clinical efficacy is not available, then comparative efficacy and safety data is generally expected. The following rules generally apply with regards to number of confirmatory clinical trials:

- If the MOA is the same for all indications, then likely one confirmatory trial would be needed in the most sensitive indication
- If MOA for the various indications is different, then likely more than one confirmatory trial would be needed

The FDA and EMA are typically open to considering new study designs for Phase III that differ from classical Phase III studies requiring very large numbers of patients to meet standard equivalence criteria. These new study designs could include dose response or time response type studies, which can require fewer patients and be of a shorter duration, if feasible.(Chow and Liu 2009) The choice of endpoints can also differ between the various regulatory agencies. It is therefore important to obtain agreement on endpoints prior to commencing clinical studies. There is also a procedure to have joint meetings with FDA and EMA for joint scientific advice, which should be considered.

Regulators do expect to see comparative immunogenicity data, this includes nature of immune response (anaphylaxis, neutralizing antibody, etc.), clinical relevance, and severity of consequences. It should be noted that the bioanalytical assay(s) should be capable of detecting antibodies against both the biosimilar and the reference molecule. Generally, FDA/EMA expect a clinical study design to establish statistical evidence that the proposed product is neither inferior nor superior to the reference product by more than a pre-specified margin (https://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf; http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf).

Allied to immunogenicity testing is the requirement by the FDA for a transition study. The aim of the transition study is to investigate any adverse effects that may occur in switching patients from the reference product to proposed biosimilar. The requirement for a transition study for an oncology product is generally waived. A transition study is not a requirement for EMA approval.

Allowance of interchangeability is another consideration for a biosimilar in the U.S. For this designation, the FDA requires a biosimilar to be "expected to produce the same clinical result as the reference product in any given patient." New FDA guidance on interchangeability—issued in January 2017 and entitled, 'Considerations in Demonstrating Interchangeability with a Reference Product Guidance for Industry: Draft Guidance' (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf)—helps clarify the agency's latest thinking in this area. The draft guidance recommends that sponsors wishing to have a biosimilar approved as interchangeable with a reference product perform one or more switching studies, designed to show that patients can alternate between the two products safely and without diminished efficacy.

FDA agreement on the interchangeability study design should be obtained prior to study initiation. If a biosimilar product gains approval as being interchangeable, it garners 1 year of marketing exclusivity and can be substituted for the innovator product at the pharmacy without authorization of the prescribing physician. In the EU, substitution is not permitted or controlled centrally and is a matter for national agencies to address.

The Pediatric Study Plan (PSP) has to be addressed for biosimilar products for the FDA but is not needed for biosimilar product approval in the EU. It should be noted that PSP is not needed in the U.S., if the sponsor is seeking interchangeability designation for their product (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM273001.pdf).

A sponsor may be able to obtain extrapolation to *other indications from a single clinical study conducted in the most sensitive indication* if the reference product's mechanism of action (MOA) is the same for all indications. The information needed for scientific justification for extrapolation should include (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf):

- Relevant target receptors for each indication
- Pattern of molecular signaling upon receptor binding
- Expression and location of target receptors
- Relevance of PD measures to MOA
- Relevance of PK values in different patient populations
- Differences seen in toxicities for the various conditions of use in the different patient populations.

Bridging Studies

A thorough review of regulatory requirements of the region/country of interest is crucial in determining what, if any, bridging studies may be needed. Based on the country or region of interest for potential marketing of the biosimilar, the sponsor should take into consideration the following:

- Is the reference product used in similarity studies acceptable to the additional region/country of interest? If not, determine what additional testing will be required.
- Does the region/country of interest require clinical data on their own population?
- Will the region/country of interest accept the similarity data generated to date as pivotal for product approval? If not, what additional CMC, nonclinical and clinical data is required.
- Have you had a scientific advice meeting with regulators of region/country of interest? If not, it is highly recommended that the sponsor seeks scientific advice and put forth robust scientific justifications for acceptance of the data that has already been generated.

- Additional considerations include whether the regulators require local manufacturing. If yes, potential partnership with a local manufacturer will be important to consider.

Pulling It All Together

Once you have collected all the required information for region/countries of interest for potential approval/marketing of your biosimilar, your regulatory strategy should take into consideration the following as you work on the overall strategic plan:

- What are the corporate goals for regulatory approval for the varies countries/regions of interest?
- What is the budget for activities to be undertaken?
- Do you need to co-develop the biosimilar with a partner, i.e., share the cost of clinical trials? If yes, consider all ramifications.
- Do you need to get the biosimilar product approved in a highly regulated market before it can be registered in emerging markets of interest?

Regulatory strategy is a living document and there can be numerous factors that may require shift in strategy over time, such as:

- Any changes in company business strategy
- Marketplace competition
- Changes in regulations for region/countries of interest
- Differences seen in biosimilarity data (CMC, Nonclinical or Clinical)
- Discussions with regulators
- Any additional testing that may be required and its impact on overall regulatory submissions/product approval timelines

Once all of the above considerations are addressed, a strategy can be outlined as to when what steps will occur with regards to manufacturing process development, biosimilarity testing, meetings with regulators, scale-up, nonclinical testing, clinical studies, marketing authorization filing in specific timeframes for countries/region of interest.

Conclusion

In conclusion, the design and implementation of successful regulatory strategies for the development of biosimilars requires a critical examination of numerous factors as outlined in this chapter. Understanding the commercial needs of your organization along with having robust competitive intelligence in hand it is possible to develop a robust streamlined regulatory strategy with the following in mind:

- Review of all available guidance documents relevant to the development of the biosimilar for countries/regions of interest.
- Use of stepwise approach for the development of the biosimilar.
- Ensuring the development of a robust CMC foundation with batch-to-batch consistency.
- Confirming physicochemical similarity of multiple batches of relevant reference product (over the shelf life of the reference product) and proposed biosimilar using validated, state-of-the-art analytical methods.
- Confirming similarity of the proposed biosimilar against the reference product(s) of choice by non-clinical testing [in vitro and in vivo (*if needed*)].
- Developing a streamlined CDP based on a quickest-to-market strategy for priority markets.
- Seeking scientific advice from regulators to obtain agreement on the CDP. This should occur once sufficient CMC and in vitro non-clinical studies have been completed to facilitate discussions, and for meaningful advice to be provided on the in vivo nonclinical and clinical development program that will be required.

All elements of the development of a biosimilar product are important, and regulators from highly regulated markets have indicated that to make a decision on the approvability of a biosimilar, they will rely on the review of the “totality of the evidence” data presented.

References

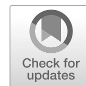
Chow S-C, Liu J-p. Statistical assessment of biosimilar products. J Biopharm Stat. 2009;20(1):10–30.

Cox TM. Gaucher disease: clinical profile and therapeutic developments. Biologics. 2010;4:299–313.

Krishnan E, Lingala B, Bruce B, Fries JF. Disability in rheumatoid arthritis in the era of biological treatments. Ann Rheum Dis. 2012;71:213–8.

US Food and Drug Administration. FDA 101: Regulating biological products. US Food and Drug Administration. August 20, 2012. Available at: http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm048341.htm. Accessed 5 Dec 2005.

Chapter 6
Clinical Information Requirements for Biosimilar Biologic Drug Approvals in Canada



Bradley J. Scott and Jian Wang

Abstract Biologic medicines have been invaluable in the treatment of conditions ranging from cancer to autoimmune disease. The expiry of patents for many successful biologic medicines has fostered a high degree of interest in the development of subsequent entry versions, known as biosimilar biologic drugs or, simply, biosimilars. Biologics are structurally complex and more difficult to replicate than small molecule pharmaceuticals. Health Canada's guidance on biosimilar biologic drugs specifies rigorous comparisons of chemistry and manufacturing attributes, non-clinical aspects, as well as clinical studies comparing the pharmacokinetics/pharmacodynamics, safety and efficacy, and immunogenicity of a biosimilar biologic drug to its reference. When taken as a whole, evidence of similarity in each of these aspects allows for a conclusion of biosimilarity, meaning that, in any given setting, the biosimilar is expected to produce clinical outcomes that are not meaningfully different from those expected with the reference biologic drug. This chapter is intended to present the Canadian clinical information requirements for biosimilars, and how they relate to the Canadian regulatory framework, while putting these requirements into context with other aspects of a complete data package in support of biosimilarity.

Keywords Biologic · Biosimilar · Canada · Clinical · Comparison · Extrapolation · Immunogenicity · Non-Clinical · Regulatory

Introduction

Biologic medicines ("biologics") have been invaluable in the treatment of diseases ranging from cancer to autoimmune disease. Take, for example, the anti-CD20 monoclonal antibody, rituximab, which has revolutionized the treatment of patients

B. J. Scott · J. Wang (✉)
Biologics and Genetic Therapies Directorate, Health Canada, Ottawa, ON, Canada
e-mail: Bradley.scott@canada.ca; jian.wang@canada.ca

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_6

with non-Hodgkin's lymphoma (NHL). Since its introduction in 1997, deaths due to NHL have decreased each year and continue to fall (NIH National Cancer Institute 2014). More recently, immuno-oncology therapies have emerged as a new treatment paradigm for patients with advanced cancers such as melanoma and lung cancer. Other biologic medicines, such as the filgrastims, epoetins and infliximab have played key roles in the management and treatment of patients with serious and often life-threatening disease.

Biosimilars are a relatively new class of biological medicine that are developed based on an already approved reference product. The development, authorization and marketing of biosimilars, through appropriate regulatory pathways, have the potential to improve the accessibility and affordability of biologic medicines. With regard to cost, it is notable that, in 2016, seven of the top ten highest grossing drugs in Canada were biologics, costing patients and payers nearly 3.1 billion Canadian dollars (Government of Canada 2017a).

The expiry of patents for many successful biologics, such as those previously mentioned, has promoted a high degree of interest in the development of subsequent entry versions, termed biosimilar biologic drugs or, simply, biosimilars. In 2013, data published concerning the number of biosimilars in development signaled a high level of activity in the field. For example, over 40 programs were reported to be underway for biosimilar versions to just three of the most successful biologics; the TNFα inhibitors known as Enbrel, Humira, and Remicade (Rader 2013).

Clearly, biosimilars represent a remarkable business opportunity for developers while offering patients and payers potentially significant cost savings. However, regardless of cost, regulators and manufacturers must ensure that biosimilar medicines provide patients with clinical outcomes that are not meaningfully different than what might be achieved through treatment with the original product (the innovator).

Notably, biosimilars are not generic biologics. Generic pharmaceuticals (small molecule drugs) can be characterized with a high degree of precision and are generally composed of a single chemical entity in addition to pharmaceutical excipients. In the case of oral dosage forms, bioequivalence studies are usually the only clinical studies required to demonstrate that formulation differences do not affect the extent or rate of absorption. Given identical active ingredients in addition to equivalent absorption characteristics, a generic pharmaceutical can be considered "bioequivalent" to the reference product. In contrast, biologics are structurally complex and are produced within living systems via intricate manufacturing processes. This method of production results in inherent within- and between- batch variability. Therefore, it is impossible for a biosimilar developer to make an identical copy of an originator biologic. In fact, because biologic medicines have within batch variability (i.e., they are often not composed of a single-chemical entity), and between batch variability, manufacturers and regulators take steps to ensure that variability is within pre-specified limits from batch to batch and after manufacturing and process changes.

The principles of assessing the comparability of biotechnological/biologic products before and after changes are made to the manufacturing processes are addressed

in the International Council on Harmonization (ICH) guidance document *Q5E*. It is from this guidance that many of the principles of biosimilarity assessment are derived (International Council on Harmonization 2004). Notably, there are situations for which the ICH Q5E guidance suggests that non-clinical and/or clinical studies may be necessary. For example, major formulation or manufacturing changes might warrant efficacy and safety studies to demonstrate that the changes are not detrimental to the safety and efficacy of the product. Given that biosimilar manufacturing processes always differ significantly in comparison to the innovators' manufacturing processes (the biosimilar manufacturer does not know the precise manufacturing process of the reference product), Canadian and international guidance on the development of biosimilars advises that manufacturers should provide comparative demonstrations of similarity in pharmacokinetics/pharmacodynamics (PK/PD), efficacy, safety and immunogenicity between the proposed biosimilar and its reference. In most cases, this requires the conduct of a comparative PK/PD study (or studies) and at least one comparative clinical trial in a representative indication. Clinical trials should be designed to be sensitive enough to detect differences between the biosimilar and its reference product.

This chapter is intended to provide discussion regarding the pathway to authorization for biosimilars in Canada with a focus on clinical aspects as outlined within Health Canada's recently updated guidance document: *Information and Submission Requirements for Biosimilar Biologic Drugs* (Health Canada 2016). By doing so, it is hoped that manufacturers, physicians, patients and payers will gain a better understanding of what it means to be approved as a biosimilar in Canada.

Regulatory Framework

In Canada, biologic drugs (or "biologics") are regulated under the *Food and Drugs Act and Regulations*. Specifically, biologics are products listed on *Schedule D* to *the Food and Drugs Act*. Technical requirements specific to biologic drugs are set forth in Part C, Division 4 of the *Food and Drug Regulations* and regulations governing the authorization and marketing of new drugs are set forth in Part C, Division 8. Other divisions, more generally applicable to all drug products, including biologics, are those pertaining to general requirements (Div. 1), establishment licensing (Div. 1A), good manufacturing practices (Div. 2) and clinical trials (Div. 5). It is notable that, as of 2018, there are no Canadian regulations that are specific to biosimilars. Biosimilar drug products are regulated by Health Canada as "new drugs" as defined in the *Food and Drugs Act Regulations, C.08.001*. In addition, because biosimilars enter the market subsequent to a biologic drug product that was previously authorized in Canada, they are subject to existing laws and regulations outlined in the *Patented Medicines (Notice of Compliance) Regulations*. Due to this, a biosimilar product cannot receive a Notice of Compliance (NOC) if it is deemed that authorization would cause the biosimilar to infringe upon existing patents protecting the reference product.

Recognizing that biosimilars require a unique drug development paradigm in comparison to innovative biologics (and generics), Health Canada first published guidance on the topic in 2010. Since then, eleven biosimilar biologic drugs have been issued a Notice of Compliance in Canada (Table 6.1). A revised version of the guidance was published in 2016 (Health Canada 2016). In an effort to provide transparency to stakeholders, the guidance document: *Submission and Information Requirements for Biosimilar Biologic Drugs* sets forth the various elements that are expected to be provided within a *New Drug Submission* intended to support a biosimilar biologic drug. These elements are discussed further in this chapter. In addition, since the guidance stresses the use of “step-wise” development paradigm, sponsors may also find the guidance document useful as a guide to aid in the development of a biosimilar biologic drug.

Reference Biologic Drug

In order for a sponsor to be able to file an NDS in support of a biosimilar, a suitable reference biologic drug (reference product or “the reference”) must have been authorized, in Canada, on the basis of a full data package consisting of extensive chemistry, manufacturing, non-clinical and clinical data. That is, a biosimilar must be “subsequent to a biologic drug that is authorized in Canada and to which reference is made”.

Recognizing the difficulties associated with choosing a reference product that satisfies the needs of various national regulatory authorities, Health Canada has determined that sponsors may use a non-Canadian sourced version of the reference product in studies intended to demonstrate biosimilarity; however, the onus is on the sponsor to demonstrate that the chosen reference product contains the same medicinal ingredient and is a suitable proxy for the Canadian product. The demonstration of reference suitability can usually be based on an assessment of product characteristics such as dosage form, and route of administration, all of which should be the same as the product that is authorized in Canada. Sponsors should pay particular attention to product concentration when reference products are sourced from outside Canada. Occasionally, reference products sourced from outside Canada may be available in multiple strengths, some of which may be of a different concentration than the Canadian version, and, for various reasons, might not be authorized for use in Canada. It is recommended that sponsors seek out a reference product that has the same concentration as the version available in Canada. By doing so, the sponsor avoids potential concerns regarding differences in absorption and subsequently, safety and efficacy (compared to the Canadian authorized product). Thus, it is important to keep in mind that the chosen reference acts as a surrogate for the Canadian authorized product, and therefore, should have characteristics similar to those of the Canadian authorized product.

Often, sponsors may wish to use the same reference product sourced from more than one jurisdiction in order to satisfy various national regulatory authorities

Table 6.1 Listing of biosimilars approved in Canada as of December 2017

Biosimilar brand name (INN)	Reference biologic drug (brand name)	Therapeutic area (s)	Date of Canadian approval (notice of compliance)[a]
Omnitrope (somatotropin)	Genotropin	Growth hormone deficiency in children and adult growth hormone deficiency	April 20, 2009
Omnitrope (somatotropin)	Genotropin	Additional indications for small for gestational age, idiopathic short stature and Turner's syndrome	May 8, 2015
Inflectra (infliximab)	Remicade	Rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and plaque psoriasis	January 15, 2014
Inflectra (infliximab)	Remicade	Additional indications for adult Crohn's disease, fistulising Crohn's disease and adult ulcerative colitis	June 10, 2016
Remsima (infliximab)	Remicade	Rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and plaque psoriasis	January 15, 2014
Remsima (infliximab)	Remicade	Additional indications for adult Crohn's disease, fistulising Crohn's disease and adult ulcerative colitis	August 5, 2016
Basaglar (insulin glargine)	Lantus	Treatment of pediatric (>6 years) and adult patients with type 1 diabetes mellitus and adult patients with type 2 diabetes mellitus	September 1, 2015
Grastofil (filgrastim)	Neupogen	Prevention or treatment of neutropenia	December 7, 2015
Brenzys (etanercept)	Enbrel	Rheumatoid arthritis, ankylosing spondylitis	August 31, 2016
Erelzi (etanercept)	Enbrel	Rheumatoid arthritis, Polyarticular juvenile idiopathic arthritis and Ankylosing spondylitis	April 6 2017
Admelog (insulin lispro)	Humalog	Treatment of patients with diabetes mellitus, initial stabilization of diabetes mellitus	November 16, 2017
Renflexis (infliximab)	Remicade	Rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, plaque psoriasis, Crohn's disease, fistulising Crohn's disease, and ulcerative colitis	December 1, 2017
Lapelga (pegfilgrastim)	Neulasta	Supportive therapy for patients with non-myeloid malignancies – prevention of febrile neutropenia and infection	April 5, 2018
Mvasi (bevacizumab)	Avastin	Metastatic colorectal cancer; locally advanced, metastatic or recurrent non-small cell lung cancer	April 30, 2018

INN international non-proprietary name

[a]Notice of Compliance dates were obtained from Health Canada's Notice of Compliance database available at: https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/notice-compliance/database.html

such as the European Medicines Agency (EMA) and the U.S. Food and Drug Administration (FDA). Health Canada prefers that sponsors choose one reference product to use throughout the development of a biosimilar; however, flexible approaches have been considered and two reference products may be acceptable if information is provided to demonstrate that they are similar with respect to chemistry and manufacturing and, if combined in the reference arm of a clinical trial, they are demonstrated to meet comparative bioavailability standards. Finally, it is noted that the chosen reference product(s) should be sourced from Canada or, preferably, from a jurisdiction that has formally adopted International Council for Harmonization guidelines and that has regulatory standards and principles for the evaluation of medicines, post-market surveillance activities, and approaches to comparability that are similar to Canada.

Development Paradigm

The development of an innovative biologic medicine involves thorough characterization of the protein's physicochemical properties and biological activities. Characterisation includes, but is not limited to, assessments of primary, secondary, tertiary and quaternary structure as applicable. Post-translational modifications are characterized as these can often be critical attributes that influence the functionality of the protein; for example, effector mechanisms associated with monoclonal antibodies. It is also important to determine the biological activity of the product, in terms of receptor binding and pharmacodynamic effect. Finally, the properties of the drug product, such as the formation of aggregates and product stability must be defined.

Beyond structural and functional characterization, sponsors of innovative new drugs must follow the appropriate ICH guidance with respect to the pre-clinical characterization of safety, and safety pharmacology. Investigations involve testing in a relevant animal species (i.e., one that expresses the appropriate receptor) with the intention of identifying safe initial doses for clinical trials and target organ toxicity and reversibility. For biologics, the most important ICH safety guideline is ICH S6(R1), which clarifies the breadth of pre-clinical safety studies that are necessary to support regulatory approvals for biotechnology-derived pharmaceuticals. It is well accepted that a number of the pre-clinical safety investigations that are necessary for small molecule pharmaceuticals are not usually required for products of biotechnology (e.g., genotoxicity and carcinogenicity studies) (International Council on Harmonisation 1997).

Finally, for an innovative biological, it is notable that the bulk of the time spent generating data to support regulatory approval is often focused on clinical studies. Safety and efficacy must be established for each indication and clinical use for which regulatory approval is sought. Establishing safety and efficacy involves various clinical studies including early phase trials, to establish safe and effective doses, as well as confirmatory studies demonstrating the efficacy and safety of the

product for each new clinical use. Confirmatory studies may involve thousands of patients, depending on the target population, that need to be treated and followed for a significant length of time.

The development paradigm in place for innovative biological medicines places a large emphasis on clinical evidence to demonstrate the safety and efficacy of the product in each indication. In contrast, for biosimilars, the emphasis is placed on the demonstration of similarity to a reference, which, if achieved, allows the biosimilar to leverage prior information (i.e., clinical safety and efficacy information) known about the reference.

Demonstrating similarity to a biologic reference requires a wide range of comparative side-by-side studies between the potential biosimilar and a reference product. While development of an innovative biologic will usually focus heavily on clinical studies, biosimilar developers are most concerned with demonstrating similarity at the physicochemical and biological level (i.e., the quality comparison) using a variety of analytical techniques. To do so involves the side-by-side comparison of numerous attributes, the most important of which are known as “critical quality attributes” or CQAs. Examples of such CQAs may include: primary amino-acid sequence, protein folding, subunit interactions, size heterogeneity, glycosylation, bioactivity, aggregation, impurities and others. These attributes can be compared using a variety of methodologies. For example, mass spectrometry and peptide mapping are useful in the determination of primary sequence while surface plasmon resonance and cellular bioassays are often useful in comparing receptor binding kinetics and bioactivity, respectively. Important to note is that the manufacturer of a biosimilar is essentially attempting to reverse engineer the innovator’s product. Thus, the biosimilar product is not a “me too” biologic that is designed to simply engage the same target as the reference product while having a significantly different structure or composition. It is an intentional “copy” of the reference product that is designed to be as close to identical to the reference product as possible; however, this comes with the caveat that producing identical biologics is impossible due to the inherent variability of biological systems and their sensitivity to differences in manufacturing processes. Therefore, biosimilar manufacturers and regulators have a responsibility to thoroughly characterize and assess any differences that might exist between a potential biosimilar and its reference and to determine whether those differences have a meaningful impact on the safety and/or efficacy of the biosimilar in comparison to the reference.

Given a successful demonstration of biosimilarity from the quality perspective, certain clinical studies are advisable in order to address the residual uncertainty that remains in light of the limitations of analytical methodologies and the subtle differences that may be identified in a thorough side-by-side comparative quality program. The remainder of this chapter will discuss the various types of non-clinical (animal) and clinical studies that may be required to support biosimilar approval, their purpose, and best practices for their conduct. For a more thorough discussion of Health Canada’s view on the importance of comparative quality data, as well as the employed methodologies, the interested reader is referred to Wang et al. (2017).

Non-Clinical Development

Early guidance from Health Canada (and other national regulatory agencies) on the development of biosimilars, indicated the need for non-clinical (i.e., animal toxicology) studies comparing the biosimilar to the reference product with the intention of identifying differences in toxicokinetics and/or other toxicological parameters; however, as experience has been gained in the review and approval of biosimilars, both in Canada and around the world, comparative non-clinical studies have sometimes been deemed unnecessary. This is due to several factors. Particularly, comparative data obtained from these studies, is difficult to interpret due to the low number of animals usually included within any given dosing group. The low power of these studies makes it difficult to draw conclusions when and if differences are noted between the biosimilar and the reference. On balance, it is also difficult to interpret whether a lack of identified differences is indicative of biosimilarity. Finally, for many biologic products, there can be difficulty in identifying a relevant species from which meaningful comparative data can be obtained (Ryan 2015). For these reasons, Health Canada has decided to include flexibility with respect to the need for comparative non-clinical studies in its guidance on biosimilars. It is, however, important to consider whether there is a need to conduct non-clinical safety studies, which may not necessarily incorporate a reference product treatment group, to support first-in-human studies with proposed biosimilar products, particularly if new excipients are used in the biosimilar product for which the toxicological profile is unknown.

Clinical Development

The development of a biosimilar product should follow a step-wise progression. Therefore, it is usually only appropriate to initiate head-to-head clinical studies once an intended biosimilar product has been shown to be similar to an appropriate reference product via extensive side-by-side analytical and pre-clinical comparisons. Once a manufacturer is satisfied that these aspects of biosimilarity have been demonstrated, clinical studies should usually include side-by-side comparisons of pharmacokinetics, pharmacodynamics (if applicable), immunogenicity and safety and efficacy as discussed in the following sections.

Pharmacokinetic Studies

The pharmacokinetic (PK) profiles of biologic drugs are dependent on many factors including product specific characteristics. Therefore, small differences in the quality attributes of an intended biosimilar in comparison to a reference product

may potentially lead to differences in drug absorption, distribution, metabolism or excretion (Putnam et al. 2010).

The role of comparative clinical pharmacokinetic studies in the assessment of biosimilarity is to rule out unacceptable PK differences that could indicate the presence of structural and/or functional differences that, in turn, could impact on the efficacy, safety or immunogenicity of the product. In order to exclude such PK differences, the manufacturers of biosimilars should design comparative PK studies to be conducted in a setting that is sensitive to change. The type of study performed, whether a cross-over or parallel group study, will depend on the biological's known PK characteristics. Cross-over studies will often not be feasible due to the long half-lives associated with many biologic products, particularly monoclonal antibodies. For biologics with relatively short half-lives, such as insulins, filgrastims or certain fusion proteins, cross-over studies are preferred. Healthy volunteers will usually be an appropriate population in which to compare the PK properties of biologics (when feasible) because, as a whole, they are considered to be a homogenous and sensitive population. On the other hand, a healthy-volunteer population may not always adequately reflect the PK parameters in the patient population for which the reference product is indicated. For example, the immunocompetence of healthy patients versus those with disease may impact on the pharmacokinetics or pharmacodynamics. Also, host factors such as receptor expression, receptor internalization rate, and patient status may affect the disposition and clearance of biologics (e.g., target mediated drug-disposition). Therefore, the Canadian biosimilar guidance recommends that comparative PK studies be carried out in a relevant patient population when the PK or PK/PD of the reference is known to be substantially altered by the disease state(s) for which approval will be sought. Additionally, testing as monotherapy is preferable if it is feasible. Concomitant medications, such as those that are immunosuppressant, may affect the PK of the products and could mask differences between the reference and the intended biosimilar. In particular, monoclonal antibodies (mAbs) used in the treatment of cancer would preferably be evaluated in patients receiving first-line therapy to reduce the heterogeneity of the population and the effect of prior therapies, both of which might impact on the PK profiles of the test and reference products.

Route of administration is an important factor to consider in the design and conduct of comparative PK studies, and the use of a route that requires an absorption step is preferred. Where the route of administration necessitates an absorption step, standard bioequivalence type testing should be applied to assess PK similarity; for example, in determining the PK similarity of a biologic that is usually administered subcutaneously. However, many biologics are administered only via intravenous injection or infusion. Therefore, no absorption step occurs and bioavailability is 100% by definition. In this situation, simply comparing Cmax and AUC using methods accepted for the determination of bioequivalence may not be adequate to assess the PK similarity of an intended biosimilar and its reference. In these situations, analyses of additional PK parameters that reflect the distribution and clearance of the products, such as the $T_{1/2}$, K_e and Cl, may be useful.

Regarding bioanalytical methodologies: the methodology employed to determine serum concentrations from volunteer/patient samples is critical. Methods must be validated and fit-for-purpose. Health Canada has adopted the European Medicines Agency's *Guideline on bioanalytical method validation* and expects that assays used to support a biosimilar drug submission are validated in line with the guideline (Committee for Medicinal Products for Human Use (CHMP) 2011). Notably, most assays used in the detection of biologicals in patient samples are of the ligand binding type. These assays should be fully validated paying special attention to the special challenges posed by ligand-binding assays, such as the use of appropriate reference standards, specificity of binding reagents and the selectivity of the assay in a given matrix.

Pharmacodynamic Studies

Studies that comparatively assess the pharmacodynamics (PD) of a biosimilar and reference biologic provide very useful information relating to the assessment of biosimilarity. Some examples of products for which reliable PD markers exist include filgrastim, peg-filgrastim and insulin. For the filgrastims, it is possible to use the absolute neutrophil count (ANC) over time to compare the PD effect of administering filgrastim or peg-filgrastim. ANC can be plotted vs. time and the area under the ANC (ANC-AUC) curves can be compared (Desai et al. 2016; Waller et al. 2010). In the case of insulin, a euglycemic clamp study is a useful way to compare the extent of blood-glucose control offered by two insulin preparations. It is noted euglycemic clamp studies are technically challenging and should only be conducted at centres that specialise, or have expertise, in this type of study (Heise et al. 2016).

For products with a reliable PD marker, a high-quality PD study (usually combined with PK) may be more sensitive than an efficacy study in terms of detecting differences between the biosimilar and the reference. However, it is recognized that, for many biologics (e.g., mAbs), suitable pharmacodynamic markers have not been identified. In these instances, sponsors must rely more heavily on the comparison of pharmacokinetic profiles and specific PK parameters in combination with clinical safety and efficacy trials in order to complete a demonstration of biosimilarity.

Another point to consider in the conduct of comparative PK/PD studies for biosimilars is the selection of a suitable dose. The dose that is approved for use in the patient population may not always be the most appropriate dose to use in a comparison that should be designed to maximise the probability of detecting a difference. The use of a dose that induces a maximal PD response may hinder the ability of a study to detect differences in a selected PD endpoint. Therefore, it may sometimes be appropriate to use a low or sub-therapeutic dose residing on the linear part of the dose response curve in order to improve the sensitivity of the comparative PD study (i.e., sensitivity to differences between products). In addition, using a lower-than-therapeutic dose may be desirable when a comparative PD study

is to be carried out among healthy volunteers. This is for several reasons; including the need to limit the occurrence of dose-related side effects, but also because a PD ceiling effect might be elicited at lower doses in healthy volunteers than in patients which could have the effect of masking differences in the PD response if the higher dose were to be used.

Safety and Efficacy

Safety

The safety assessment for a biosimilar should involve a descriptive comparison of the overall adverse event profile as well as the types and severity of specific adverse drug reactions (ADR) occurring after the initiation of treatment. However, an assessment of this type is unlikely to be capable of detecting differences in the incidence or severity of what are often rare events to begin with. Therefore, it is also useful to compare the types and severity of adverse events/reactions to those that have been observed throughout the reference product's life-cycle in order to determine whether the biosimilar has elicited new safety signals. Again, choosing a patient population that enhances the likelihood of detecting a difference is critical to the assessment of clinical differences. This may be difficult to do for the assessment of safety; however, testing the products side-by-side in a monotherapy setting would be considered one step that could benefit the safety assessment since the safety profile would not be confounded by the use of concomitant therapies. Furthermore, testing in a relatively homogeneous population may increase the ability to detect differences in safety by reducing contributing factors such as the use of prior therapies. A relatively homogenous study population may also reduce confounding that could occur due to the use of concomitant medications and/or the presence of concomitant conditions. In general, detecting meaningful differences between safety profiles is likely to be difficult. For this reason, appropriate risk management plans and post-marketing surveillance, as required for all new drugs, are critical to the further strengthening of the safety database. With respect to biosimilar labelling, it is assumed that the biosimilar product poses the same risks to patients as the reference product. This is true even when particular adverse drug reactions have not yet been observed in subjects administered the biosimilar. Therefore, warnings and precautions (and guidance for ADR management) present in the labelling of the reference product should usually be included in the labelling of the biosimilar.

Efficacy

The comparative assessment of efficacy is a key component of the clinical assessment for biosimilars, particularly when no reliable PD marker is available. Accordingly, extensive forethought should be given to the setting in which clinical

comparability is to be tested. When choosing the clinical study model, it should be ensured that the model interrogates the relevant mechanism(s) of action considering all of the indications that will be sought for approval. It is known that some biologics can function through multiple mechanisms of action, and the mechanisms involved with treating one disease may differ from those involved with another.

Study sensitivity (i.e., the ability to detect a difference) is of critical importance to biosimilar development. To maximise the sensitivity of a clinical efficacy study, investigators should choose both populations and endpoints that, in combination, will be sensitive to differences that may impact efficacy (and/or safety). To do so, sponsors should perform a thorough review of the available clinical data for the reference product in order to determine the population-endpoint combination that is associated with both a large effect size and a robust historical dataset. By performing a thorough systematic review of studies performed with the reference product, the biosimilar manufacturer can identify critical pieces of information such as the magnitude of effect and the timing of response that are necessary to guide study design and establish clinically meaningful similarity margins. Notably, the study design will be dependent on many factors, which may not always lead to the selection of the most commonly studied disease and/or population. Similarly, the primary endpoint(s) may not be the one typically recommended for authorization of an innovative new biologic, but should be an endpoint that will provide adequate sensitivity to detect differences in efficacy. For example, there are several endpoints commonly used to assess the efficacy of biologic products in the setting of rheumatoid arthritis (RA) (Hobbs and Cohen 2012). A 20% reduction in symptoms, based on American College of Rheumatology response (ACR20) criteria, has historically been the most commonly used in determining the efficacy of an innovative new product; however, it is a dichotomous endpoint that does not provide information on the similarity of response at different magnitudes. On the other hand, a continuous endpoint, such as the disease activity score (DAS28), could be considered more appropriate in comparing two RA treatments for similarity. As another example, overall survival (OS) is considered the gold standard for the demonstration of efficacy for an innovative new product in the oncology setting; however, measuring OS can require lengthy follow-up and can be difficult due to confounding factors such as the receipt of new treatments. Therefore, comparing endpoints that read out earlier, such as response rate or progression free survival may be more appropriate in some oncology settings. Important factors that will affect the choices of study population and study endpoints include the timing of effect, the expected magnitude of effect, the homogeneity of the population with respect to baseline characteristics, the use of concomitant medications and the duration of treatment and follow-up.

Studies to compare the efficacy of biosimilars should usually be designed and powered to test a hypothesis of equivalence, should always be protected by randomisation, and should be double-blinded whenever possible. The choice of equivalence margins should be pre-defined and supported by statistical estimation based on historical data available for the reference product and by a comparison of the current and prior study designs. Pre-defined margins should exclude differences

in efficacy that would be considered as clinically meaningful. Although non-inferiority designs may sometimes be considered, their use is discouraged in the Canadian biosimilar guidance since an indication of superiority in a non-inferiority trial could lead to new questions that could lead to a rejection of the claimed biosimilarity.

Recently, some biosimilar manufacturers have included planned switches, between the biosimilar and the reference, in the comparative phase 3 studies for biosimilar biologic drugs. Health Canada does not require a switch study to be submitted to support the approval of a biosimilar biologic drug. Switch studies may provide additional information on the safety of switching from an innovator's biologic to the biosimilar product. In addition to manufacturers, governments and other interested parties have conducted their own switching studies to further support the use of biosimilar products by physicians and patients (Moots et al. 2017). These studies have been undertaken largely to allay the concerns of physicians and patients regarding switching from an innovative biologic to a biosimilar version, particularly for patients that have achieved stable disease control on an innovator's product. Although switching may provide useful information, it is important to consider how switching enrolled patients from one treatment to the other may affect the ability of the study to provide interpretable results with respect to safety and efficacy. Certainly, if the study to be used to support regulatory approval incorporates a switch, the primary efficacy endpoint should be read out before any patients are switched from the biosimilar to the reference product; however, sufficient numbers of patients should also remain on their originally randomized treatment to enable meaningful comparisons of safety and key secondary efficacy endpoints.

Immunogenicity

Immunogenic responses can have consequences that impact on safety, efficacy and pharmacokinetics of biologicals. Indeed, even minor changes to a product can have major unintended consequences. For example, formulation changes to a product containing epoetin alfa resulted in a dramatic rise in the number of cases of pure red cell aplasia in chronic kidney disease patients due to the generation of neutralizing antibodies that cross-reacted with endogenous protein (Casadevall et al. 2002). With respect to mAbs, the generation of anti-drug antibodies (ADA) has been associated with severe acute infusion reactions making mAb immunogenicity a key concern for patients (Steenholdt et al. 2011). In addition, ADAs are known to interfere with the efficacy of biologic drugs such as the anti-TNF monoclonal antibodies that are useful in the treatment of a number of auto-immune diseases (Kalden and Schulze-Koops 2017). The complexity of some biologics is such that differences in post-translational modifications, folding and conformations could lead to differences in the tendency of the product to elicit immunogenicity (Maas et al. 2007; Jahn and Schneider 2009). Thus, since no two biologic drugs can

be completely identical, it is essential that the developers of intended biosimilars assess the formation of anti-drug antibodies (ADA) in comparative clinical trials to determine whether minor molecular differences might lead to differences in the immune response and, subsequently, affect pharmacology, safety and/or efficacy.

Designing clinical studies to assess immunogenic differences can be challenging. It is important to keep in mind that patient specific factors can play a role in immunogenicity. For example, the underlying disease may influence the rate of ADA against a particular biologic. For infliximab, reported ADA rates range from 7 to 61% in patients with psoriasis, ankylosing spondylitis, Crohn's disease or rheumatoid arthritis (Ducourau et al. 2011; Perdriger 2009). However, it is difficult to compare ADA rates across studies due to differences and advances in the methodology used to detect ADA. Therefore, these differences should be interpreted with caution. Genetic factors, age, concomitant medication, duration and route of administration, and previous exposure to similar products are also important to consider as each of these may affect a patient's risk of developing anti-drug antibodies (Brinks 2013). All of these factors should be considered in order to maximize the likelihood that a study would be able to detect clinically significant differences in the immunogenic response. Also, factors that suppress the immunogenic response should be minimized to the extent possible. As indicated previously, an important aspect to consider is the use of concomitant immunosuppressive medications. For some biologic medicines, it is known that immunogenicity is observed more frequently among patients that receive monotherapy. For example, infliximab administered to patients with rheumatoid arthritis has been shown to induce higher titres of ADA when administered without methotrexate (Pascual-Salcedo et al. 2011). Therefore, it is probable that the most sensitive population to detect differences in the immunogenic response is one in which immunogenicity is not suppressed by concomitant therapies. As well, it has been suggested that the subcutaneous route of administration is associated with a stronger immunogenic response than the intravenous route (Brennan et al. 2010). Therefore, investigations of immunogenicity should be performed in studies that administer the product as monotherapy and/or via the subcutaneous route if possible. However, the feasibility of these recommendations will depend on the clinical indications and uses that are authorized for the reference product, and, in general, the product should be administered via a route of administration that is approved for the reference product in question.

Assay development is a crucial part of the strategy to determine similarity in immunogenicity. The assays employed to detect ADA should be considered state-of-the-art and may not necessarily be the same as those performed by the innovator at the time the reference product was developed. This may require biosimilar manufacturers to develop new assays that are capable of detecting differences in the ADA response between the biosimilar and the reference product. The EMA's Guideline on immunogenicity assessment of monoclonal antibodies intended for in vivo clinical use provides an excellent discussion of the problems inherent in immunoassays used to measure antibodies against mAbs and, in combination with the Guideline for similar biological medicinal products containing biotechnology-

derived proteins as active substances: non-clinical and clinical issues, provides recommendations on the types of assays that are useful for comparatively assessing the development of neutralizing and non-neutralizing ADA (Committee for Medicinal Products for Human Use (CHMP) 2012a, b). Briefly, immunogenicity testing should be conducted using a tiered approach that involves screening assays, confirmatory assays and assays to determine whether binding ADA are neutralizing. An important consideration in the assessment of immunogenicity is the use of an appropriate capture ligand. Ideally, two assays (one using biosimilar and one using reference product as a capture ligand) would be validated and employed in parallel in order to and to avoid false negatives. However, if only one assay is to be used, it would preferably incorporate the biosimilar drug as the capture ligand since this should provide a more conservative comparison of immunogenicity between the biosimilar and the reference (Cai et al. 2013). Finally, it is critical that developers ensure that immunogenicity assays have adequate sensitivity and are sufficiently tolerant of residual drug to enable the detection of ADA even at relatively low concentrations.

The comparative assessment of immunogenicity, between a biosimilar and its reference product, should involve more than simple comparisons of the incidences of anti-drug antibodies between biosimilar and reference product treatment arms. The assessment of ADA titers (magnitude) and titer distributions among the population are important aspects that should be compared. Assessments of the ADA response, time-course of ADA development and ADA persistence are also important aspects of immunogenicity to be explored. Additionally, since the generation of ADA against mAbs may have significant effects on safety, PK and efficacy, their influence on these aspects should be examined and comparatively assessed.

Authorizing Indications (Extrapolation)

Importantly, Canadian guidance regarding biosimlars emphasizes the need for extensive side-by-side chemistry and manufacturing comparisons. By laying the foundation for the demonstration of biosimilarity, this extensive comparative characterization of critical quality attributes, including assessments of biological function, facilitates a reduced clinical data package. The clinical components of the demonstration of biosimilarity, discussed above, are required to address residual uncertainty that exists due to the fact that minor differences will always exist between a biosimilar and its reference biologic drug, and evidence should be provided that these differences are unlikely to impact clinical outcomes. When taken together, demonstrations of similarity in each of these aspects (chemistry and manufacturing, biological function, clinical outcomes etc.) establish a scientific bridge that allows the biosimilar to draw on the established safety and efficacy of the reference biologic, which may be approved for indications beyond those studied during biosimilar development. By drawing on this established safety and efficacy information a biosimilar can support its own approval for each of the indications

held by the reference product without the need to re-establish safety and efficacy for each indication. As follows, a product that has demonstrated biosimilarity is eligible to seek approval for each indication held by the Canadian version of the reference biologic drug with the expectation that clinical outcomes will not differ meaningfully whether a patient receives the biosimilar or its reference. Similar to other jurisdictions, the sponsors of biosimilars are also asked to provide rationales justifying the authorization of clinical indications for which the biosimilar has not been explicitly studied. Rationales should discuss how disease specific characteristics, in aspects such as the critical mechanism of action, pharmacokinetics/pharmacodynamics, disease pathophysiology, safety and immunogenicity, might impact the biosimilar in comparison to the reference product. Sponsors should determine whether there could be concerns about biosimilarity in a particular indication that were not addressed within the context of the biosimilarity development program.

Patent Hold: A Unique Canadian Issue

As mentioned briefly in section "Regulatory Framework", biosimilar biologic drugs are subject to regulation under the *Patented Medicines (Notice of Compliance) Regulations (PM(NOC) Regulations)* because they enter the market subsequent to an innovative drug. The PM(NOC) regulations are administered by the Office of Patented Medicines and Liaison (OPML), within the Therapeutic Products Directorate, on behalf of Health Canada. Health Canada has published guidance, titled: *Guidance Document: Patented Medicines (Notice of Compliance) Regulations*, to aid sponsors in understanding and navigating the PM(NOC) Regulations (Health Canada 2012). The guidance states specifically that:

> Under the PM(NOC) Regulations, a subsequent manufacturer seeking to copy a patented innovative drug is required to address the patents listed on the Patent Register against that innovative drug. The subsequent manufacturer may either agree to wait for expiry of the patent before receiving its notice of compliance (NOC) or challenge the patent by making an allegation justifying the issuance of the NOC. The allegation may be accepted by the innovator or upheld through a Federal Court decision.

Due to the requirements of the PM(NOC) Regulations, Health Canada cannot issue an NOC to a biosimilar when there are patents listed on the patent register that have not been addressed by the subsequent entry product's sponsor; either through the agreement of the innovator or a court decision. When this occurs, a biosimilar product that has undergone review and has been recommended for approval will be placed on "patent hold". These patents may be in respect of a variety of aspects including specific clinical indications. Therefore, a biosimilar sponsor that files an NDS in respect of several indications held by a reference product may not be granted an NOC if one or more of the applied for indications is protected by a patent on the Patent Register. To prevent the blocking of an NOC for their product, the sponsors of biosimilars can limit the breadth of indications for which they request authorization

within an NDS for a biosimilar. In many cases (but not all), the application of this strategy accounts for the differences between the indications authorized for a biosimilar compared to the indications that have been authorized for the reference innovative drug.

Switching and Interchangeability

The ability to either switch or interchange a biosimilar for its reference is a key issue among biosimilar stakeholders. Physicians and patients may be reluctant to switch from an innovator's product to a biosimilar product when the disease being treated is adequately controlled. Because biosimilars have minor differences compared to their reference, patients and physicians alike may be concerned that switching to a new product could cause a loss of disease control or new adverse drug reactions. On the other hand, payers, both private and public, are likely to want to increase the use of biosimilar products in place of innovative products in order to realize substantial cost savings.

Interchangeability

One important aspect of this debate is the terminology used. It is important to ensure that there is a common understanding of what it means to be designated as "interchangeable". In Canada, various provincial legislations address interchangeability as it relates to an innovative drug and a generic copy. While there are some differences from jurisdiction to jurisdiction, "interchangeability" generally refers to the status granted to a generic pharmaceutical, and sometimes a therapeutic equivalent, that allows for the generic to be considered interchangeable for the brand name product at the pharmacy level. For further clarification, this means that the dispensing health care professional can legally substitute an interchangeable product for the brand name product that the physician has prescribed. This practice is often known as "generic or automatic substitution". For generic small molecule pharmaceuticals, provincial governments rely on Health Canada's Notice of Compliance (NOC) for products approved through the Abbreviated New Drug Submission (ANDS) pathway as evidence that a generic pharmaceutical can be declared as interchangeable with its Canadian reference product. The NOC granted for an ANDS constitutes a declaration of equivalence with a Canadian reference product (CRP).

When Health Canada issues an NOC for an ANDS, it does so based on several criteria. Most importantly, the *Food and Drug Regulations, C.08.002.1(1)(a),* state that the new drug (i.e., the generic) must be the "*pharmaceutical equivalent of the Canadian reference product*". The *Food and Drug Regulations* define *pharmaceutical equivalent* as meaning "*a new drug that, in comparison with another drug,*

contains identical amounts of the identical medicinal ingredients . . . " As has been discussed, biologic drugs are inherently variable and cannot be copied in an identical manner. Therefore, a biosimilar cannot be considered pharmaceutically equivalent to its reference, as per the definition in the regulations, and cannot be filed to Health Canada as an ANDS. Thus, since biosimilars are not eligible for approval through the ANDS pathway, Health Canada's issuance of an NOC for a biosimilar does not constitute a declaration of equivalence. Furthermore, Health Canada has previously clarified that the authorization (NOC) granted to a biosimilar is "not a declaration of equivalence" (Government of Canada 2017b).

Currently, no Canadian jurisdiction has implemented legislation or guidelines declaring that biosimilars are, or will be, interchangeable with their reference product. However, it is noted that the provinces have the authority to decide whether they want to use biosimilars interchangeably with their innovative counterparts (Klein et al. 2017). However, the decision to do so would have to be made without a declaration of equivalence from Health Canada since Health Canada currently has no legal framework that would allow such a declaration to be made. This is in contrast to the situation in the United States where legislation exists that allows for the FDA to designate a biosimilar product as interchangeable with its reference product (i.e., eligible for automatic substitution without the consent of the prescriber) (U.S. Food and Drug Administration 2017). However, no biosimilar has, at this time, been approved as an interchangeable biosimilar in the United States. In fact, the US FDA has only recently issued draft guidance on the matter (January 2017). Nevertheless, it is expected that the finalization of the FDA's *Guidance for Industry: Considerations in Demonstrating Interchangeability with a Reference Product* will provide clarity to manufacturers wishing to pursue interchangeable status in the United States. In Europe, the situation differs from both the United States and Canada. Notably, in a fairly recent paper authored by European regulators, interchangeability was defined differently than it has been described here. In their paper, the authors defined interchangeability as "the medical practice of changing one medicine for another that is expected to achieve the same clinical effect in a given clinical setting and in any patient on the initiative, or with the agreement of, the prescriber" (Kurki et al. 2017). As was previously described, interchangeability, as the term is used in Canada (and the United States) refers to the practice of automatic substitution without the consent of the prescriber. On the other hand, Health Canada has used the term switching to refer to the practice described by the European regulators as "interchangeability" (quoted above). Switching will be described in more detail below. Due to these differences in defining various terms, it is necessary to understand the terminology, as it is used in each jurisdiction, in order to have a meaningful discussion on interchangeability/automatic-substitution/switching. However, regardless of how the term "interchangeability" is used, it always refers to the practice of using the biosimilar in place of the reference product.

Notably, the EMA does not make decisions on whether a biosimilar can be automatically substituted, at the pharmacy level, for its reference product. In Europe, decisions on whether biosimilars are interchangeable, and can therefore

be automatically substituted for a reference, are made at the national level. Various member states of the European Union have made varying decisions on whether automatic substitution should be allowed. For example, France introduced legislation indicating that biosimilar substitution could occur for patients that were treatment naïve and for whom the prescriber had not prohibited it, while Greece's medicines agency [National Organization for Medicines (EOF)] recommended against automatic substitution/interchangeability (GaBI Online 2017).

Switching

In the context of biosimilars, Health Canada considers switching to refer to a one-time change from a reference biologic to a biosimilar that is initiated only after prescriber recommendation and consultation with the patient. In general, Health Canada is in agreement with the EMA, who have supported this type of transition from one product to another when it involves input from prescribers (Government of Canada 2017b; Kurki et al. 2017).

One of the most cited concerns with switching from a reference product to a biosimilar, be it with or without prescriber intervention, is concern regarding a greater potential for immunogenicity than might be encountered if one were to remain on their originally prescribed treatment. In the recent paper by Kurki et al., European regulators addressed the concern of switch related immunogenicity. Using examples of clinical data involving "high-risk" switches, the authors pointed out that, even when switching between non-comparable products, such as different versions of recombinant coagulation factor VIII (FVIII), neutralising antibody production was not significantly increased (Kurki et al. 2017). Additionally, the concern regarding enhanced immunogenicity has not borne out among clinical switching studies. In a recent systematic review by Moots et al., 19/37 (51%) identified switch studies assessed immunogenicity pre- and post-switch (Moots et al. 2017). The studies involved switches from infliximab, etanercept, adalimumab or rituximab to their respective biosimilars (most studies involved switching from Remicade® to CT-P13—a biosimilar to Remicade®). In no case was it observed that the incidence of ADA was significantly increased after switching from an innovative biologic to a biosimilar; however, these observations come with the caveats that these studies are underpowered to detect such differences and did not report on such parameters as boosted ADA titer after switching. Also, the majority of studies examined the switch from infliximab to CT-P13, which may not be reflective of other innovator/biosimilar pairs. Therefore, more clinical data, collected for more biosimilar products, is needed before the risk of increased immunogenicity is definitively dismissed as a concern related to switching.

Besides clinical data, there is some indirect evidence that immunogenicity is unlikely to be elicited at a greater rate after switching than if one were to continue to receive the innovative product. This is suggested by investigations of ADA cross-reactivity against CT-P13 and infliximab (Reinisch et al. 2017). In this investigation,

ADA positive samples from patients who participated in two large trials of CT-P13 vs. infliximab, conducted in patients with rheumatoid arthritis and ankylosing spondylitis, were assessed for cross-reactivity with either tagged CT-P13 or tagged infliximab. In the vast majority of cases, ADAs elicited against the reference product were cross-reactive with the biosimilar and vice versa. Based on these results, the authors concluded that infliximab and CT-P13 share immunodominant epitopes (i.e., epitopes likely of eliciting an immune response), which argues against an increased risk of ADA after switching from infliximab to CT-P13. Certainly, it would be interesting to see more of these types of data, and investigations of ADA cross-reactivity are encouraged by the Canadian guidance pertaining to biosimilars.

Post-Market Monitoring

As is a requirement for all new biologic drugs in Canada, the sponsors/manufacturers of biosimilar biologic drugs must provide a comprehensive risk management plan (RMP). The RMP should detail the activities that will be undertaken to monitor and detect the known and potentially unknown safety signals associated with the use of the biosimilar. In addition to this, both the manufacturer and Health Canada have responsibilities to ensure the safety of drugs marketed for use in Canada, including biosimilars. Health Canada conducts routine market surveillance, monitors adverse reaction reports and investigates complaints and reports of problems taking appropriate action when necessary. Manufacturers are responsible for setting up systems to monitor reported adverse reactions and for reporting any new information received about serious adverse reactions to Health Canada. In addition, manufacturers must notify Health Canada about any studies with new safety information that may come to light during the product's life cycle.

Conclusions

Biosimilars have quickly emerged as a new class of biologic drug that has the potential to improve access to many critical medicines through the reduction of costs. Of greater importance, compared to cost, is the need to ensure that these medicines are as safe and effective as their innovative counterparts. To this end, Canada has implemented requirements for a rigorous comparative analytical, biological and clinical development paradigm with the objective of demonstrating that proposed biosimilars are highly similar to their respective reference biological drugs. To date, there are eleven biosimilar products, spanning a range of indications—from autoimmune disease to growth deficiency, that have fulfilled these requirements and now offer Canadians more choice when it comes to their treatments. In the near future, more disease areas will see biosimilars developed and marketed. In particular, there is a robust pipeline for biosimilars intended to be used in oncology. As experience is gained in this area, both manufacturers and regulators

need to continually refine their approaches to biosimilars. Manufacturers will need to continually improve upon existing technologies to further reduce the residual uncertainty that exists with respect to minor differences between the biosimilar and the reference. At the same time regulators must stay abreast of new technologies and grapple with new challenges such as the interpretation of data from switching and interchangeability studies. Finally, if biosimilars are to make a meaningful contribution to the health care system, it is critical that health care practitioners become familiar with the way in which biosimilars are developed and determine how best to incorporate these high-quality products into the practice of medicine.

References

Brennan FR, Morton LD, Spindeldreher S, et al. Safety and immunotoxicity assessment of immunomodulatory monoclonal antibodies. MAbs. 2010;2:233–55.

Brinks V. Immunogenicity of biosimilar monoclonal antibodies. GaBI J. 2013;2:188–93.

Cai X, Wake A, Gouty D. Analytical and bioanalytical assay challenges to support comparability studies for biosimilar drug development. Bioanalysis. 2013;5:517–20. https://doi.org/10.4155/bio.13.1.

Casadevall N, Nataf J, Viron B, et al. Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. N Engl J Med. 2002;346:469–75. https://doi.org/10.1056/NEJMoa011931.

Committee for Medicinal Products for Human Use (CHMP). Guideline on bioanalytical method validation. 2011.

Committee for Medicinal Products for Human Use (CHMP). Guideline on immunogenicity assessment of monoclonal antibodies intended for in vivo clinical use. 2012a.

Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products containing monoclonal antibodies - non-clinical and clinical issues. 2012b.

Desai K, Catalano T, Rai G, et al. Confirmation of biosimilarity in a pharmacokinetic/pharmacodynamic study in healthy volunteers for an analytically highly similar pegfilgrastim. Clin Pharmacol Drug Dev. 2016;5:354–63. https://doi.org/10.1002/cpdd.269.

Ducourau E, Mulleman D, Paintaud G, et al. Antibodies toward infliximab are associated with low infliximab concentration at treatment initiation and poor infliximab maintenance in rheumatic diseases. Arthritis Res Ther. 2011;13:R105. https://doi.org/10.1186/ar3386.

GaBI Online. Biosimilar substitution in Europe. 2017.

Government of Canada. Pharmaceutical Industry Profile. 2017a. Available at: https://www.ic.gc.ca/eic/site/lsg-pdsv.nsf/eng/h_hn01703.html. Accessed 22 Nov 2017.

Government of Canada. Fact sheet: biosimilars. 2017b. Available at: https://www.canada.ca/en/health-canada/services/drugs-health-products/biologics-radiopharmaceuticals-genetic-therapies/applications-submissions/guidance-documents/fact-sheet-biosimilars.html.

Health Canada. Guidance document: patented medicines (notice of compliance) regulations. 2012.

Health Canada. Guidance document: information and submission requirements for biosimilar biologic drugs. 2016.

Heise T, Zijlstra E, Nosek L, et al. Euglycaemic glucose clamp: what it can and cannot do, and how to do it. Diabetes Obes Metab. 2016;18:962–72. https://doi.org/10.1111/dom.12703.

Hobbs KF, Cohen MD. Rheumatoid arthritis disease measurement: a new old idea. Rheumatology (Oxford). 2012;51 Suppl 6:vi21–7. https://doi.org/10.1093/rheumatology/kes282.

International Council on Harmonisation. ICH harmonized tripartite guideline: preclinical safety evaluation of biotechnology-derived pharmaceuticals. 1997;S6(R1).

International Council on Harmonization. ICH harmonised tripartite guideline: comparability of biotechnological/biological products subject to changes in their manufacturing process 2004;4.

Jahn EM, Schneider CK. How to systematically evaluate immunogenicity of therapeutic proteins - regulatory considerations. New Biotechnol. 2009;25:280–6. https://doi.org/10.1016/j.nbt.2009.03.012.

Kalden JR, Schulze-Koops H. Immunogenicity and loss of response to TNF inhibitors: implications for rheumatoid arthritis treatment. Nat Rev Rheumatol. 2017;13:707–18. https://doi.org/10.1038/nrrheum.2017.187.

Klein AV, Wang J, Feagan BG, et al. Biosimilars: state of clinical and regulatory science. J Pharm Pharm Sci. 2017;20:332–48. https://doi.org/10.18433/J3VK90.

Kurki P, van Aerts L, Wolff-Holz E, et al. Interchangeability of Biosimilars: a European perspective. BioDrugs. 2017;31:83–91. https://doi.org/10.1007/s40259-017-0210-0.

Maas C, Hermeling S, Bouma B, et al. A role for protein misfolding in immunogenicity of biopharmaceuticals. J Biol Chem. 2007;282:2229–36.

Moots R, Azevedo V, Coindreau JL, et al. Switching between reference biologics and biosimilars for the treatment of rheumatology, gastroenterology, and dermatology inflammatory conditions: considerations for the clinician. Curr Rheumatol Rep. 2017;19:37. https://doi.org/10.1007/s11926-017-0658-4.

NIH National Cancer Institute. Using the immune system in the fight against cancer: discovery of rituximab. 2014. Available at: https://www.cancer.gov/research/progress/discovery/blood-cancer. Accessed 22 Nov 2017.

Pascual-Salcedo D, Plasencia C, Ramiro S, et al. Influence of immunogenicity on the efficacy of long-term treatment with infliximab in rheumatoid arthritis. Rheumatology (Oxford). 2011;50:1445–52. https://doi.org/10.1093/rheumatology/ker124.

Perdriger A. Infliximab in the treatment of rheumatoid arthritis. Biologics. 2009;3:183–91.

Putnam WS, Prabhu S, Zheng Y, et al. Pharmacokinetic, pharmacodynamic and immunogenicity comparability assessment strategies for monoclonal antibodies. Trends Biotechnol. 2010;28:509–16. https://doi.org/10.1016/j.tibtech.2010.07.001.

Rader RA. An analysis of the US biosimilars development pipeline and likely market evolution. Bioprocess Int. 2013;11:S16–23.

Reinisch W, Jahnsen J, Schreiber S, et al. Evaluation of the cross-reactivity of antidrug antibodies to CT-P13 and infliximab reference product (Remicade): an analysis using immunoassays tagged with both agents. BioDrugs. 2017;31:223–37. https://doi.org/10.1007/s40259-017-0219-4.

Ryan AM. Frontiers in nonclinical drug development: biosimilars. Vet Pathol. 2015;52:419–26. https://doi.org/10.1177/0300985814547282.

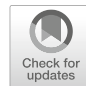
Steenholdt C, Svenson M, Bendtzen K, et al. Severe infusion reactions to infliximab: aetiology, immunogenicity and risk factors in patients with inflammatory bowel disease. Aliment Pharmacol Ther. 2011;34:51–8. https://doi.org/10.1111/j.1365-2036.2011.04682.x.

U.S. Food and Drug Administration. Guidance for Industry: considerations in demonstrating interchangeability with a reference product (*draft guidance*). 2017.

Waller CF, Bronchud M, Mair S, et al. Pharmacokinetic profiles of a biosimilar filgrastim and amgen filgrastim: results from a randomized, phase I trial. Ann Hematol. 2010;89:927–33. https://doi.org/10.1007/s00277-010-0961-x.

Wang J, Lauzon W, Njue C, et al. Extrapolation of indications for biosimilars: opportunity for developers and challenges for regulators. In: Endrenyi L, Declerck P, Chow S, editors. Biosimilar drug product development. New York: CRC Press; 2017. p. 265–8.

Chapter 7
EU Perspective on Biosimilars



Pekka Kurki and Niklas Ekman

Abstract Biosimilars are regulated by the centralized marketing authorization procedure co-ordinated by the European Medicines Agency (EMA). Marketing authorization is granted by the Commission of the European Union (EU). Thus, each biosimilar has one regulatory assessment as well as the same product information and conditions of use in the EU. The current regulatory framework is a result of 20 years evolution of legislation and regulatory guidelines. The concept of biosimilarity is based on the long experience of comparability studies of individual biological products after manufacturing changes. Therefore, the development of a biosimilar is an extensive comparability exercise with head-to-head comparisons to its reference product that must have a full dossier of quality, safety and efficacy. For the sake of global development, the current guidance allows the use of a reference sourced from a non-EU country in certain non-clinical and clinical studies provided that it can be shown to be a relevant as a comparator. The high similarity of a biosimilar and its reference is demonstrated by physico-chemical and structural as well as in vitro functional comparability studies. Non-clinical in vivo studies are rarely needed. The extent of clinical studies depends on the possibilities to demonstrate high similarity by analytical tests. When the high analytical similarity has been established and the comparable pharmacokinetics and -dynamics, safety and efficacy have been demonstrated in one therapeutic indication, extrapolation safety and efficacy to other indications is expected. Several national regulatory agencies endorse physician-guided switches between biosimilars and reference products without additional clinical trials.

P. Kurki (✉)
Clinical Immunology, University of Helsinki, Helsinki, Finland
e-mail: pekka.kurki@fimea.fi

N. Ekman
Finnish Medicines Agency, Biotechnology, Helsinki, Finland
e-mail: niklas.ekman@fimea.fi

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_7

Keywords Biosimilars · European medicines agency · Regulation · Comparability · Safety · Immunogenicity extrapolation · Interchangeability · Product information

Introduction

The knowledge of the regulatory system in the European Union (EU) helps to understand the European perspective on biosimilars (EMA 2017a). EU has a harmonized pharmaceutical legislation to guarantee high standards for quality, safety and efficacy as well as for promoting well-functioning internal market that encourages innovation and competitiveness (EudraLex 2017). However, there are more than 50 national regulatory agencies (NCAs) for medicinal products in the 28 EU Member States. This kind of a de-centralized regulatory system may solve effectively purely national issues. However, evaluation and surveillance of medicinal products require highly specialized expertise and large data bases that are difficult to obtain in small NCAs.

The European Medicines Evaluation Agency (EMEA), later named as European Medicines Agency (EMA), was established in 1995 to co-ordinate the scientific resources of the NCAs and to enable a thorough and uniform evaluation of the most important medicinal products, including similar biological medicinal products (biosimilars) in the centralized marketing authorization procedure. The EU Commission may grant the marketing authorization if it has received a positive scientific assessment report ("opinion") from EMA.

The guidelines and scientific assessment reports of EMA are prepared by its scientific committees. The Committee for Human Medicinal Products (CHMP) assesses marketing authorization applications of human medicinal products, including biosimilars. Two of CHMP working parties, Biologicals Working Party (BWP) and Biosimilar Medicines Working Party (BMWP) have been in the key position in drafting the CHMP guidelines for biosimilars (EMA 2017b).

> The European Union has a harmonised legislation and guidance for the quality, efficacy, and safety of medicinal products. Marketing authorisation applications of biosimilars are evaluated in the centralised marketing authorization procedure co-ordinated by the European Medicines Agency. A centralised marketing authorisation provides the same conditions of marketing and use in all member states of the European Economic Area (EEA).

Evolution of the EU Policy on Biosimilars

Biosimilars are still not well known outside the regulatory agencies and the pharmaceutical industry. Thus, biosimilars may appear as a new regulatory invention. However, it is important to understand that the biosimilar concept is based on a long experience and evolution.

The first serious European discussions on a possibility to copy a biological product took place in 1998 when CHMP asked BWP to develop a guideline on comparability (EMEA 1998, 2003). BWP decided to include two scenarios in the guideline; first, maintenance of comparable quality, safety and efficacy after a manufacturing change and, second, demonstration of comparability of two independent products, i.e. a biosimilar and its reference product. The first guideline focused on quality issues. During the preparation of the guideline, it became clear that a demanding comparability program ("exercise") will require multidisciplinary collaboration. Therefore, CHMP established a multidisciplinary *ad hoc* working group of comparability that later became the Biosimilar Medicines Working Party, BMWP.

Since there was uncertainty of the legal basis of biosimilars, EU Commission drafted the necessary legal text that was introduced to medicinal product directive in 2005 (European Commission 2017a). The directive states that the information to be supplied for marketing authorization of a biosimilar shall not be limited to pharmaceutical, chemical and biological data and bio-equivalence and bio-availability data. The type of additional data (i.e. toxicological and other non-clinical and appropriate clinical data) will be determined on a case by case basis in accordance with *relevant scientific guidelines*.

According to the directive, the general principles to be applied are addressed in a *guideline* taking into account the characteristics of the concerned biological medicinal product published by the Agency (EMA). In case the originally authorized medicinal product has more than one indication, the efficacy and safety of the medicinal product claimed to be similar has to be justified or, if necessary, demonstrated separately for each of the claimed indications.

> The development and assessment of biosimilars is definitely based on the concept of comparability. The legal text has proven to be suitable for biosimilars as it gives the possibility to modify the regulatory requirements according to the advances of scientific knowledge and clinical experience in a flexible way.

Guidelines for Biosimilars

CHMP guidelines are drafted in a transparent process that includes two public consultations (Fig. 7.1) (Kurki and Ekman 2015).

In addition, workshops with stakeholders are often organized before finalization of the guideline. The trigger for the development of a guideline comes often from the scientific advices or marketing authorization reviews.

Guidelines will be updated regularly, usually every 5 years. Guidelines for the development and assessment of biosimilars are typically at a rather general level focusing on main regulatory requirements and avoid technical details that may be outdated rapidly. As such, the EMA guidelines provide freedom for the Applicants to use state of the art methods and best practises in their development programs. It is also possible to deviate from the guidance with good justification usually following consultation with a relevant regulatory authority.

CHMP, together with its working parties BWP and BMWP have created an extensive set of guidelines for biosimilars (EMA 2017c) (Table 7.1). There is a hierarchy at three levels among the guidelines. The guideline for “Similar biological medicinal products” presents the definition of a biosimilar, main policy issues, such as requirements for the reference product and the main requirements of and strategy for demonstration of comparability between the biosimilar and its reference. This guideline acts as a bridge between the legislation and other guidelines for biosimilars.

On the second level are the two overarching guidelines for “Similar biological medicinal products containing biotechnology-derived proteins as active substance”, one for quality issues and the other for non-clinical and clinical issues. They describe the general principles of the demonstration of comparability of therapeutic proteins manufactured by rDNA technology, including the physico-chemical and structural comparability exercise, non-clinical studies, human pharmacokinetic (PK), pharmacodynamic (PD), efficacy and safety studies, and post-marketing safety surveillance.

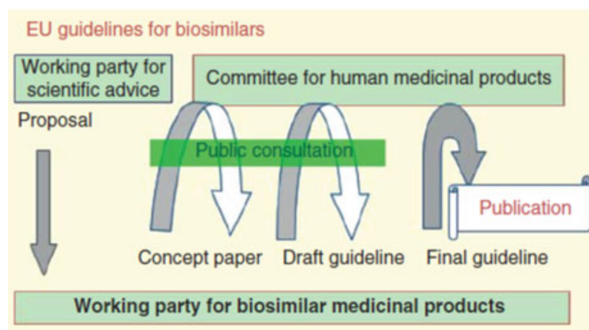


Fig. 7.1 The process of preparing a guideline in the EU

Table 7.1 Biosimilar products and the scope of guidelines in the EU

Proteins/peptides		Guideline for similar biological products	Overarching guidelines for biotechnology derived proteins	Product class-specific guideline
Growth factors	Epoetins	√	√	√
	Filgrastim	√	√	√
Hormones	Follitropin alfa	√	√	
	Insulins	√	√	√
	Teriparatide	√	√	
	Somatropin	√	√	√
Monoclonal antibodies	Adalimumab	√	√	√
	Infliximab	√	√	√
	Rituximab	√	√	√
	Bevacizumab	√	√	√
Fusion proteins	Etanercept	√	√	
Others				
	Enoxaparin	√		

The third layer of guidelines, product class-specific guidelines present more detailed guidance on non-clinical and clinical studies tailored for certain product classes.

The maintenance of the large set of guidelines for biosimilars at the level of scientific knowledge and experience is laborious and challenging, especially for consistency. Nevertheless, the European guidelines have allowed the development of high quality, safe and efficacious biosimilars over the years.

EU has an extensive set of guidelines for the development and assessment of biosimilars. The main guideline acts as a bridge from the legislation providing the main "policy issues" whereas the "overarching guidelines" give the main scientific principles for biosimilars of biotechnology-derived proteins. More tailored guidance is given in guidelines for certain product classes. Regular updates allow the evolution of the guidelines according to the scientific progress and clinical experience.

Assessment of Biosimilars

The assessment of the marketing authorisation applications of biosimilars is conducted in the same way as assessment of new innovative products and using the same standards of quality, safety, and efficacy (Fig. 7.2). For each application,

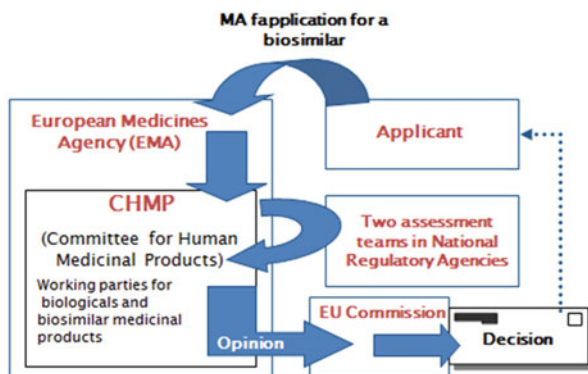


Fig. 7.2 Marketing authorisation (MA) process

the CHMP will nominate two rapporteurs. The rapporteurs, usually experts from national regulatory agencies, will create their assessment teams with expertise in pharmacy and biotechnology, non-clinical testing, clinical studies of PK, PD, efficacy and safety, including immunogenicity, statistics, and risk assessment. The two assessment reports are then submitted for CHMP review. On the basis of the comments, a joint assessment report with questions to the Applicant is created.

The Applicant will normally respond to the questions within 3 months. The responses are evaluated by the rapporteurs' teams that will create a list of outstanding issues. CHMP will then send the list to the Applicant for speedy responses. A third round is often needed before CHMP gives its opinion on the approvability of a biosimilar. EU Commission may then take the decision on marketing authorisation.

The two independent reviews, repeated interactions with the applicant and several rounds of discussions in the CHMP guarantee a balanced assessment of applications. The quality and reliability of the manufacturing as well as the non-clinical and clinical data are controlled by GMP-, GLP- and GCP-inspections. The scientific assessment does not involve re-analysis of raw data except in exceptional cases when the analyses of the Applicant are not satisfactory or reliable.

The two independent reviews of the marketing authorisation applications, repeated interactions with the applicant and several rounds of discussions guarantee a balanced assessment of applications of biosimilars that often raise new and controversial issues. The quality of data is ensured by GMP-, GLP, and GCP inspections and by requests for additional data or analyses as with other products. Systematic analyses of data of individual subjects (raw data) are rarely performed by the regulators in EU.

Reference Product

The medicinal product that is used as the reference for a new candidate biosimilar must be or have been authorized in the European Economic Area (EEA) on the basis of a complete dossier in accordance with the provisions of Directive 2001/83/EC (European Commission 2017a). A single reference medicinal product should be used as the comparator throughout the comparability program for quality, safety and efficacy during the development of a biosimilar. In principle, the reference product should be sourced from the EEA-countries.

The use of locally authorized reference products is problematic for global development of biosimilars. Repeating the same clinical comparability studies in all jurisdictions would be unethical and unnecessary, especially because the licenses of the reference product in different jurisdictions refer mainly to the same clinical trials. Therefore, the EU legislation has been interpreted a way that allows the use a non-EEA authorized comparator in certain non-clinical and clinical studies. However, such product will need to be authorized by a regulatory authority with similar scientific and regulatory standards as EMA.

The Applicant has to present "bridging data" to demonstrate that the comparator authorized outside the EEA is representative of the reference product authorized in the EEA. The type of bridging data will always include information and data from analytical studies (e.g., physico-chemical, structural and functional data) that compare the biosimilar, the EEA-authorized reference product, and the non-EEA-authorized comparator, and may also include data from clinical PK and/or PD bridging studies for all three products as outlined in the guideline "Similar Biological Medicinal Products" (EMA 2017c).

The candidate biosimilar should be compared to a product (reference product) authorized in the EEA on the basis of a full dossier on quality, safety, and efficacy. In order to allow global development of biosimilars, it is possible to replace the reference product with a non-EEA-licensed comparator for certain non-clinical and clinical studies. In this case, the Applicant has to present information and analytical bridging studies and, if deemed necessary, pharmacokinetic and—dynamic studies to demonstrate that the selected comparator is relevant to its purpose.

Quality

The starting point for the development of a biosimilar is a thorough understanding of the properties and characteristics of the reference product. In addition to publicly available information, extensive characterization studies of the reference product are required already at an early stage of product development. Furthermore, the

criticality of individual quality attributes need to be understood and batch-to-batch variability of the reference product quality attributes assessed.

As defined in ICH guideline Q8 (R2) "Pharmaceutical development" (EMA 2017d), the quality target product profile (QTPP) is a prospective summary of the quality characteristics of a drug product that ideally will be achieved. For biosimilar development, the goal is to manufacture a close copy of the reference medicinal product. Consequently, the QTPP for a biosimilar is primarily based on the detailed understanding of the reference product gained through extensive characterization studies. The characterization data forms the basis for the design of the biosimilar manufacturing process and should ensure manufacture of a product that is, from the analytical point of view, highly similar to the reference product authorized in the EEA.

In order to achieve the QTPP goal, the expression system and the individual cell clone to be used for manufacturing need to be carefully considered and selected. In addition, cell culture conditions and purification process principles have to be optimized. This is a highly critical step in the development of the biosimilar which combines the thorough understanding of the reference product with manufacturing process knowledge. Although the manufacturing process development is time consuming, it should, as far as possible, be completed before continuing to the next stage of development. Introduction of major changes into the manufacturing process later during product development could have an unfavorable impact on the overall evidence for similarity. From the experience gained, it is evident that entering into clinical studies using a biosimilar candidate derived from an early and potentially suboptimal manufacturing process is not recommended. Manufacturing process changes introduced at a later stage and potentially affecting the quality profile of the biosimilar are typically difficult to accept without new clinical trials.

During product development, comparability between the candidate and the reference is continuously evaluated. Additionally, once the intended final commercial manufacturing process has been established, the analytical comparability is typically confirmed in a large and, as far as possible, side-by-side comparability exercise using sensitive and orthogonal analytical methods. The results from the biosimilar and reference product batches should be compared directly to each other, but also against comparability ranges pre-established and pre-justified based on the characterization results of the reference product. It should be noted that, according to the EU guideline "Similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues" (EMA 2017c), the comparability ranges cannot be wider than the range of variability present in the reference medicinal product.

The principal difference between quantification of uncertainty and establishment of comparability ranges has to be acknowledged. Statistical ranges, such as tolerance ranges that estimate the unknown actual range of the data set (in this case the reference product batches) are usually not considered appropriate for establishment of comparability ranges. Such ranges could, however, be accepted in case of low batch-to-batch variability and/or if a large number of reference batches have been analyzed.

Other approaches, for example comparability ranges based on min-max characterization values of the reference product, or inferential statistics approaches can likewise be applied. However, the methodological strengths and weaknesses of the chosen approach always need be understood, appreciated and discussed. It should be emphasized that, independently of the approach taken, the statistical method itself cannot define what is sufficiently similar in order to achieve equivalent clinical performance. This needs to be evaluated based on understanding of the product, including quality attribute criticality assessment.

The biosimilar is not expected to be analytically identical to the reference product. In part due to the rapid development in the analytical methods used for characterization of biomolecules, an increasing number of small but confirmed differences between the biosimilars and their reference products are seen. Since the authorization of the first biosimilar, it has been the view of the EU regulators that analytical differences between the biosimilar and the reference product are acceptable, provided that the differences do not affect the clinical performance of the product. In principle, this approach does not make a difference between quality attributes of high and low criticality, but differences in low criticality attributes will inevitably be easier to justify compared to differences in highly critical quality attributes.

In most cases, the soundest approach for addressing the clinical relevance of confirmed quality attribute differences is by applying relevant, sensitive and properly qualified functional assays. Non-clinical in vivo studies and human clinical trials are usually not sufficiently sensitive in addressing the relevance of minor differences in quality attributes. On the other hand, if major quality difference exists between the biosimilar and the originator, then the use of the biosimilar approval route should be questioned.

Functional assays have been successfully used to justify quality differences for many approved products, as well as products under development. For CT-P13 infliximab, the Applicant used successfully a number of functional assays, including "indication-specific" assays in order to address the clinical impact of differences in afucosylation level, FcγRIIIa binding and ADCC activity which were detected between the biosimilar and the reference product (EMA 2017e). Since then, similar approaches have been used for several other biosimilar products, demonstrating the strength of the extended functional testing for justifying differences detected in quality attributes.

Non–Clinical Development

Most adverse effects of a biological product are due to its pharmacological effect that is already known from studies and post-marketing surveillance of the reference product. The majority of the biosimilars are polypeptides or proteins that will rarely cause unexpected toxic reactions related to their metabolic routes. Considering the high similarity proven by physico-chemical and structural as well as in vitro

functional comparability analyses, it is highly unlikely that the active substance of the biosimilar candidate would cause unexpected toxicity (van Aerts et al. 2014). Therefore, the approach to non-clinical development differs from the non-clinical development of a product with a new active substance.

The guideline for "Similar biological medicinal products containing biotechnology-derived proteins as active substance—non-clinical and clinical issues" describes the stepwise approach to non-clinical development (EMA 2017c). Head-to-head comparisons by in vitro functional tests create the foundation for non-clinical studies. The tests should be relevant to the pharmacological action and be sensitive enough to detect differences. The product class-specific guidelines may suggest suitable in vitro tests.

The non-clinical in vitro studies should compare the concentration–activity/binding relationship of the biosimilar and the reference medicinal product at the pharmacological target(s), covering a concentration range where potential differences are most sensitively detected. These tests should involve several batches of the reference product and of the biosimilar that are representative of the material intended for clinical use and should cover the whole spectrum of pharmacological/toxicological aspects known to be of clinical relevance for the reference product and for the product class.

According to the guideline, these head-to-head comparative studies should include relevant assays on:

– *Binding* to target(s) (e.g. receptors, antigens, enzymes) known to be involved in the pharmaco-toxicological effects and/or pharmacokinetics of the reference product.
– *Signal transduction and functional activity/viability* of cells known to be relevant for the pharmaco-toxicological effects of the reference product.

These assays can be considered as paramount for the non-clinical biosimilar comparability exercise since in vitro assays may often be more specific and sensitive to detect differences between the biosimilar and the reference product than in vivo studies in animals.

At the next stage, the need for non-clinical in vivo toxicological studies will be considered. There are several factors that reduce the feasibility of in vivo studies of biosimilars:

- Toxic effects of a therapeutic proteins are usually mediated by the known pharmacological effects of drug substance.
- The function of drug substance is often species-specific
- Biological drug substances are often immunogenic in conventional animal models
- Conventional animal models are not sensitive enough to detect differences in the biosimilar setting
- The risk of transit to clinical studies can be predicted by the knowledge of the reference product and the outcome of physico-chemical, structural and in vitro functional tests

Furthermore, the EU legislation requires the application of the 3Rs principle (Reduction, Refinement and Replacement of animal experiments) for product development in order to reduce suffering of animals (European Union 2017). Especially, the use of non-human primates is discouraged while for many biosimilars, non-human primates are the only species that share the function of the molecule with man.

The development will proceed to the third stage in the rare cases when in vivo non-clinical studies of safety are deemed necessary. Under the current regulatory praxis, CHMP has not required non-clinical in vivo studies in the development of biosimilars.

Predictive in vitro or non-clinical in vivo tests for immunogenicity may be useful for developers in selecting their biosimilar candidate. However, these tests are not yet sufficiently validated for determination of relative immunogenicity for regulatory purposes.

The adverse effects of a biosimilar can be usually predicted on the basis of the experience of the reference product and the outcome of the physico-chemical, structural and in vitro functional tests. Predictive in silico tests or animal models are not yet sufficiently validated for regulatory purposes in the EU. In general, the in vivo non-clinical comparability studies are regarded questionable if preceding physico-chemical, structural and in vitro functional tests have demonstrated a high similarity of the biosimilar to its reference.

Clinical Studies

In the stepwise development of a biosimilar, clinical studies represent the last, confirmatory part of the biosimilar comparability exercise. Therefore, it is expected that the batches used in clinical studies represent the manufacturing process that will be used to produce commercial batches. Adequate additional bridging data should be presented if the Applicant wishes to deviate from this recommendation. It is also recommended that the clinical study program will start by demonstration of comparable pharmacokinetics and—dynamics and followed, if necessary, by confirmatory efficacy and safety studies.

Pharmacokinetic Comparability

Comparative pharmacokinetic (PK) studies designed to demonstrate a comparable PK profile are always required for biosimilars with a protein or polypeptide as the

active substance. Products containing peptides or non-protein active substances may sometimes be exempted from PK studies for scientific and technical reasons.

Comparative pharmacokinetic studies should be planned on the basis of the characteristics of the reference product, including its mode of action, safety profile and pharmacokinetic properties, such as target-mediated disposition, linear or non-linear PK, time-dependency, and half-life.

The EU guideline on non-clinical and clinical issues (EMA 2017f) recommends a single dose cross-over study design with full characterisation of the PK profile, including the late elimination phase. A parallel group design suits better for proteins with a long half-life and potential for significant immunogenicity, such as monoclonal antibodies. Provided that there are no safety concerns, a single dose study in healthy volunteers is recommended as it minimises most confounding factors, such as target-interference. The doses in the single dose PK biosimilar comparability study in healthy volunteers may be lower than the recommended therapeutic doses provided that the dose is still on the linear part of the kinetics. If a study in healthy volunteers is not feasible, a multiple dose study in patients needs to be conducted. This study may be part of a safety and efficacy study. However, it is recommended that the comparative PK will be demonstrated before exposing a large group of patients. A sensitive model/population, i.e. that has fewer factors that cause major inter-individual or time-dependent variation, should be explored.

In a single dose PK study, the primary parameters are the $AUC_{(0\text{–}inf)}$ for intravenous administration and $AUC_{(0\text{–}inf)}$ and usually C_{max} for subcutaneous administration. Secondary parameters such as t_{max}, volume of distribution, and half-life, should also be estimated. In a multiple dose study, the primary parameters should be the truncated AUC after the first administration until the second administration ($AUC_{0\text{–}t}$) and AUC over a dosage interval ($AUC\tau$) at steady state. Secondary parameters are C_{max} and C_{trough} at steady state.

Sometimes, a multiple dose PK study in patients may not be feasible due to major variability in target expression, including variability over time. In these cases, in vitro studies may show comparable interaction between the biosimilar and its target(s), including FcRn for a monoclonal antibody, and a pivotal PK study in the target population may not be needed. In this case, additional PK data should be collected during the efficacy, safety and/or PD studies. This allows further investigation of the clinical impact of variable pharmacokinetics and possible changes in the PK over time, for example by population pharmacokinetics. It is recommended that some PK parameters, notably the trough concentration, are included in the multiple dose safety and efficacy studies in order to detect changes in PK over time.

The acceptance criteria used in standard clinical bioequivalence studies of chemical substances may be applicable for the planning comparative pharmacokinetic trials for biologicals. However for biologicals, PK is used to measure not only absorption but also possible differences in the interaction with the body between the originator and the biosimilar. Thus, 90% CIs of ratios of biosimilar to reference product within a pre-specified, justified acceptance range may not be sufficient. The location and the width of the confidence interval should also be taken into account in

the interpretation of similarity. Thus, statistically significant differences in 90% CIs within the justified acceptance range regarding relevant PK parameters would need to be explained and justified. In addition, if the 90% CI crosses the pre-specified boundaries the applicant would need to explain such difference and explore root causes. Pre-specified correction for protein content may be acceptable if adequately justified.

If the reference product can be administered both intravenously and subcutaneously, the evaluation of subcutaneous administration will usually be sufficient as it covers both absorption and elimination. In a pivotal PK study, anti-drug antibodies should be measured in parallel to PK assessment using appropriate sampling time points.

Comparative pharmacokinetic (PK) studies designed to demonstrate a comparable PK profile are always required from biosimilars with a protein or polypeptide as the active substance. A single dose study in healthy volunteers preceding multiple efficacy and safety is regarded as the most sensitive test for comparative PK. The standard bioequivalence range may not always be sufficient for demonstration of comparable PK profiles of the biosimilar and its reference product. Introduction of some parameters into the multiple dose efficacy and safety studies is strongly recommended.

Pharmacodynamic Comparability

The role of pharmacodynamic tests is often supportive. It is recommended that relevant pharmacodynamic endpoints are added to comparative pharmacokinetic and efficacy and safety studies.

In certain cases, pharmacodynamic studies may be sufficient to demonstrate clinical comparability of the biosimilar and the reference medicinal product (EMA 2017f). If there is an accepted surrogate marker of efficacy, the clinical study may be shorter and sample size smaller. Examples of recommended endpoints for confirmatory clinical studies include absolute neutrophil count to assess the effect of granulocyte-colony stimulating factor (G-CSF), early viral load reduction in chronic hepatitis C to assess the effect of alpha interferons, and euglycaemic clamp test to compare two insulins. Magnetic resonance imaging of disease lesions can be used to compare two β-interferons in multiple sclerosis.

There may be relevant PD-markers that are not established surrogates for efficacy. If such a PD marker has a clear dose-response or a concentration-response relationship, a single or multiple dose-exposure-response study at two or more dose levels may provide sufficient data for the omission of the confirmatory efficacy study.

In exceptional cases, the confirmatory clinical trial may be waived if physico-chemical, structural and in vitro biological analyses and human PK studies together with a combination of PD markers that reflect the pharmacological action and concentration of the active substance can provide robust evidence for biosimilar comparability. Currently, this strategy may be applied for peptides. The Applicants should ask for scientific advice if they plan to use this “fingerprinting” concept.

In most cases, PD studies will provide supportive evidence for clinical comparability. Accepted surrogate markers may replace the clinical endpoints in a safety and efficacy study. A multiple dose pharmacodynamic study may replace the clinical safety and efficacy study if there is a relevant PD marker with linear dose-concentration-response curve. A “fingerprinting” approach without clinical trials is acceptable for peptides.

Clinical Comparability

For time being, the comparative efficacy and similar safety profile will normally be confirmed in randomized, preferably blinded clinical trials (EMA 2017f). However, it should be remembered that clinical studies are usually not very sensitive for differences. Thus, they cannot compensate major differences in physico-chemical and structural studies. In addition, the recording of a large number of efficacy and safety parameters will always lead to differences just by chance. The aim of the efficacy and safety studies is to confirm the clinical comparability, not to re-demonstrate the efficacy and safety of the biosimilar that, according to the EU definition, will contain just a new version of the same active substance as the reference product. As a result, the confirmatory clinical trials of biosimilars will not be identical to pivotal efficacy studies of the reference product.

In principle, a biosimilar should target the same therapeutic indications as its reference product. Usually, one of the therapeutic indications is selected for a confirmatory clinical efficacy and safety study. This “lead” therapeutic indication should provide a sensitive model for a comparative clinical trial that, in addition to efficacy, addresses all relevant safety issues of the reference product. In some special situations, it may be possible to study an off label indication, for example when the therapeutic approach to combination chemotherapy as changed since the conduct of pivotal clinical trials with the reference product.

The EU guidelines recommend equivalence design but non-inferiority design is also possible if justified. The equivalence/non-inferiority range should be based on data of the reference product. Ideally, the delta should represent the largest difference that is clinically insignificant. The primary efficacy endpoints should be selected on the basis of their sensitivity for differences. Therefore, the selected

primary efficacy endpoints may not be the same as in the pivotal clinical trial of the reference product, such as overall survival or disease progression. Instead, overall response rate in cancer or disease activity in rheumatoid arthritis may provide better discriminatory value. It is recommended that the confirmatory efficacy and safety studies will include PK measurements, especially trough levels, to facilitate the interpretation of immunogenicity results.

Randomized, usually blinded, parallel group clinical efficacy and safety trials are still required for biosimilars of more complex biosimilars, such as monoclonal antibodies. The role of these studies is confirmatory to the results of the physico-chemical, structural, and in vitro functional comparability studies. The selected therapeutic indication should be representative for other therapeutic indications in terms of efficacy and safety. It is important to integrate the testing of efficacy and safety, on one hand, and PK, PD, and immunogenicity, on the other hand. The endpoints should be sensitive for differences.

Immunogenicity of Biosimilars

The human immune system has evolved to recognising proteins. The purpose of the recognition is to distinguish foreign or denatured self-proteins from body's normal proteins. If a protein is deemed foreign or denatured, the immune system will mount an immune reaction against the protein. The immune system is normally tolerant to normal human proteins and their copies produced using biotechnological methods. The immunological tolerance varies between individuals as it is partly genetically determined.

Many therapeutic proteins are similar to normal proteins of the body. Therefore, they may be recognised by the immune system as "self" and no or a limited activation of the immune system will take place. However, other therapeutic proteins, especially if denatured, evoke an immune reaction that is usually detected by measuring anti-drug antibodies (ADAs).

Safety problems may arise if the ADA-response will continue to evolve. Life-threatening hypersensitive reactions may occur if the ADAs will have a class switch to IgE or if pathogenic immune complexes are formed. Another serious reaction is possible if therapeutic protein has an endogenous counterpart that is targeted by cross-reacting neutralizing ADAs. Cross-reactive epoetin alfa-ADAs may cause a serious complication, pure red cell aplasia, by neutralizing the endogenous erythropoetin. Patients with enzyme-deficiencies do not have immune tolerance to the normal enzyme used in the replacement therapy. Therefore, patients often develop ADAs that may neutralise the therapeutic enzyme or may cause serious

hypersensitivity reactions if there is a switch to IgE class ADAs. It is important to analyse the risk factors of immunogenicity to design an appropriate immunogenicity program for a given biosimilar.

EU has two guidelines for the assessment of immunogenicity of therapeutic proteins, one general guideline (EMA 2017g) and the other for monoclonal antibodies (EMA 2017h). Regulatory guidance on the immunogenicity of biosimilars is also found in the guidelines specific for biosimilars.

The basic immunogenicity package for a biosimilar consists of comparative data on ADA incidence, titres, neutralizing capacity, and persistence as well as clinical correlations. The ultimate purpose is to detect harmful immunogenicity. The first step is to compare the incidence, titre, and neutralising capacity. Secondly, possible clinical correlations should be looked for. Differences in the immunological impact on efficacy and safety are not tolerated. The third step is to monitor immunogenicity post-marketing.

The foundation for understanding the impact of immunogenicity of a therapeutic protein is the assay of ADAs. The guideline describes the assay strategy that starts with a sensitive screening assay. The positive samples are then confirmed by an assay that is more specific. Confirmed samples are further titrated and characterized for neutralising ADAs. Cell-based assays for neutralizing ADAs are usually giving more complete data on neutralization when the therapeutic protein will bind to a receptor or an antigen on the cell surface. Unfortunately, cell-based assays are often insensitive and difficult to standardise. Therefore, assays measuring binding to the relevant target may also be acceptable, especially if the therapeutic protein is used to capture and eliminate its target in the circulation.

The EU guidance on ADA assays is at a rather general level and concentrating on the regulatory requirements. The guidelines do not promote any particular assay methodology but gives a list of their advantages and handicaps. For technical details, reference is made to the guidance of learned societies, such as AAPS and ABIRISK consortium (EMA 2017g). The general guideline emphasizes the need for sensitive assays that tolerate circulating therapeutic protein in the samples. The assays should be validated, including the matrix effects in different populations.

The general guidance on immunogenicity applies also for biosimilars. The biosimilar comparability program should always include the determination of relative immunogenicity of the biosimilar and its reference in head-to-head comparative clinical studies (EMA 2017f, g). The guidelines present two options for assay methodology. The most common approach is to develop a single assay for ADAs in samples from both biosimilar- and reference product-treated patients by using active substance of the biosimilar as the antigen/capture agent. The use of as single positive control is preferred. For the time being, lack of suitable human positive controls forces the Applicants to use purified animal antisera. Unfortunately, antisera from hyperimmunized animals may not be representative for human ADAs, especially in studies of drug tolerance of the ADA assay.

The other option is to develop separate but similar assays for the antibodies to the biosimilar and to its reference. This is challenging because the assays should have equivalent specificity and sensitivity. The benefit of this assay is that cross-

testing of all samples with both assays will add to the information about relative immunogenicity. For example, good concordance between the results from the two assays will suggest that the immunogenic epitopes of the biosimilar and the reference product are the same.

The ADA testing should be integrated into the pivotal PK/PD and safety and efficacy studies utilising the information on the reference product. Correlation of the ADA-results to the PK, efficacy, and safety requires an adequate number of samples and sufficient follow up to study the evolution the immune response. The length of the follow up must be justified. In chronic treatment, 1 year follow is usually required. The sampling schedule should allow the investigation of the evolution of the immune response; persistence, titres and neutralizing antibodies.

Neither the current guidelines nor the CHMP will give any pre-defined range of immunological comparability because of the multiple dimensions of immunogenicity. A difference in the incidence of ADAs does not exclude biosimilarity. However, a root cause of a difference in immunogenicity should always be investigated. First of all, ADA assay should be re-evaluated for a possible bias. The most common problem in the ADA assays is drug interference which will result in false negative results due to residual product in the blood samples for ADA analysis. In such cases, the drug tolerance of the assay(s) should be revisited and the drug concentrations in the samples compared. If no technical problem is discovered, the Applicant will be asked to review all differences observed in the analytical, structural, and functional comparisons and discus their possible role in immunogenicity.

In case of truly different ADA-incidences or titres, including neutralising ADAs, the persistence of the ADA responses and possible clinical correlations should be explored by analysing pharmacokinetics, relevant symptom complexes (e.g. hypersensitivity, autoimmunity), and doses of the comparators in relevant clinical studies. The burden of evidence is on the Applicant who must convince regulators of the lack of clinical impact of the increased ADA incidence. If no harmful effects are observed, the manufacturer will have to commit to post-marketing studies to exclude potential late and rare immunological complications of the biosimilar.

> The goal of immunogenicity studies is to detect harmful immunogenicity. The Applicants should present a summary of immunogenicity studies that contains an analysis strategy based on risk analysis, ADA results with clinical correlations, and, if necessary, the post-marketing risk detection and mitigation measures. The basis of a meaningful evaluation of immunogenicity is on ADA assays that should be sensitive enough to allow a meaningful analysis of clinical correlations. The basic "immunogenicity package" contains data on the incidence, titre, neutralizing capacity and persistence of the ADAs. An increased incidence of ADAs may be acceptable if the Applicant can demonstrate the lack of clinical impact.

Extrapolation of Efficacy and Safety

A medicinal product interfering with pathogenic pathways that affect several diseases may also have several therapeutic indications. Different etiology, pathology, clinical presentation and the fact that some *non-comparable* biological medicinal products may not be effective and safe in same therapeutic indications are not relevant in the discussion of the risks in extrapolating safety and efficacy data of a biosimilar in one therapeutic indication to others. For example, some TNF-inhibitors are effective in rheumatoid arthritis but not in inflammatory bowel disease. Such examples are not valid for biosimilars that contain highly similar versions of the active substance of the reference products (Weise et al. 2014).

Once the close similarity between the biosimilar and its reference product has been demonstrated in the physico-chemical, structural, and in vitro functional analyses as well as in human PK-, PD- and efficacy and safety studies, extrapolation of efficacy and safety to other therapeutic indications is expected but has to be justified. An essential argument is that the lead therapeutic indication is representative of other therapeutic indications in terms of efficacy and safety, including immunogenicity (EMA 2017f).

In the justification, the Applicant has to provide information on the mode of action of the active substance in the therapeutic indications. The mode of action of the active substance has to be distinguished from the pathogenesis of the disease—the binding of the active sites to their targets is relevant, not the types of the target cells. Extrapolation may be challenging if the molecule has several active sites, like monoclonal antibodies. Fortunately, there are assays to study different active sites of a monoclonal antibody to cover those functions that may not have been relevant in the clinical model used to demonstrate similar efficacy, safety and immunogenicity. The Applicant should provide these in vitro functional results as additional support to extrapolation (Weise et al. 2014).

Immunogenicity is another key factor in the justification to extrapolate. In general, it is very difficult to compare the immunogenicity studies conducted at different times, in different populations by using different immunogenicity assays. Therefore, nomination of the “most sensitive” therapeutic indication may be an illusion. It is sufficient that the model is sensitive for differences in immunogenicity and its clinical consequences.

The first biosimilar monoclonal antibody licensed in EU, CT-P13 (infliximab) raised concerns whether comparable efficacy, safety, and immunogenicity demonstrated in rheumatoid arthritis and ankylosing spondylitis can be extrapolated to inflammatory bowel disease (IBD). CHMP asked the Applicant to perform further in vitro functional assays that would better reflect the pathophysiological situation in vivo. Data from in vitro functional tests with different conditions and with different effector and target cells demonstrated comparability between the biosimilar and the reference infliximab (EMA 2017e). Post-marketing clinical trials have confirmed the clinical comparability in IBD and the extrapolation is now accepted by the gastroenterological community in Europe (Danese et al. 2017). The efficacy and

safety have been extrapolated also in the approval of other biosimilars without any unexpected problems post-marketing (Weise et al. 2014).

Once the biosimilar and its reference product have been shown to be comparable in physico-chemical, structural and in vitro functional tests and in PK-, PD-, efficacy and safety studies in one therapeutic indication, extrapolation of safety and efficacy to other therapeutic indications is expected. However, the Applicant has to justify the representativeness of the "lead" indication to the others in terms of safety, immunogenicity and efficacy, Additional data, such as in vitro non-clinical functional studies are needed in cases where the active substance of the reference product has more than one active site that may lead to variation in the therapeutic effect between the therapeutic indications.

Interchangeability

Interchangeability is an essential element in using biosimilars for the benefit of the patients and healthcare in general. It is important to note that, in the EU, interchangeability means changing one medicine for another that is expected to achieve the same clinical effect in a given clinical setting and in any patient *on the initiative, or with the agreement of the prescriber*. Automatic *substitution* of biosimilars is a practice of dispensing one medicine instead of another equivalent and interchangeable medicine at the pharmacy level *without* consulting the prescriber (European Commission and DG Enterprise and Industry 2013). Thus, the term "interchangeability" is narrower in the EU than in the US legislation that couples interchangeability with substitutability. In Europe, interchangeability is seen as a scientific and medical term whereas substitution is a political, administrative and practical (information and training of patients) measure.

The mandate of EMA does not include the assessment of interchangeability whereas the legislation allows the Member States to take a position on the interchangeability of biosimilars. In many EU Member States, the national legislation is tailored for generic substitution and reflects the outdated view that a biological product cannot be copied. However, no Member State has prohibited or given negative guidance on physician-guided switching of reference products to its biosimilar. Instead, an increasing number of Member States recommend the prescription of biosimilars and several national regulatory agencies have deemed biosimilars interchangeable with their reference products (Medicines for Europe 2017).

For time being, no EU Member State has yet formally instituted automatic substitution of biosimilars for both treatment naïve and reference product-treated patients. However, certain organizations, such as hospitals and insurance companies,

use tendering processes for the procurement of medicinal products for their list of preferred medicinal products. Addition of a biosimilar to the list will de facto lead to automatic switches in patients treated with the reference product. Furthermore, payers in Europe have become impatient with the slow pace of physician-led switches (QuintilesIMS 2017) and have introduced strong economic incentives (Mack 2015) or mandated nation-wide "non-medical" switches (Glintborg et al. 2017).

Regulatory positions on interchangeability of biosimilars in EU: Examples of Member States (regulatory authorities) supporting physician-guided switches between reference product and its biosimilars by 2017:

Austria—Austrian Medicines and Medical Devices Agency (2017). Belgium—Federal Agency for Medicines and Health Products (2016) Denmark—Danish Medicines Agency (2016) Finland—Fimea (2015) France—ANSM (2016) Germany—Paul-Ehrlich Institut (2015) Italy—Procurement law (2016) Ireland—HPRA (2015) The Netherlands—MEB (2015) Norway—Norwegian Medicines Agency (2017) Poland—Minister of Health (2014) Portugal—Infarmed (2015) Scotland—HIS (2015) Sweden—SLL (2015) United Kingdom—NHS (2015)

From a theoretical point of view, there is no reason to expect that switches between different comparable versions of the same active substance would have adverse effect on the safety or efficacy. However, such a situation may be possible if the new version, such as a biosimilar, is inferior in quality. The very rare examples of adverse effects of switching between old and new versions of a biological product come from manufacturing process changes of individual products, such as epoetin alfa-containing products. Thus, it is important to note that the problems have been due to the failure in the demonstration of comparability, not to the switch itself.

Interestingly, the problem with epoetins has been in the formulation and stability, not in differences between the active substances. Expert regulators familiar with biosimilars have pointed out that differences in the active substances, even if non-comparable, have rarely caused problems upon switching (Kurki et al. 2017). In contrast to most manufacturing changes, a biosimilar is always compared to the reference product in a thorough physico-chemical, structural, in vitro functional and human clinical comparability exercise, including relative immunogenicity. The developer of a biosimilar product has the major advantage of knowing the safety history of the reference product. Thus, the comparability exercise can be tailored according to the safety profile of the reference product.

The main concern in the interchangeability has been immunogenicity (Ebbers et al. 2012). As mentioned, very rare immunological problems have been encountered upon switches when the new product is inferior to the original product. It is also known that intermittent treatment of a given immunogenic therapeutic protein, such

as infliximab, may trigger delayed immunological reactions due to a temporary change in the immune complexes of the active substance and ADAs. In theory, such a reaction could be possible upon a switch that is delayed as compared to normal treatment schedule.

It has also been suggested that "drifting" or even single manufacturing change of a biosimilar or its reference product would make them non-interchangeable (Ramanan and Grampp 2014). It has also been argued that small differences between the biosimilar and its reference product might become clinically significant upon repeated switches (FDA 2017). For time being, these concerns are only theoretical. Nevertheless, they highlight the importance of a rigorous approach to comparability testing of both the reference and biosimilar products. There are no plans to introduce regulatory requirements for clinical interchangeability studies or automatic post-marketing comparability studies of biosimilars and their reference products in EU. Thus, there is only one level of biosimilarity in the EU.

It is important to note that, in the EU, interchangeability is a scientific/medical term that means changing one medicine for another that is expected to achieve the same clinical effect in a given clinical setting and in any patient *on the initiative, or with the agreement of the prescriber*. Substitution is a practical measure that is based on political and administrative decision to dispense a certain medicinal product instead of another interchangeable product. There is no EU-wide regulatory assessment of the interchangeability. Several EEA-countries have national regulatory positions endorsing the interchangeability of biosimilars. There are no plans in EU to introduce regulatory requirements for clinical interchangeability studies and, thus, create two levels of biosimilarity. From the scientific point of view, it is highly unlikely that two comparable biological products would not be interchangeable.

Product Information

The purpose of the European "Summary of product characteristics" (SmPC) is to provide information on the safe and rational use of the product to prescribers (European Commission 2017b). According to the current guidance, the SmPC of a biosimilar should contain the same information as the reference product SmPC with the exception of product-specific information (EMA 2017i). This is logical because the information in the SmPC of the reference product contains information of the long safety follow up and the pivotal efficacy and safety studies. The biosimilar comparability exercise is described in the European Public Assessment Report (EMA 2017j).

Information of immunogenicity is problematic since the information in the SmPC of the reference product can be based on data obtained with ADA assays that do not meet the current requirements. It is not uncommon that the incidence of ADAs is significantly lower in the SmPC of the reference product than in comparative studies of biosimilar development program. This can be explained by the higher sensitivity of current ADA assays. Therefore, the EU SmPCs of biosimilars will contain only an abbreviated description of the clinical significance of ADAs without exact numbers.

In the medical community, the introduction of biosimilars raised concerns that are partly due to the lack of understanding of the principles of the development of biosimilars. The divergence of the regulatory and medical communities is amplified by commercially tainted information on biosimilars. In the past, many learned societies in the EU discouraged the use of biosimilars. The questioning of the regulatory benefit risk assessment of biosimilars forced the EU Commission (Medicines for Europe 2017) and EMA (Vermeer et al. 2013) to intensify the distribution of unbiased information of biosimilars to prescribers and patients. This information and accumulating re-assuring clinical experience has narrowed the gap between regulators and medical community in Europe.

The most important product information for a prescriber in EU is the Summary of Product Characteristics (SmPC) that crystallizes the safe and rational use of a product. As biosimilar is a copy of the reference product, the SmPC of a biosimilar must be a copy of the reference product's SmPC with the exception of some product-specific data. Another exception is the information on immunogenicity. Detailed data on the ADAs in the SmPCs of the reference products is often outdated because of the improvement in assay methodology. In this situation, the immunogenicity information of the biosimilar is abbreviated. The data on the biosimilar comparability exercise is given in the European Public Assessment report.

Biosimilars have created an unusual situation for regulators since learned societies and patient organisations have initially discouraged the use of biosimilars and questioned the benefit/risk assessment of the EMA. These worries have been amplified by biased commercial information. Therefore, the EU Commission, EMA and national regulatory agencies have produced neutral information and collaborated with their stakeholders to clarify the controversies.

Post-Marketing Surveillance

Every new medicinal product in the EU, including biosimilars, has to have a risk management plan that includes a description of the applicant's pharmacovigilance system. In addition, it contains a list of potential and observed risks, and the measures to detect and mitigate these risks. Routine pharmacovigilance is mostly based on spontaneous reporting by health-care professionals and the patients. The package of all new products, including biosimilars will have a "black triangle" encouraging the user to report adverse events. Almost all biosimilars that were licensed in the EU have had non-standard risk detection and mitigation obligations, most often registries, for a particular therapeutic indication or risk.

The EU legislation requires that all adverse event reports of a biological product shall contain International Nonproprietary Name (INN), the brand name and the batch number. The EU is not in favor of the proposed four letter code to be attached to the INN. It is encouraging that the traceability of biosimilars has been outstanding in the EU (Vermeer et al. 2013). However, the batch numbers have not been given in the majority of the reports. This is a general problem for biologicals. It is expected that the traceability will improve in 2019 when the new safety features, including unique identifier is introduced to packages of all prescription medicines (European Commission 2017c).

In addition, the applicants are encouraged to participate in already existing pharmaco-epidemiological risk detection studies in place for the reference product. Ongoing risk minimisation activities agreed for the reference product should, in principle, also be included into the risk management programme of its biosimilar.

There is a trend to boost the routine pharmacovigilance and traceability of medicinal products, including biosimilars. All new biosimilars have a risk management plan with post-marketing measures that go beyond the routine pharmacovigilance, such as patient registers or additional clinical studies. In EU, it is not believed that the four letter codes attached to the INN would improve the reporting of possible adverse effects of biologicals. The real problem for all biologicals is the reporting of batch numbers. The situation may improve upon adoption of the new falsified medicines directive that introduces a unique identifier for each package of medicinal products.

References

van Aerts L, De Smet K, Reichmann G, et al. Biosimilars entering the clinic without animal studies: a paradigm shift in the European Union. mAbs. 2014;6:1155–62.

Danese S, Fiorino G, Raine T, et al. ECCO position statement on the use of biosimilars for inflammatory bowel disease—an update. J Crohn's Colitis. 2017;2016:26–34.

Ebbers H, Crow S, Vulto A, Schellekens H. Interchangeability, immunogenicity and biosimilars. Nat Biotechnol. 2012;30:1186–90.
EMA. History of EMA. 2017a. http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000628.jsp&mid=WC0b01ac058087addd. Accessed 24 Oct 2017.
EMA. Committee for medicinal products for human use (CHMP). 2017b. http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000094.jsp&mid=WC0b01ac0580028c79. Accessed 9 Dec 2017.
EMA. Multidisciplinary guidelines: biosimilar. 2017c. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000408.jsp&mid=WC0b01ac058002958c. Accessed 25 Nov 2017.
EMA. ICH guideline Q8 (R2) on pharmaceutical development. 2017d. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002872.pdf. Accessed 9 Dec 2017.
EMA. Remsima: public assessment report. 2017e. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_Public_assessment_report/human/002576/WC500151486.pdf. Accessed 9 Dec 2017.
EMA. Similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. EMEA/CHMP/BMWP/42832/2005 Rev. 1. 2017f. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001378.jsp&mid=WC0b01ac058002958c. Accessed 9 Dec 2017.
EMA. Guideline on Immunogenicity assessment of therapeutic proteins. 2017g. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/06/WC500228861.pdf. Accessed 9 Dec 2017.
EMA. Guideline on immunogenicity assessment of monoclonal antibodies intended for in vivo clinical use. 2017h. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128688.pdf. Accessed 9 Dec 2017.
EMA. QRD general principles regarding the SmPC information for a generic/hybrid/biosimilar product. 2017i. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2012/05/WC500127589.pdf. Accessed 9 Dec 2017.
EMA. European public assessment reports. 2017j. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124. Accessed 10 Dec 2017.
EMEA. Concept paper on the development of a committee of proprietary medicinal products (CPMP) guideline on the comparability of biotechnology-derived medicinal products. 1998. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003966.pdf. Accessed 23 Oct 2017.
EMEA. Guideline on the comparability of medicinal products containing biotechnology-derived proteins as active substance. Non-clinical and clinical issues. 2003. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003963.pdf. Accessed 23 Oct 2017.
EudraLex. Volume 1 - pharmaceutical legislation for medicinal products for human use. 2017. https://ec.europa.eu/health/documents/eudralex/vol-1_en. Accessed 9 Dec 2017.
European Commission. 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 (consolidated version). 2017a. https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir_2001_83_consol_2012/dir_2001_83_cons_2012_en.pdf.
European Commission. A guideline on the summary of product characteristics. 2017b. https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-2/c/smpc_guideline_rev2_en.pdf. Accessed 9 Dec 2017.
European Commission. Commission delegated regulation (EU)2016/161. 2017c. https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2016_161/reg_2016_161_en.pdf. Accessed 10 Dec 2017.

European Commission, DG Enterprise and Industry. What you need to know about biosimilar medicinal products. Process on corporate responsibility in the field of pharmaceuticals access to medicines in Europe. A consensus Information Document. 2013. http://europa.eu/geninfo/query/resultaction.jsp?query_source=GROWTH&QueryText=biosimilars&op=Search&swlang=en&form_build_id=form-CA4kk1hS9th2Qw_AyEvYWFKAUUlCNvY6kFflINvXlD4&form_id=nexteuropa_europa_search_search_form. Accessed 28 Oct 2017.

European Union. 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. 2017. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010L0063&from=EN. Accessed 9 Dec 2017.

FDA. Considerations in demonstrating interchangeability with a reference product. 2017. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf. Accessed 10 Dec 2017.

Glintborg B, Sorensen I, Loft A, et al. A nationwide non-medical switch from originator infliximab to biosimilar CT-P13 in 802 patients with inflammatory arthritis: 1-year clinical outcomes from the DANBIO registry. Ann Rheum Dis. 2017;76(8):1426–31.

Kurki P, Ekman N. Biosimilar regulation in the EU. Expert Rev Clin Pharmacol. 2015;8(5):649–59.

Kurki P, van Aerts L, Wolff-Holz E, et al. Interchangeability of biosimilars: a European perspective. BioDrugs. 2017;31:83–91.

Mack A. Norway, biosimilars in different funding systems. What works? GaBI J. 2015;4:90–2.

Medicines for Europe. Positioning statements on physician-led switching for biosimilar medicines. 2017. http://www.medicinesforeurope.com/wp-content/uploads/2017/03/M-Biosimilars-Overview-of-positions-on-physician-led-switching.pdf. Accessed 7 Dec 2017.

QuintilesIMS. The impact of biosimilar competition in Europe. 2017. http://www.medicinesforeurope.com/wp-content/uploads/2017/05/IMS-Biosimilar-2017_V9.pdf. Accessed 29 Oct 2017.

Ramanan S, Grampp G. Drift, evolution, and divergence in biologics and biosimilars manufacturing. BioDrugs. 2014;28:363–72.

Vermeer N, Straus S, Mantel-Teeuwisse A, et al. Traceability of biopharmaceuticals in spontaneous reporting systems: a cross-sectional study in the FDA adverse event reporting system (FAERS) and EudraVigilance data bases. Drug Saf. 2013;36:617–25.

Weise M, Kurki P, Wolff-Holz E, et al. Biosimilars: the science of extrapolation. Blood. 2014;124:3191–6.

Part III
Biopharmaceutical Development and Manufacturing of Biosimilars (Chemistry and Manufacturing Controls (CMC), Quality by Design (QbD), Quality Target Product Profile (QTPP))

Chapter 8
Biosimilars Drug Substance Development and Manufacturing: Effective CMC Strategy

Adriana E. Manzi and Michiel E. Ultee

Abstract The Chemistry, Manufacturing and Controls (CMC) section of a regulatory submission details aspects of the drug's chemical properties, its manufacturing process development and production, and control mechanisms, including analytical testing both in-process and of the finished drug, to show the process is controlled and reproducible. For a candidate biosimilar much is already understood about the chemistry and properties of the drug by thorough characterization of reference drug product. The challenge is to develop a manufacturing process with adequate controls to ensure that the biosimilar product closely matches the reference drug product and to demonstrate this biosimilarity with a strong analytical package.

Keywords Biosimilar development · Biosimilar manufacturing · Chemistry manufacturing controls · CMC

The Chemistry, Manufacturing and Controls (CMC) section of a regulatory submission details aspects of the drug's chemical properties, its manufacturing process development and production, and control mechanisms, including analytical testing both in-process and of the finished drug, to show the process is controlled and reproducible. For a candidate biosimilar much is already understood about the chemistry and properties of the drug by thorough characterization of reference drug product. The challenge is to develop a manufacturing process with adequate controls to ensure that the biosimilar product closely matches the reference drug product and to demonstrate this biosimilarity with a strong analytical package.

A. E. Manzi (✉)
Atheln, Inc., San Diego, CA, USA
e-mail: adriana@athelnbiomed.com

M. E. Ultee
Ulteemit BioConsulting, LLC, Hillsborough, NJ, USA
e-mail: mike.ultee@comcast.net

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_8

Understanding the Reference Drug

The first step in an effective biosimilar CMC strategy is to develop an extensive and sound understanding of the biochemical properties of the reference or innovator drug product, which for biosimilars we shall consider to be a recombinantly produced purified protein. The biosimilar developer lacks access to any proprietary information about the reference product, including for example development reports, batch records, and release specifications for the drug substance and product. Therefore, the developer must mine public sources of information, from information disclosed by the innovator through publication, presentation or information accessible from regulatory agencies. Additionally, a Certificate of Analysis is also available for some biologics reference drug products procured from certain geographical regions.

Generally, one can readily obtain basic aspects of the drug, such as the amino-acid sequence, type of product (IgG antibody, fusion protein, enzyme, etc), mechanism of action, dosage and formulation. Those with a background in development of similar types of protein products can identify likely Critical Quality Attributes (CQAs) that must closely match those of the reference product in order to avoid having any clinically meaningful differences from the reference product, the key test of biosimilarity. Some CQAs represent the protein functional aspects, such as an antibody binding to its target or an enzyme activity, which determine the mechanism of action of the protein. Others are structural, such as post-translational modifications (PTMs) like glycosylation, which can affect secondary aspects like effector function, as well as pharmacokinetic and pharmacodynamic (PK/PD) profiles. Other CQAs may represent the protein molecular state, for example aggregation or degradation, as well as secondary, tertiary or higher order structures that may impact biological activity. Aggregation and degradation are special concerns for protein therapeutics due to the higher immunogenicity of aggregated and certain degraded proteins.

Additionally, as with any biologic product, it is necessary to evaluate the process-related residual impurities such as host-cell proteins (HCP), host-cell DNA (HCDNA), and residual Protein A when Protein A chromatography is used in a process. HCP and Protein A impurities are generally detected by sensitive immunoassays (ELISA), while DNA is typically measured using the quantitative polymerase chain reaction (qPCR). In early development HCP can be detected using a standard assay for the HCP of a particular cell line, such as CHO cells. Later stage development needs to address the specific HCP produced by the cell line of interest. Such a cell-line specific ELISA is produced by generating detection antibodies against the non-transfected form of the cell line (sometimes called null-set cells) that had been used with transfection to produce the protein. These tests, together with the general safety assays (i.e. endotoxin, sterility) with appropriately selected acceptance criteria for the corresponding dose(s) and modes of administration will ensure the product safety of the biosimilar.

Any therapeutic protein represents a family of closely related variants, which all have the same amino-acid sequence but varying levels of modifications, either from variation within PTMs or due to degradations such as deamidation and oxidation. Thus, a typical monoclonal antibody shows at least 4–6 bands on isoelectric focusing, with each band representing a charge variant. All of these variants may have comparable activity such as target binding, but can vary in terms of effector function or PK/PD profiles (Schiestl et al. 2011). The biosimilar developer needs to understand the allowed range of these variants. This can be ascertained by extensive testing of many lots of the reference product, which will reveal the extent of variation for each CQA. How many lots? This is both a statistical and logistical question, and is addressed in more detail in Chap. 17. Recent approvals of some biosimilars suggest at least 20 lots should be assayed (Gray 2017), with more being better in terms of range setting. Testing of the originator lots may reveal considerable difference in one or more CQAs defining groups of lots with different profiles. This observation likely indicates that the innovator conducted a process changes(s), assessed comparability before and after the change(s), and obtained approval by the regulatory agencies for the new process/product. Furthermore, lots aged for varying lengths of time under the indicated storage conditions will reveal stability parameters.

Expression Systems and Clonal Selection

One has a choice of biological expression systems for production of a recombinant protein therapeutic, ranging from the simplest microbial systems to complex cellular systems such as mammalian cell lines. The simplest microbial cells, such as *E. coli*, can only express the protein without post-translational modifications (PTMs), such as glycosylation (see Table 8.1). For simpler proteins lacking such PTMs bacterial systems can be an attractive choice due to their rapid growth rate and productivity. They can present challenges in terms of recovery of the protein, which is typically produced in reduced, unfolded form in inclusion bodies. There are other bacterial systems that secrete the folded protein, such as that based on *Corynebacteria* (Ajinomoto), but again the protein would lack any PTMs. Higher order microbes such as yeast expression systems (i.e., *P. Pastoris* or *S. cerevisia*) will produce some PTMs, but usually in a more primitive form such as high-mannose glycosylation. For ready secretion of recombinant proteins with full PTMs, mammalian cell lines such as CHO, NS0 and PER.C6 have been the expression system of choice. Occasionally an alternative expression system such as one based on insect-cells (Protein Sciences) or even green plants (Medicago) is utilized for specialty products, such as vaccines (Fig. 8.1).

With these considerations in mind, the biosimilar developer would in most cases select a comparable system to what was used for the reference product. However, for simple proteins lacking PTMs, there could be economic advantages in going to a high-yielding microbial system. Secondly, in selecting a mammalian cell, attention

Table 8.1 Some common post-translational modifications (PTMs) (Walsh and Jefferis 2006)

Type	Amino-acid modified	Comments
Glycosylation, N-linked	Asparagine in consensus sequence Ser/Thr, X, Asn	Range from simple high-mannose to complex multi-antennary structures
Glycosylation, O-linked	Serine or threonine	Simpler structures than N-linked, but sites less predictable
Glycosylation, C-terminal linked	C-terminal amino acid linked to phosphoethanolamine	GPI-anchored proteins, typically enzymes
Phosphorylation	Tyrosine, serine or threonine, with others possible occasionally	Associated with enzymes
Hydroxylation	Proline and lysine	Associated with collagens
Gamma-carboxy glutamic acid	Glutamic acid	Associated with blood proteins involved in coagulation

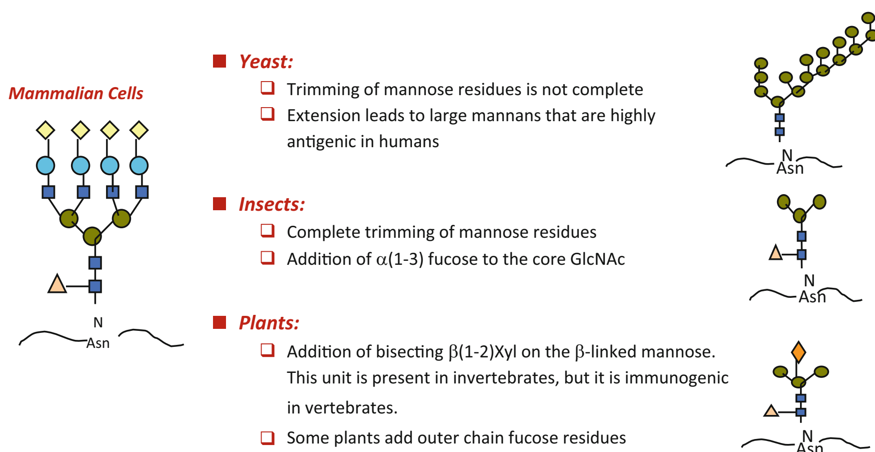


Fig. 8.1 Glycosylation control is critical for attaining biosimilarity

should be given to subtle differences in some PTMs between types of cells, such as the type of sialic acid added (NANA vs. NGNA) or types of linkages between sugars (Ghaderi et al. 2012). The key focus is to produce PTMs as close as possible to those of the reference product. In this regard, the safest approach is to use the same type of expression system, even down to the specific strain of mammalian cell (i.e., CHO-DG44 vs. CHO-S), as used for the reference product.

In selecting clones for production of recombinant proteins following transduction, one generally seeks those with the highest productivity. Productivity is a function of both the specific expression rate, typically expressed as picograms/cell/day (pg/c/d), and the viable cell density (vcd). Thus, high productivity may be due to

high specific productivity in spite of low cell density, or vice versa. The ideal would be for both pg/c/d and vcd to be high.

In development of a biosimilar, however, a new parameter enters into the selection process—how closely does the product from any given producing clone match that of the reference product? While this is mostly a function of PTMs, it may also be influenced by potential degradative processes in some clones, such as oxidation and deamidation. This new criterion supersedes productivity in terms of importance, since any biosimilar structural features must match as closely as possible to the reference product. Therefore, some highly productive clones may need to be passed over in favor of less productive ones that produce a protein more comparable to that of the reference product.

Finally, as with any development of clones for production of a recombinant protein, clonal stability needs to be assessed by repeatedly passaging the cells, typically for 20 passages or 50 generations, and measuring productivity and protein quality at selected time points during these passages. Some clones lose their ability to express a protein or alter their expression over time, and these need to be eliminated from consideration. With this consideration in mind it is recommended to select a primary clone and at least two back-ups for the clonal-stability study. Additionally, the product expressed by all selected clones should be evaluated and pass all criteria for the identified CQAs for the full culture period.

Development of Upstream and Downstream Processes and Scale-Up to Manufacturing

Biosimilar development follows the same path as new biopharmaceutical development in terms of upstream (culture) and downstream (purification) process development. Thus, following selection of a suitable clone, the goal is to develop consistent and high-yielding processes that produce protein with only trace levels of impurities. Upstream development is typically carried out in small bioreactors, ranging from multiple mini-bioreactors such as ambr® (Sartorius, see Fig. 8.2) or DasBox (Eppendorf) to 2–10 L benchtop bioreactors. By varying the culture media, feeds, and bioreactor parameters one seeks to produce the most protein in the shortest time. On the downstream side, the typical approach is to use a system of filters and chromatographic columns to purify the protein of interest by either binding/retaining it while impurities pass through, or vice versa.

However, there is one major difference for biosimilars—comparison to the reference product is needed throughout the upstream and downstream development program to ensure that the selected conditions produce protein highly similar to the reference product. Again, one may need to forego a high-yielding upstream process in favor of a lower-yielding one that produces protein more similar to the reference product. Similarly, some downstream process steps have the potential to induce structural changes in the protein, such as deamidation, which would produce

Fig. 8.2 Photo of ambr® mini-bioreactors in 15 mL (left) and 250 mL (right) sizes. (Photo courtesy of Sartorius Stedim Biotech & used with permission)

acidic variants that may pull the protein outside of the isoelectric-point range of the reference product. Using newer charge-based techniques (Müller-Späth et al. 2015), it may be possible to separate charge variants in order to remove undesirable species in terms of biosimilarity. Such approaches, however, would add complexity and additional processing compared to tailoring the clone and production process to produce material similar to the reference product.

Of course, should the new process yield a higher purity in terms of lower residual host-cell proteins, or lower levels of aggregation or product-related degradants, that would be beneficial since such components of the reference product are undesirable. Indeed, as bioprocessing techniques improve, increasing lower levels of such impurities are expected, and it may well be that the reference product has already seen several process improvements over its lifetime to produce purer material. Finally, a full viral-clearance study is required for a biosimilar downstream process just as it is for a new biopharmaceutical.

Scale-up of biosimilars follows the same path as new biopharmaceuticals except that one is constantly assessing the effect of scale-up on the similarity of the biosimilar to the reference product. Some factors to assess as bioreactor scale increases are levels of oxygenation and carbon dioxide, pH and nutrients/waste products, as well as rate and efficiency of mixing. These effects need to be assessed over the range of bioreactors that could be used in production. Following the initial small-scale bioreactors, a series of increasingly larger bioreactors are used and the product quality assessed along the way. As with new biopharmaceutical development, consistency and robustness in production and purification is the goal.

While it is possible to selectively remove undesirable charge variants in the purification process by specialized techniques like charge-displacement cation-exchange chromatography (Zhang et al. 2011), this is challenging given the subtleness of the variations, which could be only a single-charge change on a large protein. Furthermore, such an additional step in the downstream process would likely reduce yields significantly due to the need to closely fractionate similar species. Therefore, it is much better to select a cell line and growth conditions that result in a protein with a highly similar profile to that of the reference product.

Regarding product process qualification, process characterization and validation, these areas of development are similar to those needed for a novel biopharmaceutical. However, for a biosimilar these steps can be easier since product and process knowledge is more advanced due to the learnings from the reference product. As for control strategies, these would be particularly strict for a biosimilar due to the need to maintain biosimilarity to the reference product at each step of the production and purification.

Effect of Raw Materials on Product Quality

Sometimes overlooked but very important is the assessment of critical raw materials, such as culture media components, process buffering chemicals, and excipients on the quality and hence biosimilarity of a protein. For example, some specific culture-media additives can alter post-translational modifications, such as the addition of N-acetyl neuraminic acid (NANA) to block production of NGNA sialic acid (D. Ghaderi et al. 2012). Furthermore, sugars such as glucose used in culture media, or sucrose and trehalose used in final formulations, can be contaminated with endotoxins, as sugars are biologically derived. This will vary from lot-to-lot and must be assayed, since endotoxins are a highly undesirable impurity. The careful consideration of critical raw materials continues through formulation development. For example, the effect of oxygen to produce peroxides in solutions of polyethers, such as polysorbates (Tween®) is well-known. Such peroxides are highly reactive and will degrade the protein. These can be controlled through a combination of sourcing ultra-high purity polysorbates meeting tight specifications for low-levels of peroxide, drawing only from previously unopened containers, and using fresh solutions. Furthermore, as with any new biopharmaceutical, one should seek to avoid animal-derived raw materials whenever possible, and if not, perform a risk-assessment on the potential for contamination from sources of transmissible spongiform encephalopathies (TSE). Other impurities in raw materials can include transition metal ions that can act as catalysts for the oxidation of sensitive amino acids such as methionine.

With the advent of single-use technologies utilizing plastics, a new potential source of impurities is present due to leachables and extractables from the plastics. This is particularly a concern with plastic bags used widely throughout the production process, from bags used to contain the cells and media of production to other

bags for storage of buffers and both intermediates and bulk drug substances. The plastic materials must therefore be rigorously tested for leachables and extractables, and their effect evaluated. The United States Pharmacopeia has developed a rigorous series of tests for such materials for medical use to qualify for USP's Type VI classification, which means that they are biologically compatible and suitable for even in vivo use in patients. Manufacturers of single-use systems have had to source very high purity plastics for their products, since previously these plastics were primarily used in industries not requiring such purity, such as the automobile industry. Customer of single-use materials should request leachable and extractable data from the supplies as well insist that they be made from USP Type VI materials.

Testing and control of residual process impurities such as host-cell proteins (HCP), host-cell DNA (HCDNA), and protein chromatographic ligands such as Protein A is the same for biosimilars as for novel biopharmaceutical. Levels of such residuals, which represent a potential safety concern, should be comparable to or lower than those of the reference drug since the biosimilar developer is relying on the positive safety history of the reference product.

Effect of Manufacturing Conditions on Product Quality and Control Strategy for Biosimilars

As described above, a tight control of the CQAs is required throughout the development of upstream and downstream processes to ensure that the selected conditions do not induce structural changes in the protein that may pull the target product outside of the range of the reference product. Many of the analytical methods ensuring the maintenance of biosimilarity continue to be used as in-process controls.

For biosimilars, as well as for any biologic, specifications need to be set following ICH Q6B. Some of the parameters tested and their acceptance criteria will reflect the specific process used for the manufacturing of the biosimilar product and as such differ from the originator (i.e. residuals from materials used in upstream media or downstream purification steps, HCP, residual host cell DNA, etc.).

The analytical comparability studies will guide the establishment of specifications for the CQAs. The principles discussed in Chaps. 11 and 17 can guide the decisions as to the acceptable ranges in the acceptance criteria set for each CQA to ensure biosimilarity.

Specifications will vary depending on the type of molecule (i.e. recombinant protein vs monoclonal antibody). For glycosylated molecules, understanding the impact of glycan composition on potency is important since variability is expected lot-to-lot. Variation on the level of sialylation can influence the PK profile of a molecule. For mAbs, oligosaccharide core-fucosylation impacts the Fc conformation and can lower the binding affinity of the molecule to Fc receptors. Therefore, it is important to monitor these structural features and ensure process controls are sufficient to

Table 8.2 Generic specifications for a biosimilar BDS

Parameter	Quality attribute	Method
Identity	Primary sequence	Peptide mapping by RP-HPLC/UV
Content	Protein concentration	UV Spectroscopy at 280 nm
Purity	Size variants	SEC-HPLC with UV detection Capillary electrophoresis—SDS (CE-SDS)
	Charge variants	Capillary isoelectric focusing (cIEF) Hydrophobic interaction chromatography
	Glycosylation	Oligosaccharide mapping/normal phase HPLC with fluorescence detection Sialic acids by RP-HPLC with fluorescence detection
Potency	Binding affinity/ADCC/CDC/ proliferation, etc.	Specific assays
Process-related impurities	Residual process impurities	Specific to process (i.e. residual protein A)
	Residual host cell proteins (HCP)	Specific to expression system (i.e. 2-D Western blot overlaid with antibody for specific cell line)
	Residual host cell DNA	qPCR

maintain them within the acceptable range. For a monoclonal antibody, control of amino-acid modifications that may impact product potency may also be required (i.e. methionine oxidation, asparagine deamidation, etc.).

An important aspect of setting specifications is the evaluation of stability behavior for the biosimilar candidate in comparison to the reference material. Degradation profiles obtained during side-by-side, forced-degradation studies can uncover differences in behavior between the biosimilar candidate and reference material, such as oxidation due to metal catalysis from process residuals, requiring additional process controls. Typical forced-degradation conditions include exposure to high temperature, light, low and high pH, and oxidizing conditions. Their impact on the protein in terms of accumulation of low and high molecular weight species (SEC-HPLC; reducing and non-reducing CE-SDS), changes in charged species (CEX-HPLC), conformational changes (CD), potency and protein concentration (UV) should be evaluated for biosimilar candidate and reference material.

Additionally, the functional assays should be developed as early as possible to measure biological activity ensuring that biosimilar candidates are effective and suited for moving along the development process. As example, the functional assays used for testing monoclonal antibodies are described in Chap. 16. One or more of these assays will also become part of the specifications after appropriate validation to ensure assay performance.

A generic list of specifications for a biosimilar BDS is presented in Table 8.2.

As the number of lots manufactured using the same process increases, results for all the parameters included in the specifications as well as in-process and stability

results, will create a database in support of process validation and enable to set robust specifications ranges for commercial lots.

Implementing a Successful CMC and Analytical Strategy for the Development of Biosimilars

The Analytical Strategy is the most critical component of planning a biosimilar development program. A chapter has been dedicated in this book (Chap. 11) to defining the principles of analytical biosimilarity that enable a carefully designed analytical plan. This section focuses on key factors to implement a successful CMC strategy fostering the constant collaboration between process development, manufacturing and analytical scientists throughout the development program of each biosimilar candidate.

The development of a biosimilar product requires a complex set of CMC elements that need high level of expertise and investment. In particular, biosimilar development is heavily frontloaded on analytics. This shifts the traditional pattern of investment during the development of *de novo* biologics that delays to later stages the considerable costs of equipment/ expertise and/or outsourcing of state-of-the-art testing. Instead early investment in this area is critical for success when developing a biosimilar product. Analytical biosimilarity of the candidate to the reference product is the first layer of biosimilarity demonstration. It needs to be monitored and confirmed at every step of the development path to ensure success (Fig. 8.3).

A biosimilar protein is, by definition, related to the originator drug (reference product) by a common primary amino-acid sequence. Therefore, it is imperative to first ensure this condition is met by the biosimilar candidate. It is recommended that the biosimilar developer confirms the amino-acid sequence of the reference product by direct analysis to circumvent any potential misrepresentations in the public information.

In addition, all therapeutic proteins, including originator drugs, are a population of product variants. The key is to establish the characteristics of the API in the reference product to define the *Quality Target Product Profile (QTPP)* for the biosimilar (Bui et al. 2015). *The QTPP is defined as a combination of Critical Quality Attributes (CQA's).* Based on ICH guidelines (Annex of ICH Q8), "A CQA is a physical, chemical, biological or microbiological property or characteristic that should be within an appropriate limit, range or distribution to ensure the desired product quality".

The CQA's, their values and ranges that define the biosimilar QTPP are established from testing of the reference product as indicated earlier. A thorough characterization of the API in the reference product using a panel of physico-

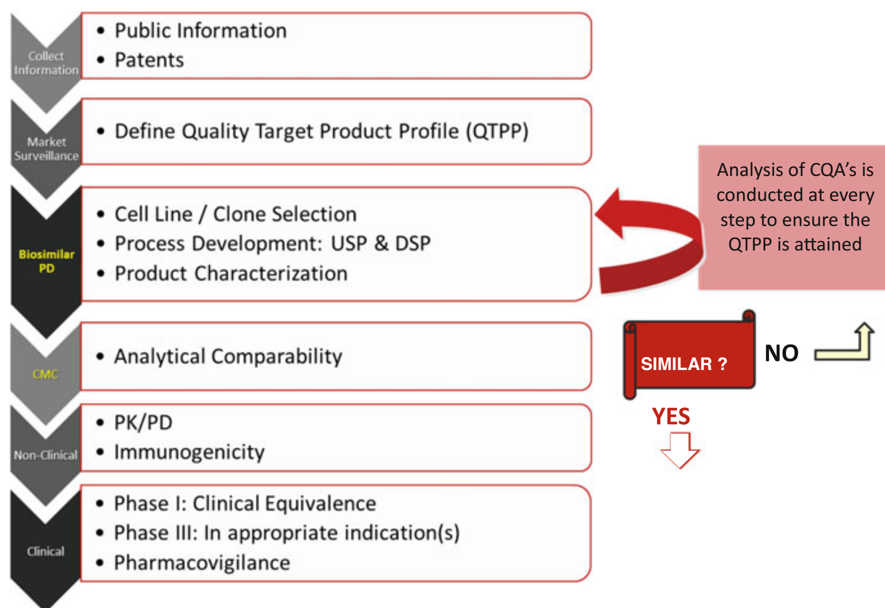


Fig. 8.3 Analytical strategy for developing a biosimilar

chemical methods (such as peptide mapping by LC-MS and including the analysis of PTMs) and biological assays (i.e. ADCC, CDC, others) originates a database to compare the biosimilar candidate. The surveillance of the reference product occurs continuously in order to form a full view of originator product variability in CQA's. This includes variability derived from process changes incorporated by the originator and approved based on comparability exercises, as well as stability changes within the approved parameters. Since the reference product is only accessible to the biosimilar developer in its final formulation, testing requires methods that are not sensitive to the composition of the formulation or extraction of the API from the formulation avoiding changes to its structure and/or properties.

The assessment of CQA's does not require a full validation of the analytical methods employed as per ICH Q2(R1). However, it is in the best interest of the biosimilar developer to use methods providing reproducible and reliable results from very early stage. It is recommended at a minimum to qualify the methods by evaluating the critical parameters for the intended purpose. Additionally, using orthogonal analytical methods based on different principles can typically aid in detecting small differences in molecular variants.

As indicated earlier, CQA's need to be evaluated when selecting an expression system, even for the individual clones, to ensure that biosimilarity is attainable. Evaluating some of these parameters (i.e. glycosylation) at the clone level requires the adaptation of the methods typically used for the analysis of late-stage, highly

purified samples to small quantities of cell culture supernatant. Additionally, testing of many samples in parallel (i.e. comparing clones, culture conditions, etc.) implies considerable dedication of resources.

Consider for example that evaluating the glycosylation produced by different clones implies obtaining the glycosylation pattern of major N- and O-glycan species and evaluating the differences observed in minor glycans present. This requires enzymatic or chemical release of oligosaccharides followed by fluorescent labeling and HPLC with fluorescence detection), identifying the individual oligosaccharides species i.e. by LC-MS of the released and tagged glycans, and estimating the glycosylation site occupancy (i.e. by CE-SDS and/or LC-MS). Once a clone(s) are selected, the effect of culture conditions and media additives, culture age, etc. should also be evaluated by testing the same glycosylation aspects. It is well known that glycosylation can be affected by these parameters and any changes may result in lack of biosimilarity.

Similarly, the assessment of biosimilarity at different steps in the development of the DSP purification needs to cover all the CQAs to ensure the biosimilar molecule obtained through the optimized USP is neither degraded nor the variant composition is changed (Chaudhari et al. 2017). As with any biologic product, a critical aspect of this evaluation is the assessment and control of aggregation.

The full comparability of the biosimilar candidate with the reference product should also include physico-chemical methods to evaluate the secondary, tertiary and higher order structures of the molecule that impact its biological functionality. A detailed evaluation of these methods can be found in Part IV of this book.

The dossier associated with a biosimilar regulatory filing requires a substantial analytical package—351(k), significantly larger than one supporting a new biologic filing—351(a) (Fig. 8.4; Ha and Kornbluth 2016). A solid analytical package demonstrating the candidate is highly similar to the reference product lowers the potential for clinical differences. The analytical similarity data is used to determine the extent and design of the non-clinical and clinical studies required.

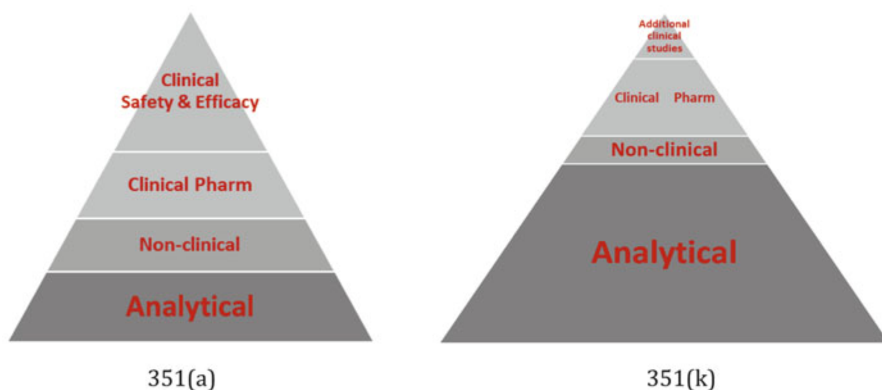


Fig. 8.4 Regulatory dossier for a new biologic [351(a)] compared to a biosimilar [351(k)]

As with the development of the bulk biosimilar, the analytical strategies for selecting the formulation and manufacturing process used for the drug product need to be supported by a strong analytical package. The closer the formulation chosen for the biosimilar matches the innovator formulation, the lower the chances of an impact on stability. However, changes in formulation may be preferred in order to generate IP and this is feasible as long as they come supported by a strong analytical package. More details on this topic can be found in Chap. 11 of this book.

The manufacturing of the final dosage form of a biosimilar does not differ from what is used for any biologic. Regularly, the selection of process steps, conditions such as temperature, time, holding times, etc. is done to ensure lack of impact on the characteristics of the product. Additionally, shipping conditions for bulk drug substances are controlled and testing of CQAs conducted at both ends to ensure no impact in biosimilarity (i.e. shear forces may impact aggregation). For biosimilar manufacturing, however, the level of control needs to ensure that no CQA is affected by the manufacturing and shipping selections beyond the acceptability ranges.

Conclusion

Development of a biosimilar is in some ways easier and other ways more challenging than a new biopharmaceutical. It is easier in that the reference drug is already well established and understood, and mechanism of action and indications known. It is more challenging in that at each step of the development program a biosimilar must be rigorously compared to the reference drug product. In a complex glycoprotein, multiple variants constitute any particular product and their ratios can vary from lot to lot. A thorough understanding of the range of this lot-to-lot variation, as well as the routes to degradation over time, is essential to developing a biosimilar that is indeed highly comparable to the reference product.

The constant collaboration between process development, manufacturing and analytical scientists throughout the development program of each biosimilar candidate is key to a successful outcome. Attaining the combination of quality attributes defining the targeted product requires constant verification as the candidate development progresses.

References

Bui LA, Hurst S, Finch GL, Ingram B, Jacobs IA, Kirchhoff CF, Ng C-K, Ryan AM. Key considerations in the preclinical development of biosimilars. Drug Discov Today. 2015;20(51):3–15.

Chaudhari PS, Nath R, Gupta SK. Opportunities and challenges in biosimilar development. Bioprocess Int. 2017;15(5):24–33.

Ghaderi D, Zhang M, Hurtado-Ziola N, Varki A. Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. Biotechnol Genet Eng Rev. 2012;28:147–75.

Gray T. Biosimilar analytical similarity assessments: are you testing sufficient innovator lots? Biosimilar Dev. 2017. https://www.biosimilardevelopment.com/doc/biosimilar-analytical-similarity-assessments-are-you-testing-sufficient-innovator-lots-0001

Ha CY, Kornbluth A. A critical review of biosimilars in IBD: the confluence of biologic drug development, regulatory requirements, clinical outcomes and big business. Inflamm Bowel Dis. 2016;22(10):2513–26.

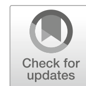
Müller-Späth T, Ulmer N, Aumann L, Bavand M. Twin-column cation-exchange chromatography for the purification of biomolecules. Biopharm Int. 2015;28(4):32–6.

Schiestl M, Stangler T, Torella C, Čepeljnk T, Toll H, Grau R. Acceptable changes in quality attributes of glycosylated pharmaceuticals. Nat Biotechnol. 2011;29(4):310–2.

Walsh G, Jefferis R. Post-translational modifications in the context of therapeutic proteins. Nature Biotech. 2006;24(10):1241–52.

Zhang T, Bourret J, Cano T. Isolation and characterization of therapeutic antibody charge variants using cation exchange displacement chromatography. J Chromatogr A. 2011;1218(31): 5079–86.

Chapter 9
QbD in Biopharmaceutical Manufacturing and Biosimilar Development



Christina Vessely and Christopher Bussineau

Abstract Over the last ten years, the development of biosimilars has transitioned from concept into approved products. The mechanisms of action of the molecules had been proven, clinical efficacy and safety profiles established, with large markets and sales margins, making them attractive targets for many biopharmaceutical companies, both large and small. However, inherent properties of the molecules result in higher levels of risk in the eyes of regulatory bodies. Therefore, a major component of the demonstration of biosimilarity is the thorough analytical characterization of the biosimilar in comparison to the reference product. The establishment of analytical biosimilarity can reduce the number of clinical studies required to support product approval.

The application of Quality by Design (QBD) in early product stages can both reduce risk to patients and streamline the development path for any biologic. The concepts become even more critical with the development of biosimilar molecules, where decisions that are made at very early stages with respect to cell lines, fermentation parameters, and purification strategy will impact the chemical and physical properties of the product. A thorough analytical tool kit is key to establishing critical quality attributes for the product across fermentation and purification development.

Critical quality attributes for a biosimilar must take into account both the knowledge gained during early manufacturing and formulation development work and also properties of the reference product. Characterization of the reference materials must occur in parallel with process development. The application of more advanced analytics during early development will allow for the selection of clones that are closest to matching the reference product in terms of glycosylation patterns. Information on the impurity profile of the reference product can be utilized to

C. Vessely (✉)
Analytical and Formulation Development, Biologics Consulting Group, Inc., Alexandria, VA, USA
e-mail: cvessely@biologicsconsulting.com

C. Bussineau
Manufacturing and Regulatory, Biologics Consulting Group, Inc., Alexandria, VA, USA
e-mail: cbussineau@biologicsconsulting.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_9

establish the target product profile as well as to gain an understanding of sensitivities of the molecule to different stresses that may be encountered during processing.

The application of QBD does not end at the drug substance stage, but should also be applied during formulation development and also in the selection of container/closure components. Understanding the sensitivity of the product to certain stresses should influence decisions for fill/finish processing equipment as well as placing limitations on exposure to light and time out of refrigeration.

The ultimate goal is patient safety, and in the world of Biosimilars, that means applying the correct tools to ensure that the product will match the reference product.

Keywords Quality by Design (QbD) · Critical quality attributes (CQA) · Quality target product profile (QTPP) · Design space · Risk assessment · Control strategy · Design of experiments (DOE)

Introduction

Over the last 10 years, the development of biosimilars in the United States has transitioned from concept to the approval of multiple products. These molecules are attractive targets for many biopharmaceutical companies, both large and small, since the mechanisms of action have been proven, clinical efficacy and safety profiles have been established, and large markets and the potential for reasonable sales margins exist. However, inherent properties of the molecules result in higher levels of risk to patient safety in the eyes of regulatory bodies. Unlike their small molecule counterparts, it is not possible to create a chemically/biologically equivalent molecule, due to differences related to host cells, processes, etc., hence the use of the term “biosimilar” instead of “biogeneric”. As a result, a major component of the demonstration of biosimilarity is the thorough analytical characterization of the biosimilar in comparison to the reference product. The establishment of analytical biosimilarity can reduce the number and complexity of clinical studies required to support product approval.

The application of Quality by Design (QbD) in early product stages can both reduce risk to patients and streamline the development path for any biologic. QbD principles become even more critical with the development of biosimilar molecules, where decisions that are made at very early stages with respect to cell lines, fermentation parameters, and purification strategy could significantly impact the chemical and physical properties of the product. A thorough analytical tool kit is key to establishing critical quality attributes for the product across fermentation and purification development.

Critical quality attributes for a biosimilar must take into account the knowledge gained during early manufacturing and formulation development work, as well as the properties of the reference listed drug (RLD). For this reason characterization of the RLD must occur in parallel with or even prior to process development of

the biosimilar. For example, the application of advanced analytics during early development of the drug substance (DS) will allow for the selection of clones that are closest to matching the reference product in terms of impurity profiles and/or product glycosylation patterns, thereby increasing the probability of successfully achieving biosimilarity of the drug product (DP). Information on the impurity profile of the RLD can be utilized to establish the target product profile of the biosimilar, as well as to gain an understanding of sensitivities of the molecule to different stresses that may be encountered during processing. This information can be especially important since samples of the innovator DS are not as available as the formulated RLP (DP).

Of course, the application of QbD does not end at the DS stage but should also be applied during formulation development and in the selection of container/closure components. Understanding the sensitivity of the product to certain stresses should influence decisions for fill/finish processing equipment as well as placing limitations on exposure to light and time out of refrigeration (TOR) to minimize any negative effects of storage on the drug.

The ultimate goal is patient safety, and in the world of biosimilars, that means applying the correct tools to ensure that the product will match the RLD.

What is QBD?

The term Quality by Design (QbD) is used frequently by product developers and agency reviewers, but what does it actually mean, and what is its value? The International Conference on Harmonization (ICH) defines QbD as a systematic approach to development that begins with predefined objectives and emphasizes product understanding and process control based on sound science and quality risk (ICH 2009). The goal is to ensure that quality is established in the product from its inception in a more holistic manner by looking beyond the individual processes that are involved in the manufacture of a drug product, into the entire development program.

Historically, the approach was to initiate biopharmaceutical product development following the discovery of a molecule believed to offer some clinical benefit (proof of concept). Preliminary development efforts centered around the need for sufficient material with which to perform toxicology studies and efficacy studies in animal models. Once data was gathered to support product safety and efficacy, processes were developed further and scaled up to supply clinical trial materials (CTM) for first-in-human studies. As clinical studies progressed into further stages of development, greater quantities of CTM were required and manufacturing scale up proceeded as required. During the clinical development phases, analytical methods were developed to evaluate product safety, identity, strength, potency, and quality (SISPQ) in a stage-wise process. Specifications were developed to define acceptable levels for each of those attributes, again in a stage-wise process, keeping pace with the increased knowledge of the product gained from more sophisticated analytical

tools. Assuming all specifications were met for each analytical release method, the resulting product was accepted for use. Process parameters were tightly controlled to maintain product consistency, and any deviations from that consistent process were fully investigated retrospectively to allow for batch release. In cases where the product did not meet the predefined specifications, the material could be reprocessed or rejected. As clinical development progressed into commercialization this strategy led to increased cost when failures occurred. Repeated failures with no assignable or predictable root cause raised concerns by regulatory agencies with respect to the overall quality and safety of the product.

Where QbD differs from the historical approach is that it begins with predefined objectives and requires an understanding of the influence of variations that may occur in the production process on product quality. The key premise of QbD is that quality cannot be tested into a product. Testing performed at the end of the process is expected to confirm that all quality attributes have been met, but it does not speak to product consistency or process control (Yu 2008). A high-quality product has been defined by Janet Woodcock as "a product free of contamination and reproducibly delivering the therapeutic benefit promised in the label" (Woodcock 2004).

It is critical that QbD efforts follow a scientific approach and reflect the degree of risk associated with changes to a product or process. The evaluation of a risk profile allows for a more deliberate progression through process and product development based on potential impact. QbD is not an exercise performed as a single event. It is intended to initiate at the pre-clinical or early clinical stage and continue through commercialization as a part of the product lifecycle.

The QbD approach is more than just a box checking exercise intended to appease agency reviewers at time of regulatory submission. It offers many advantages to the product developer as well. QbD promotes a more complete understanding of the impact of process variability on product quality thereby highlighting the most critical points of the process. By understanding which factors have the greatest impact on overall product quality, sufficient process controls can be implemented which will lead to increased success rates in manufacturing. Optimization for process robustness at the most critical steps leads to a higher probability of success during process scale-up and/or process transfer. Additionally, the application of QbD enables regulatory flexibility. The application of process understanding in the design of acceptable operating ranges (the design space) can streamline post-approval process changes that are made within those ranges.

The elements of QbD have been defined in ICH (Q8) (ICH 2009) and are further discussed by Yu et al. (2014). These elements include the quality target product profile (QTPP), critical quality attributes (CQAs), the performance of risk assessments to provide the connection between critical material attributes (CMAs) with critical process parameters (CPPs), and the control strategy that includes both the predefined acceptance criteria for the DS, raw materials, and drug product as well as in-process controls. Also key to QbD are the concepts of process capability and continuous improvement. The compilation of all of these elements leads to the definition of the design space for the process. By defining the design space, it becomes easier to determine which modifications to process parameters would

be considered a major change to the current process, requiring regulatory approval prior to implementation.

The Quality Target Product Profile

The first element of QbD is the definition of the product, also known as the Quality Target Product Profile, or QTPP. The QTPP is a prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account the safety and efficacy of the drug product (ICH 2009).

The QTPP includes a definition of all parameters of the intended drug product, including the dosage form, strength, route of administration, intended indication, etc. In addition to the definition of the product itself, the QTPP includes the drug product quality criteria for the intended marketed product, in terms of purity, stability, potency, safety, etc.

For a novel product, the definition of the product can be quite complicated. It requires an understanding of the disease state, tolerance of the patient population to particular dosing regimens, knowledge of the molecule of interest and allowable administration routes, estimation of efficacious dosage in the intended patient population, compatibility of the product with different container/closure systems, etc. Often the product definition will evolve as more knowledge is gained about the product during the development lifecycle.

For a biosimilar product, this first step of the QbD process is simplified because the product has already been defined by the originator. As such, the dosage strength, route of administration, container/closure system, and other physical parameters can be viewed as pre-established attributes.

Critical Quality Attributes

The QTPP drives the preliminary definition of product critical quality attributes (CQAs). A CQA is defined by the ICH as a physical, chemical, biological, or microbiological property or characteristic of an output material including finished drug product that should be within an appropriate limit, range, or distribution to ensure the desired product quality (ICH 2009). CQA's may be established to ensure that a parameter on the QTPP has been met. For example, if the QTPP requires a 30 mg dose, then content may be set to 30 mg $\pm 10\%$. If a shelf life of 24 months is required to support market and supply chain needs, additional CQA's must be applied to demonstrate product quality upon storage. The criticality of a product attribute is determined based on its propensity to cause harm to the patient should the product fall outside the defined range for that attribute.

In the development of a biotechnology product, there are several categories of critical quality attributes that are most often included in the release profile for a product. Those include quality, assay, potency, purity, impurities, and safety. A typical release panel for an early stage product will include an evaluation of

appearance, purity, one to two assays for the determination of product related impurities, as well as assays for process related impurities such as residual host cell proteins or residual DNA. The release panel will also include the determination of microbial product attributes such as endotoxin levels, as well as the presence of bioburden (in DS) or the confirmation of sterility (in DP).

As previously stated, however, the CQAs are not just acknowledged and tested for at the end of DS or DP manufacture as a specification. Instead, they can and should be used to inform on process development decisions. The potential CQA list is a document that evolves over the development life cycle and may experience both the addition and removal of specific CQAs from the list based on knowledge gained on the product and process. In the case of biotechnology products, the potential CQA list for DS tends to be more extensive than that for DP (Sangshetti et al. 2017). This is due to the fact that there are considerably more steps involved in the synthesis, harvesting, and purification of the DS, and because the majority of the potential CQAs are a consequence of imparting the necessary quality at each step of the DS manufacturing process.

Risk Assessments

Risk assessments are mentioned in the ICH Q8 (R2) guidance as providing the linkage between material attributes and process parameters to drug product CQA's (ICH 2009). Risk assessments are described in significant detail in ICH Q9 (ICH 2005), and are used to determine which of the material attributes and process parameters are most likely to have an impact on product quality, as demonstrated by the CQAs. The intent of the regulatory agencies is that companies will perform risk assessments early in the product development lifecycle, and that the risk assessment exercises will be repeated as needed based on knowledge gained during process development or based on the need for process changes to accommodate increased manufacturing scale or other challenges.

There are three basic steps that are performed as part of the preliminary risk assessments. Those include risk identification, risk analysis, risk evaluation. Risk identification is the process of evaluating a manufacturing process or product, or potential changes to that process or product, to determine the hazards that may result from those activities. In the identification process, scientists should take into account historical trends and/or lessons learned, industry best practices, as well as subject matter expert experience and opinion. The goal at the first stage is to select all possible risks and consequences.

Risk analysis follows the identification of risks and is intended to put a qualitative or quantitative value to the risk in terms of likelihood of occurrence and severity of impact. In some cases, the ability to detect the hazard is part of the evaluation and scoring process, as an easily detectable hazard is much more likely to result in batch rejection and therefore offers additional protection to consumers compared to a difficult to detect hazard. It is critical during the risk assessment process that participants acknowledge that there are assumptions and uncertainties inherent to

the risk assessment process. Typical sources of uncertainty may relate to lack of process and product experience, both in manufacturing and in the clinical setting.

Risk evaluation is a summary of the information determined in the risk identification and risk assessment steps. The values applied in the risk analysis lead to a quantitative estimate or ranking of risk for each process input or parameter. The risk evaluation is generally the product of the risk components for a given parameter. The quantitative estimate may be then translated into qualitative risk descriptors such as "high", "medium" or "low" (ICH 2005). When quantitative values are used for each of the components, it is best if they differ significantly for different risk levels. For example, instead of rating low, medium and high risks for a parameter as 1, 2, and 3, it's much better to use values such as 4, 40, and 400 to represent low, medium and high respectively. Having an order of magnitude difference between risk levels improves resolution, making it much easier to categorize the overall risk as low, medium or high.

The initial risk assessment exercise may result in a very extensive list of potential parameters for evaluation. Through process experience and the use of specific tools such as design of experiments or the use of scale-down models to evaluate specific manipulations, the significance of each parameter can be further elucidated, allowing for less critical parameters to be removed from the evaluation in subsequent risk assessment exercises.

There are multiple tools available that may be useful in the performance of risk assessments, many of which are described in the risk management guidance document (ICH 2005).

Design Space

ICH Q8 (ICH 2009) describes the design space as the relationship between process inputs, which include both the attributes of material inputs as well as the process parameters, and critical quality attributes. The design space includes an evaluation of the potential variables within the manufacturing process to determine the impact of variation of those parameters on product quality. The design space defines ranges around each of those variables such that as long as the process is operated within those ranges, consistent product quality can be achieved.

In most cases, the data set that defines the design space is achieved through intentional process variations that are performed as part of process characterization and process performance qualification (PPQ). One Factor at a Time (OFAT) studies, in which all process parameters except one are held constant from one run to the next, are useful for gaining process understanding and a preliminary determination of acceptable ranges for a given parameter. However, it is rare in practice that only a single parameter would deviate from the center point of its intended operating range. As such, a more sophisticated Design of Experiments (DOE) approach or other statistically derived experimental design may be required. The DOE approach allows for the determination of the impact of not just a single parameter on the overall product quality. DOE provides statistically based evidence of which factors

interact with one another resulting in a potentially greater impact to product quality than either factor would have in isolation.

Although much of the discussion within ICH Q8 is focused on process evaluation, it is important to note that the determination of design space also requires a thorough understanding of analytical methodology and method capability. Methods that are not capable of detecting critical impurities will not provide an adequate definition of the design space. Conversely, a lack of understanding of variability inherent to a particular analytical method may result in an unnecessarily restrictive design space.

There are two major components to the design space; the product design and the process design. The product design includes the parameters that were determined as part of the product definition (i.e. physical, chemical and biological attributes of the product itself as captured in the QTPP). The product design should also include an understanding of variability of these physical characteristics on product quality and stability. For example, the impact of excipient type and quality, influence of solution pH, and exposure to certain stresses during processing and storage. Thorough formulation studies allow for the scientist to become intimately familiar with the impact of variations in physical and chemical attributes during long term storage. Degradation pathways can be minimized or even eliminated when the formulation is optimized. The excipients can also impact product identification or patient acceptability (Yu et al. 2014).

The formulation optimization includes not just the selection of the most appropriate type of excipient or stabilizer, but also the determination of suitable quality for these items (Nazzal et al. 2002; Awotwe-Otoo et al. 2012). As an example, surfactants are often included in protein formulations to prevent agitation induced damage. However, surfactants are notorious for contamination with peroxides (Ha et al. 2002; Singh et al. 2012; Herman et al. 1996). Choosing a poor quality surfactant can result not only in a lack of protection from interfacial damage, but can promote chemical degradation of the product as well.

The process design component involves the evaluation of the unit operations that are utilized for the manufacture of a product. A unit operation is defined as a discrete activity that involves physical or chemical changes to the product (Little 2013). The process for the manufacture of a biotechnology involves multiple steps or unit operations. The design space may be determined for each individual unit operation or may include a combination of multiple unit operations (MacGregor and Bruwer 2008).

A process is generally considered well-understood when (1) all critical sources of variability are identified and explained, (2) variability is managed by the process, and (3) product quality attributes can be accurately and reliably predicted (U.S. Food and Drug Administration CDER 2004). Process parameters consist of both the operating parameters (input process parameters) and process state variables (parameters that are changed based on the input process parameters). Process parameters whose variability have a direct impact on overall product quality are considered as critical process parameters (CPP's) and should be both monitored and controlled to assure that product will meet acceptable quality standards.

The consistency in product quality depends not only on control of the process, but also the robustness of the process. Robustness is defined as the ability of a process to demonstrate acceptable quality and performance while tolerating variability in inputs (Goldek et al. 2006). Robustness is a function of both formulation and process design.

From a practical standpoint, the establishment of the design space does more than just allow us a comfort zone in which to operate our process. It provides a definitive guidance as to whether changes in process parameters constitute a major change to the process (thus immediately reportable to the agency, or even requiring preapproval in the case of late stage and/or marketed products), vs. a minor change that would be included in an annual report. This helps us to understand also the level of risk to product quality incurred resulting from a specific process change, as well as knowing whether a formalized comparability exercise will be necessary to justify that the post-change material is at least as good (safe, pure, effective) as the pre-change materials.

It should be noted that the design space is not simply a set of parameters that is defined by the product sponsor. The design space is proposed to the agency by the applicant and requires agency assessment and approval.

Control Strategy

Once the process has been adequately designed and defined, critical quality attributes determined, and the design space established, it is necessary to implement systems to maintain that design space, thereby facilitating consistency in product quality. Aspects of the control strategy include, but are not limited to, raw material specifications, in-process controls, intermediate product testing, and finished product testing.

The control strategy is defined by the ICH as a planned set of controls, derived from current product and process understanding that ensures process performance and product quality (ICH 2009, 2012). Some aspects of the control strategy may rely on real time monitoring of CQA's through the implementation of process analytical technology (PAT). In the best case, process parameters maintain the flexibility to be adjusted based on the outcomes of the real-time testing. The control strategy is tied closely with the risk assessment, in that CQA's which are determined to have a higher risk with respect to patient safety will warrant a greater degree of process control.

The control strategy should take into account raw material quality as well, with a focus on point of entry to the process. Materials which are used in early process stages, and can be subsequently removed by purification, generally impose a lower degree of risk to the final product quality compared to those introduced in the final process steps.

Process Capability and Continual Improvement

Manufacturing processes include a series of steps which lead to the drug product. Each step includes a degree of variability, the impact of which propagates to variability in the final product. The capable process is one in which variability is both understood and controlled statistically in relation to the product acceptance criteria. The goal of QbD is to utilize experience gained during development of the process to identify potential sources of variability. Once identified, mitigation strategies can be employed to reduce overall variability both to the particular process steps and to the ultimate quality of the product. Process capability indices are used to quantitate process performance, allowing for an evaluation of process improvement across the product development lifecycle.

Continual improvement is the ultimate goal of QbD. Any process will have some degree of inherent variability. Experience gained across product development stages allow for the implementation of control that will reduce variability, resulting in a more favorable risk/benefit ratio for patients.

Prior Knowledge

Prior knowledge is a broad concept that refers to the experience gained by individuals over the course of their careers and gained by companies across the development of their portfolios. While prior knowledge cannot in and of itself be used as justification for a particular set of process controls, it should be utilized for the establishment of best practices. It should also be used to guide process development and is key to the risk assessment process. Prior knowledge is generally not information that can be easily gathered through literature sources or other public information. Instead, it is based on direct experience with like products and processes, or through studies performed to assess the current product.

Application of QbD to Biosimilar Product Development

A key premise of QbD is to start with the end in mind. In the development of a novel product, the definition of the product is complex, and requires input CMC, toxicology, clinical, regulatory, marketing. The process is iterative, as knowledge is gained on the product itself in terms of dosing requirements as well as the biophysical chemistry of the final product. With novel products, the product definition document is a living document that may face significant changes from one phase to the next.

In the development of a biosimilar, on the other hand, many aspects of the product have already been defined by the reference listed drug (RLD). The key to establishing biosimilarity is to develop a product that is the same in dosage strength, route of administration, and pharmacokinetic/pharmacodynamic profiles.

Companies developing biosimilar products will often work to mimic not only the chemistry and biological properties of the molecule, but also the delivery device (i.e. pre-filled syringe, vial, cartridge and autoinjector, etc) in order to compete with the RLD in the market. In the following paragraphs, a description is provided for how a QTPP can be arrived at for a proposed biosimilar candidate product under development. A chapter in the next section "Principles of Analytical Similarity" describes this approach in greater detail.

The Biosimilar Quality Target Product Profile (QTPP)

As is true in the development of a novel therapeutic, the first step in the QbD approach to biosimilar development is the determination of the QTPP. Although there are many elements that are defined by the reference products, there are certain areas that may require modification. As an example, for certain biosimilar targets, process patents and formulation patents may not expire at the same time. It may therefore be necessary to modify the formulation or other aspects of the product in order to avoid patent infringement. There may also be aspects of the RLD that are less than desirable for the biosimilar developer. For example, if the RLD has limited stability, or particular impurity species that are associated with adverse events for the RLD, the biosimilar manufacturer may design the process and/or formulation to mitigate those issues. At early development stages, the QTPP may be reasonably simplistic with respect to evaluation of formulation and speculative with respect to stability. The QTPP will likely be revisited at subsequent clinical stages as experience is gained with the product and its critical quality attributes.

In order to better illustrate what a QTPP should look like, an example is provided below. This example is based loosely on an example provided by the FDA for the development of modified release dosage forms (Generic Drugs 2011). The parameters described are similar to those reported for a model monoclonal antibody product. The QTPP should include not just the targets for each parameter, but also a justification for those choices. The justification will help the developer to think through the criticality of each attribute, as well as their feasibility. In the case of a biosimilar, in many cases, the justification is simply going to be that the parameters are intended to match those of the RLD. It should be noted, however, that the matching of critical quality attributes to the RLD must be assessed based on the assays developed and implemented by the biosimilar manufacturer. The exact methods used to test the RLD by the originator will most likely not be available for this comparison (Table 9.1).

Table 9.1 Example quality target profile for a biosimilar monoclonal antibody

QTTP element	Target	Justification
Dosage form	Lyophilized powder in a multidose vial for reconstitution Liquid solution in a single-dose pre-filled syringe Liquid solution in a single-dose pre-filled cartridge	Match to RLD[a]
Route of administration	Subcutaneous injection	Match to RLD
Dosage strength	50 mg/mL API, 0.5 mL/dose	Match to RLD
Formulation	*Liquid* 5% xylitol 15 mM sodium phosphate 2.5% meglumine 2.5 mM NaCl 2.5 wt% sucrose *Lyophilized* 40 mg mannitol 10 mg sucrose 1.2 mg tromethamine	*Liquid* Equivalent or improved purity and stability compared to RLD[b] *Lyophilized* Match to RLD
Pharmacokinetics	*Single 25 mg dose* Mean ± standard deviation half-life of 102 ± 30 h Clearance of 160 ± 80 mL/h. A maximum serum concentration (Cmax) of 1.1 ± 0.6 mcg/mL Time to Cmax of 69 ± 34 h was observed in these patients following a single 25 mg dose *Dosed twice weekly for 6 months* Mean Cmax was 2.4 ± 1.0 mcg/mL (N = 23) Two- to sevenfold increase in peak serum concentrations and approximately fourfold increase in AUC 0–72 h (range 1- to 17-fold) with repeated dosing	Bioequivalence requirement

Long term stability	At least 24-month shelf life under refrigeration (2–8 °C)	Equivalent to or better than RLD shelf life
Drug product quality attributes	Physical attributes	Match to RLD
	Identification	
	Assay	
	Potency	
	Purity	
	Product related impurities	
	Process related impurities	
	Moisture content (Lyo only)	
	Microbial attributes	
	Microbial limits	
Container/closure system	Sufficient container/closure integrity to prevent moisture ingress/egress and to maintain sterility of the product	Stability, patient safety
Administration/concurrence with labelling	Single-use prefilled syringes and cartridges simplify administration	Match to RLD, patient compliance. Additional information is provided in the RLD labelling
Alternate routes of administration	None	None are listed in the RLD labelling

[a]Although the physical state of the dosage form is intended to match the RLD, modifications to formulation excipients or concentrations of excipients may be acceptable as long as analytical biosimilarity can be demonstrated

[b]Excipients listed for formulation are based on those reported in Patent number US939330582. The values listed in this table were selected as the midpoints of the ranges described within the patent, and are not intended to predict the ideal formulation for the product described within the patent

Critical Quality Attribute (CQA) Development for a Biosimilar Product

As discussed earlier, the next step in the QBD process is the definition of the product CQAs. While in the development of a QTPP one can lean heavily on published data for the RLD, this is not necessarily the case for the development of our product specific CQAs. In some cases, such as in the case of Filgrastim, an entire book chapter was devoted to the description of product characterization (Herman et al. 1996, 2002). Additionally, monographs exist for Filgrastim within both the United States Pharmacopeia (USP) and the European Pharmacopeia (EP) (U.S. Pharmacopoeia 2016; EDQM 2017). In other cases, it may be difficult to even obtain a certificate of analysis for a product, let alone a full description of analytical methods.

It is important to note that regardless of what has been published on a product, critical quality attributes must still be established for a biosimilar product. There are multiple reasons for this. First, adherence to previously reported data may be misleading. Again, using Filgrastim as an example, the size exclusion chromatography method reported by Herman et al. 2002 matches the method contained in the USP monograph. The assay requires the preparation of a mobile phase containing phosphoric acid adjusted to pH 2.5. It is possible that under these conditions, soluble aggregate species may be less likely to form or may even dissociate due to charge repulsion between molecules. Additionally, the size exclusion methods that are presented in the USP and EP for this molecule are vastly different (different pH, different buffering species, etc). Furthermore, the biosimilar company cannot simply take the assumption that the analytical methods used in the testing of a product in 2002 would still be applied today. As an example, chromatography columns/resins that were used in the development of the RLD may no longer be available at the time the biosimilar molecule is pursued, leading to the expensive choice of a custom resin/custom column, or the development of a new assay. It is also important to think critically about whether the methods described within the publication or monograph are the most appropriate for the evaluation of product quality, or if there are other techniques or methodologies that should be applied instead.

If the biosimilar company intends to sell the product in a single region only, it may be acceptable to align to the monograph appropriate to that region. This alignment must still be justified based on current scientific best practices. This includes the use of orthogonal techniques for the demonstration that the method is performing adequately.

If the intent is to distribute to both the US and the EP, differences in monographs may leave the biosimilar company with the conundrum of determining which method or methods to use for testing of their molecule. There may a case for applying both assays for release and stability testing to demonstrate alignment in both regions. However, this requires additional resources to be applied for release and stability testing across the product development lifecycle.

Alternatively, the company has the option to select one method or the other, or even a completely different method or analytical technology for evaluation of the specific attribute. The onus is then to demonstrate the equivalence and/or superiority

of the chosen assay through side by side testing of both relatively pure materials and those containing reasonable levels of degraded species.

It should also be clearly stated that a pharmacopeial monograph represents the minimum testing that should be applied for the release and characterization of a biosimilar. Additional tests are required in order to ensure product quality and safety to patients at the point of DS or drug product release. Furthermore, pharmacopeial monographs do not prescribe in-process testing that should be performed to ensure the removal of specific impurity species for DS intermediates.

The establishment of CQAs should be performed at very early stages of pharmaceutical development. This activity is even more critical in the development of biosimilar molecules, as regulatory bodies tend to have higher expectations for the level of analytical characterization applied to these products even at Phase 1. For most of the products that are currently targets for biosimilar development, certain CQAs can be assumed even at the earliest program stages. As described in ICH Q6B (ICH 1999), an evaluation of product attributes including physicochemical properties, assay, purity, potency, and safety must be evaluated. Table 9.2 provides a list of parameters and methods that are often applied in the development of therapeutic proteins and monoclonal antibodies.

Though they are closely linked, the release panel should not be confused with the list of DS or DP critical quality attributes. In most cases, product manufacturing requires performance assessment across product steps, as well as the determination of parameters that are required for forward processing. In some cases, in-process methods are identical to those used for product release and are intended to monitor the removal of specific impurity species across purification steps. In other cases, the method may be unique to the process due to matrix interference in the release method. For example, for early process streams (immediately post fermentation), it may be necessary to implement an HPLC method with relatively poor resolution in place of a UV or more traditional RP-HPLC purity/impurity assay due to matrix interference or other practical considerations (such as the rapid degradation of columns that occurs when running lower purity samples).

Risk Assessments

As stated previously, the criticality of a product attribute is determined based on its propensity to cause harm to the patient should the product fall outside the defined range for that attribute. DS and DP release panels are designed to address major concerns with respect to product safety and quality. However, even within the release panel, the risks associated with each parameter will vary. In order to account for those differences, it is recommended that sponsors apply a tier-based structure for the assessment of biosimilarity (Burdick et al. 2017; U.S. Food and Drug Administration CDER 2017). In order to completely define the criticality of any particular product attribute, a risk assessment must be performed.

A tier-based strategy divides the assays on the release and characterization testing panel into one of three categories, based on the criticality of data produced in the

Table 9.2 Example release panel for therapeutic protein and monoclonal antibody drug substances

Category	Attribute	Analytical Method
Assay	Concentration	USP <1057>
Potency	Relative potency[a]	Cell based assay[b]
Physicochemical properties	Appearance	Visual inspection
	pH	USP <797>
	Osmolality[c]	USP <785>
Product related purity/impurity	Soluble aggregates	SEC-HPLC
	Charge variants	IEX-HPLC
	Oxidized and other product related impurity species	RP-HPLC[d]
	Covalent vs. non-covalent aggregates	SDS-PAGE, reduced and non-reduced Or CE-SDS, reduced and non-reduced
Process related purity/impurity	Residual DNA	qPCR
	Residual host cell proteins	ELISA[e]
	Residual protein A[f]	ELISA
Microbiological attributes	Bioburden	USP <61>
	Endotoxin	USP <85>

[a]Determined in comparison to a product specific reference standard
[b]Relatively simplistic assays may be acceptable for the demonstration of potency at early stage. By the end of phase 3, the potency assay must be representative of the mechanism of action for the product
[c]May not be included at the drug substance stage if fill/finish includes additional formulation of the product
[d]May not be required for certain mAb products for which adequate separation of oxidized species is achieved by IEX-HPLC
[e]Kit based assays are acceptable at early clinical stage. By Phase 3, a process specific assay must be implemented
[f]Specific to monoclonal antibodies. Not relevant for therapeutic proteins

assay. The assays which are most critical to the confirmation of biosimilarity fall into tier 1, and usually consist of the product potency assays. The assessment of biosimilarity for a tier 1 assay requires a statistical determination of equivalence between the RLD and the biosimilar product.

Tier 2 assays include most of the purity/impurity assays, as well as the determination of strength, microbial purity, and other parameters that are numerically driven. Tier 2 assays also require a statistical evaluation in comparison to the RLD, but in this case it's a comparison to the range of experience determined for the RLD as opposed to a more stringent statistical determination of equivalence.

Tier 3 assays include all assays for which numerical data is not achievable/relevant. For example, assessment of appearance, which may just be comments around a visual inspection, or spectroscopic methods for the evaluation of protein secondary or tertiary structure. There is no statistical evaluation of results required for tier 3 assays.

Risk assessments help to tie the parameter in question to the risk to patient safety. The impact to patient safety if a failure should occur, in conjunction with the ability to detect the impurity or product quality failure if it were to occur and can also assist in assigning quality attributes to the tiers as described above.

Risk levels may vary significantly from impurity to impurity within a single product. Product use and administration may also impact patient risk for a given attribute. For example, with respect to aggregated proteins, it is reported in the literature that adverse immunological responses to an analyte generally don't occur upon the first exposure. However, the first exposure opens the door for the immune system to generate anti-drug antibodies (ADAs), which can reduce the safety and/or effectiveness of the treatment during subsequent administration. Therefore, products that are administered a single time within a patient's life will result in a lower risk rating with respect to the presence of aggregates compared to those that are administered repeatedly.

Risk assessments require contribution from all members of the product development team, including analytical chemists, process development scientists, toxicologists, clinical specialists, quality assurance specialists and regulatory experts. The exact methodology used for a risk assessment will vary from company to company. Regardless of the methodology utilized, the process should be proceduralized to ensure consistency across the company, as well as to provide the agency with the framework upon which risk has been evaluated. Risk assessment outcomes can easily be influenced by the risk tolerance of the participants. Those who have worked in large pharma with well-defined practices and procedural expectations may be more risk averse compared to individuals who have built their careers at startup companies. This is not to say that persons at start-up companies have any less concern for their patients compared to a larger company. It's more that they are required on a daily basis to define and understand the difference between a patient risk, a business, and a regulatory risk, and to take those differences into account during the risk assessment process.

Although the purpose of a risk assessment is relatively easy to understand, the methodologies used in the process may be more difficult to grasp. How do you set a rank score for a risk? The answer is that you have to take the assumption that overall risk is not based on a single parameter but is instead reflected as the propagation of multiple inputs. In the case of a biosimilar, those inputs may include risk to patient should the defect occur, and the ability to detect the defect if it were present in a sample. Because the risks are propagated through one another, quantitation of risk levels is not additive but instead multiplicative. Even with a quantitation based on the multiplication of risk input values, there is a danger of misinterpretation of the grey areas (i.e. lack of separation between cumulative risk that is low, moderate, or high). In order to allow for easier assessment of that cumulative risk, it is recommended that during the assignment of risk values to a particular input, the numerical assignments for risk levels should vary by almost an order of magnitude rather than trying to assess risk on a scale of 1–5.

In order to illustrate the risk assessment process, an example for a model monoclonal antibody product is provided in Table 9.3. When performing a risk

Table 9.3 Risk assessment example for a model monoclonal antibody

Attribute	Impact	Reason	Uncertainty	Reason	Final score	CQA	Testing	Control strategy
Appearance	16	Difference in color or clarity could indicate new impurity	7	Materials produced to date match RLD	112	Yes	Release test	Control of process and raw materials
pH	16	pH outside target range could significantly impact degradation rates	7	Not tested between in process and DS release	112	Yes	Release test	pH of final formulation buffer
Charge variants	16	Different glycosylation patterns could impact PK/PD	7	Exact fermentation conditions for RLD not known	112	Yes	Release test	Selection of clone that produces glycosylation pattern with closest match to RLD
Amino acid sequence	16	Amino acid substitutions could impact activity/stability	3	RLD amino acid sequence known	48	Yes	Release test for identity	Controlled through validation of master cell bank
Secondary structure	12	Changes could impact stability or potency	1	Major change would impact binding	12	No		
Tertiary structure	12	Changes could impact stability or potency	1	Major change would impact binding	12	No		
Bioburden/ sterility	20	Immediate risk to patient safety	7	Specific contaminating organism may be unknown	140	Yes	Release test	Filtration of intermediates, cleaning, in-process testing, filtration prior to fill/finish
Endotoxin	16	Immediate risk to patient safety	7	Known contaminant, increases with storage if bioburden present	112	Yes	Release test	Control of incoming RM's, control of bioburden

Oxidation	16	Oxidation within binding site impacts potency	7	Oxidation rates and formulation impact unknown	112	Yes	Release test	Formulation development/excipient range study
Aggregates	16	Potential for immune response to IgG	7	Acceptable level of aggregates not known	112	Yes	Release test	Formulation development/excipient range study
Fragments	4	Depending on size of fragment, could be active	7	Impact on immunogenicity unknown	28	No	Release tests—fragment would be observed by CE-SDS R/NR or SEC	
Osmolality	2	Dilution of product in isotonic saline prior to administration	1	Based on excipient addition	2	No		
Host cell proteins	12	Potential immune response	5	Process specific assay for RLD not available	60	Yes	DS release test	Track removal in process, minimize levels in final product
Host DNA	12	Potential immune response	5	Process specific assay for RLD not available	60	Yes	DS release test	Track removal in process, minimize levels in final product during process qualification/validation
Concentration	12	Potential for over or under dosing	5	Exact impact to PK/PD not yet determined, potency assay highly variable	60	Yes	Release test	Control of final concentration through two-step final dilution
Potency	20	Impact on clinical response	7	Highly variable potency assay	140	Yes	Release test	Potency assay improvements, increased number of replicate analyses

assessment, it is valuable to document not just the applied risk scores, but also the logic that was used to determine those scores, as that will serve as a backbone for later risk assessments. As knowledge is gained about the product and process, risk levels will naturally change. The logic used for previous risk assessments in conjunction with the newly gained knowledge, can be used to further justify the change in risk level.

It should be noted that there is no defined “right way” to perform a risk assessment. It is more important to have a process that you can understand and defend than to follow a specific process. At the same time, the process that is adopted for use in your facility should be adequately documented and applied consistently across programs and functional areas.

As was stated previously, risk assessments should not be performed at a single point within a development program. Instead, risk should be evaluated throughout the development process. Triggers for performing a risk assessment would include development milestones (i.e. transition from one clinical phase to another), or major changes to the product (i.e. changes in manufacturing scale or location, application of new analytical methods, changes to DS or DP formulation, etc).

While it may be tempting to limit your risk assessments to your assumed critical quality attributes, extended characterization may be necessary to fully justify the evaluation of risk levels. This is especially true for a biosimilar product, where a higher level of analytical characterization is expected at early program stages.

One of the most difficult aspects of analytical method and formulation development at early development stages for a novel product is having adequate quantities of material to perform the studies. Preliminary fermentation and purification runs may result in the production of milligram quantities of product. Depending on the methodologies used for analytical and formulation development, those activities can easily consume hundreds of milligrams of material. Also, the material that is produced at that stage is unlikely to be representative of materials that will actually go in to the clinic. So, you may find yourself chasing impurities and trying to mitigate degradation pathways that are exaggerated compared to what will be seen in later stages. In the best case, this results in methods and formulations that are more robust than they would have been otherwise. In the worst case, this results in loss of time and a drain of resources with little useable data, as those studies may need to be re-evaluated once processes have been optimized.

In biosimilar product development, there is a distinct advantage with respect to risk assessment activities. Specifically, the reference product, which is representative of the final QTPP for the program, is available before process development even begins, and there are a number of reputable sources for acquiring such material. As a result, analytical development activities, including forced degradation studies to help define the most likely product degradation pathways, can and should be initiated even before the first materials are produced in the biosimilar development laboratories.

For the formal establishment of biosimilarity, the agency requires that multiple lots of the RLD are evaluated and statistics applied for the comparisons between

RLD and biosimilar product. While this is not an absolute requirement during the establishment of the QTPP, the value of evaluating multiple RLD lots at even the earliest development stages is that you can begin to assess the variability across certain parameters within the RLD. Understanding that variability may help you in the risk assessment process. For example, if you see significant variability in charge variant profiles across different lots of RLD, you may be able to conclude that variations of this magnitude have minimal impact on product performance in patients.

Design Space

The design space is the relationship between the process inputs and the critical quality attributes. Risk assessments are closely tied to the development of the design space, as they provide a preliminary view of which parameters are likely to have the greatest impact on product quality. This can serve to optimize process characterization studies that are required to add definition to that design space. By focusing first on process variables that are most likely to impact patient safety, the greatest risks can be more efficiently mitigated through the implementation of in-process controls and critical operating ranges even at phase 1.

The ultimate goal of the design space is to create a robust process. A robust process is defined as one that is not sensitive to "small" changes. The objective of process development and characterization is to determine acceptable ranges to ensure the process consistently produces the same quality of material. The key inputs in the development of a robust product include information about the final intended dosage form, the quality attributes, and the general manufacturing pathway. The previous sections of this chapter have provided examples for how the intended dosage form and quality attributes are determined for a biosimilar product. Similar to the establishment of critical quality attributes, the biosimilar developer has an advantage in the design of the manufacturing process compared to innovator companies. Specifically, analytical testing that is performed in the establishment of CQAs can also provide information about the process steps and purification techniques that will need to be applied for removal of particular impurity species or avoidance of a particular degradation pathway. Optimization of a fermentation process is the key to establishing glycosylation patterns for a mammalian cell culture product and gives the sponsor the greatest chance of matching properties for the RLD.

The application of Design of Experiments can be utilized to further optimize process characterization studies and the establishment of the design space. In the DOE approach, multiple parameters are evaluated, and interactions between the parameters may be determined as well. The design itself, as well as the evaluation of resulting data, is based on statistics.

In the development of a novel molecule, the manufacturing process will go through multiple changes and iterations as additional experience is gained on the molecule, and critical quality attributes will be better defined as the program progresses. The biosimilar program, conversely, requires a very early definition of CQAs and establishment of acceptable ranges around those attributes in comparison

to the RLD. As such, there is increased pressure on manufacturing to produce a particular quality of product at early stages. DOE can significantly reduce the time required to develop a capable process compared to the more traditional One Factor At a Time (OFAT) studies.

An example for a DOE study design is provided below. In this example, the goal is the optimization of glycosylation profiles based on fermentation parameters. The specific parameters under investigation include temperature, pH, and glucose concentration. The output is based on a specific desired glycosylation pattern which is intended to match the RLD. In this example, a statistical analysis software, JMP (SAS) is utilized. The first illustration, shown in Fig. 9.1, describes the type of design (full factorial vs. fractional factorial), the output Response parameters, the input parameters and their ranges, the number of center point runs, and the number of replicates. The second illustration, provided in Table 9.4, shows the resulting experimental design table.

The experimental process runs are performed as described in Table 9.4, then resulting output results added to the evaluation. There are multiple options for how to view the output of the data including cube plots, response surface diagrams, and interaction plots. In this case, the output parameters are assessed using interaction plots, as shown in Fig. 9.2. With this example, there appears to be an additive effect of increasing both temperature and glucose concentration with respect to the number of charge variant peaks observed.

Once the interactions have been defined, and influence of parameters on one another have been evaluated, it is possible to select more appropriate process control ranges in order to ensure consistency in manufactured product. The determination

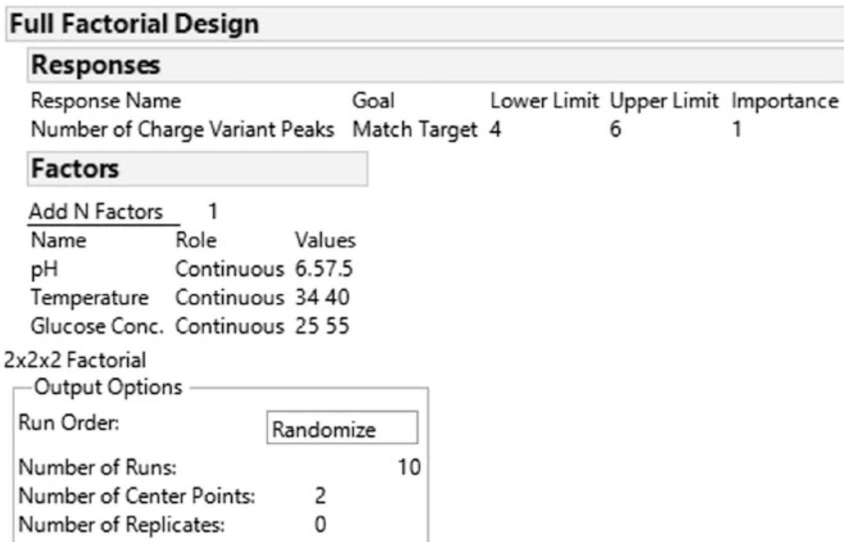


Fig. 9.1 JMP DOE study parameters

Table 9.4 Statistically derived experimental design

Pattern	pH	Temperature	Glucose conc.
− + +	6.5	40	55
− − +	6.5	34	55
− − +	6.5	34	55
+ − +	7.5	34	55
+ + +	7.5	40	55
+ − −	7.5	34	25
− + −	6.5	40	25
− − −	6.5	34	25
+ + −	7.5	40	25
0 0 0	7	37	40
0 0 0	7	37	40

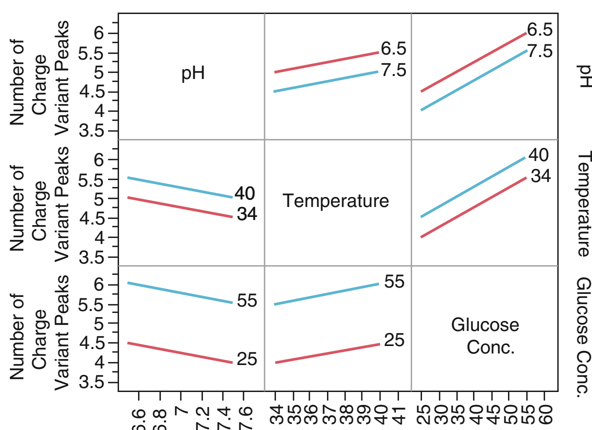


Fig. 9.2 Interaction plots based on DOE results

of design space is not limited to the development of the manufacturing process but should also be applied in the development of product formulation and analytical methods for evaluation of CQAs.

Control Strategy

The control strategy for a biosimilar is much like that applied to a novel product. It requires an evaluation of the process to determine which quality attributes are most critical to product performance and the robustness of each step of the manufacturing process. The integrated process strategy for a biosimilar will employ controls around procedures, raw materials, in-process control (IPC) testing, process monitoring, release and stability testing, process validation, product characterization, and evaluation of product comparability following any major changes to product formulation or manufacturing.

One major differences between a control strategy for a biosimilar vs. a novel product is related to the development timeline. While for a novel product the control strategy may evolve over multiple years and three clinical phases, the manufacturing process and control strategy for a biosimilar product may be locked as early as first phase 1 in support of the first in human studies. Even in the event that process development continues past phase 1, phase 2 is generally skipped and the process locked early in phase 3. The limits to the timeline mean less time to determine critical process parameters as well as fewer manufactured lots upon which to gain product/process experience. This puts a much higher burden on the QbD process for establishing the capable process very early in the program.

The key to a successful biosimilar program is the establishment of analytical comparability. While that has an obvious implication on the analytical development teams, it additionally impacts the process development teams who must develop a process that meets much tighter analytical acceptance criteria for product release at an early program stage compared to a novel product. Tighter analytical acceptance criteria by necessity leads to more restrictive process controls.

A further challenge in the development of a control strategy for a biosimilar is that it must have the flexibility to evolve based on performance of the originator product over time. Although many of the RLD products that are pursued as biosimilar candidates have been on the market for a number of years, there may still be a need for manufacturing changes at the originator company over time due capacity constraints, changes in availability of a critical raw material or starting material, or just the need for modernization of certain aspects of the manufacturing process. The comparability acceptance criteria that are applied to the RLD may be broader than those applied in a biosimilarity exercise. It is therefore possible to design your process to meet a particular RLD product profile, only to find that between the time you performed your biosimilarity assessment in support of first in human studies and the time that you perform the follow-up assessment in support of your 351(k) biosimilar filing, the goal posts around critical quality attributes for the RLD may have shifted. That being the case, it is critical during the development of a biosimilar program that you have an adequate procurement strategy that includes purchasing RLD not just once but over time.

Process Capability and Continuous Improvement

The capable process is one that incorporates the concepts discussed above to create a robust process that will consistently deliver product of expected quality, safety and performance. For a novel product, the understanding of process variability gained in the definition of the design space allows for an objective mechanism for evaluating normal variability and its potential to impact overall product quality. The application of statistics to data collected during process and analytical development and validation studies allow for the setting of appropriate specifications. In-process control testing, release and stability testing results that fall within predefined acceptance criteria provide evidence of success of the process as well as the quality of the product.

When developing a biosimilar, the capability of a process is not judged solely on its own merits, but also on the performance of the product in comparison to the RLD. While the same basic equations and calculations are performed to evaluate process capability for either a biosimilar or a novel product, it is important to remember that the properties of the RLD are taken into account as part of the statistical data analysis used to justify product specifications, leading to more stringent requirements for capability in the manufacture of a biosimilar.

The idea of continuous improvement almost seems to contradict the development strategy and path for most biosimilars in that the goal is to lock the commercial manufacturing process as early in development as possible. Process and analytical validation studies are initiated after a much shorter development period than would be applied for a novel product. It is only logical, therefore, to conclude that continuous improvement would be not only applicable to biosimilar, but critical to assuring continued product quality and patient safety as additional product experience is gained post approval. It is important to recognize that a locked commercial process is not a process that can never be changed. It is instead one that requires a higher level of justification for making changes, and where (especially post licensure) agency alignment on both the intended changes and the planned comparability assessments should be obtained prior to the initiation of any change.

Prior Knowledge

Prior knowledge, in relation to biosimilar development, is a complicated topic. In this era, it is common for people to work at four to five, or even as many as 10 different companies within their career. It is therefore very common for a person who once worked on the development of an RLD to find themselves years later at a company that is working on the development of a biosimilar to that RLD.

As an individual, one must carefully separate prior knowledge of product development in general from prior knowledge of the specific RLD. Information gained through the public sector (i.e. publications, press releases, etc) may be immediately applicable to the development of a biosimilar product. General knowledge of process or formulation development concepts may be applied to the development of a biosimilar and used to justify decisions during development. Specific knowledge of why a particular development pathway was chosen for the RLD cannot be used to justify a set of process parameters, specifications, or other acceptance criteria. Specific knowledge of product degradation or sensitivity based on experience with the RLD is not sufficient evidence for a risk assessment, even if it has also been reported in the literature.

Although the distinction between general and specific prior knowledge is important to keep ourselves out of the courtroom and to provide justification to agencies, it is also important for our own decision making processes. We must remember that the reason we refer to these products as biosimilars instead of biogenerics is that it is not possible to create an exact copy of a biological molecule. There will always be some differences related to the exact fermentation or purification process, or even

the host cell that is used for development of the biosimilar. Without a thorough understanding of those differences, however, minor, one cannot assume that those differences will have no impact on product quality or on processes that are necessary for successful biosimilar product development.

QbD and Analytical Development

In the discussion on the setting of critical quality attributes, references were made to different resources that may be available for use in analytical method selection. Alignment to a historical resource or even a compendial monograph does not guarantee successful method execution. QbD principles should be applied to the development of new analytical methods for the same reasons that they should be applied to process development; reduction in development time and method understanding compared to OFAT studies.

If you ask a room full of analytical development scientists whether they are currently applying QbD principles to their analytical method development, a few will nod and smile. The rest will just look puzzled and potentially avoid eye contact. However, once a dialogue is initiated on the topic, we see that most of us are already incorporating at least some QbD principles in our method development activities. We may just not think of it in those formal terms. The application of the formalized QbD principles to the development of analytical methods is described below.

The Quality Target Analytical Method Profile (QTAMP)

The first element of analytical QbD is the definition of the method. The QTAMP, like the QTPP, is a prospective summary of the desired characteristics of the analytical method. The QTAMP includes definition of all assay parameters and measures of success. For example, the goal may be to develop an RP-HPLC for evaluation of product oxidation. The QTAMP should include also a definition of desired performance characteristics. For example, desired resolution of impurity peaks from the main peak, or acceptable levels of method variability.

Critical Method Attributes (CMAs)

CMA's are not specifically defined in any ICH or FDA guidance document. However, they should be thought of as any method parameter that, if altered, has a high probability of impacting the ability to meet system suitability criteria and/or the correct reported value determined in the method.

CMAs are typically determined through the performance of assay robustness studies. The QbD mindset strongly promotes performance of assay robustness studies early in the analytical development lifecycle so that knowledge can be applied for the sake of method improvements, analytical investigations, and the setting of operating ranges that will ensure adequate method performance. CMAs may also include critical assay reagents or environmental conditions (i.e. laboratory temperature or humidity).

Risk Assessments

Risk assessments are performed in analytical QbD to predict the most likely failure points for an assay and to determine how the failure will be determined. Risk assessments are key to establishment of system suitability criteria, and to the implementation of adequate control strategies for critical assay reagents.

Design Space

The design space for an assay is also defined primarily through assay robustness studies. The goal in setting the design space for an analytical method is to confirm that all assay ranges (i.e. column temperature, incubation times, mobile phase makeup, etc.) are capable of providing consistent method performance. The method performance is evaluated based on properties specific to the method. For example, in an HPLC method, method performance may be evaluated based on retention time for main peak (and other major peaks) in comparison with the reference standard, adequate separation between main peak and impurity peaks, or the ability to accurately detect impurities in the presence of potential interfering species.

As was described for process development, DOE or other statistically designed studies may be performed for method optimization and definition of assay design space. The primary goal is to employ adequate assay controls to allow for consistent performance of the analytical method. A thorough DOE study also provides evidence that can be used in laboratory investigations down the road for the determination of root cause in the event of assay failure.

Control Strategy

Once the analytical method has been adequately designed and defined, critical quality attributes determined, and the design space established, controls must be implemented to assure future analytical method performance. It is recognized that analytical methods fail for a number reasons, including age of columns, inconsistent

quality of critical reagents, as well as simple human error. Of those failure mechanisms, human error is the most difficult to control as it is generally random. Procedural controls may be implemented to reduce the level of perfection required by the analysis. For example, proportions of aqueous and organic components of mobile phases may be designed to decrease the impact of gradient slope on analyte separation. For the determination of protein concentration by A280, gravimetric dilutions may be employed to improve accuracy and precision especially where significant product dilution is required.

Analytical Method Capability and Continual Improvement

Analytical methods are constantly under evaluation, through both formal activities such as routine release/stability testing of product, training of new analysts, or transfer of methods to outsource vendors. Analytical method capability may be more easily defined based on the number of successful runs vs. analytical investigations for a given method.

Continual improvement is a goal of QbD analytical method development. The consequence of continual improvement is often the need for method revalidation following method improvement. Utilization of QbD principles in method development at an early development stage, in conjunction with applying adequate resources to analytical activities, can reduce the need for assay changes down the road. The application of design space may negate the need for revalidation in some cases, or reduce regulatory requirements for post approval change notification in others.

Prior Knowledge

Prior knowledge is applied to analytical method development frequently. Corporate prior knowledge may result in the implementation of platform methods for early stage studies, which will be replaced or supplemented with product specific methods prior to licensure.

QbD and Formulation Development

Similar to analytical development, many are already utilizing some QbD concepts in the development of product formulations. To some degree, the first QbD step has already been completed with the development of the QTPP. The product definition and determination of the QTPP provide guidance on the intended formulation. In the development of biosimilars there may be even further definition around the QTPP for the drug product since it has been pre-defined by the innovator. However, one should not assume necessarily that formulation development activities are not

required. Even if a decision is made to match the formulation of the RLD, acceptable ranges for each excipient must be determined.

As was true for analytical development, we often have an advantage in the development of a formulation for our biosimilar product in that the RLD is already available to us through reputable sources. It is therefore possible to purchase sufficient quantities of that originator product for initiation of formulation development. However, depending on the product, that approach may be cost prohibitive.

The example provided in Table 9.1 represents an interesting case where there are two different formulations that have been defined for the biosimilar product; one which matches the formulation for the RLD and one which does not. As was previously mentioned, there may be multiple drivers for not aligning to the RLD formulation. In the case where the decision is related to a patent on the RLD formulation, the development scientists are aware at even the earliest program stage that a new formulation will be required. Conversely, the need to change formulation in order to facilitate improved stability may take more time and data to justify.

Critical Formulation Attributes (CFAs)

CFAs are simply the excipients or formulation parameters that have the greatest impact on product performance and stability. One example of a CFA is solution pH. It is widely reported in the literature that oxidation and deamidation occur at faster rates at pH extremes and are minimized at neutral pH. Aggregation may be worst at or near the product isoelectric point.

The choice of excipients may have an impact beyond the stability of a product in solution. Phosphate buffered saline was a very common excipient in some early biotechnology product formulations, due to its similarity with physiologic conditions. However, this excipient combination can lead to significant damage of product upon freeze thaw due to crystallization of the monobasic and dibasic forms of phosphate at different temperatures, leading to sometimes extreme pH shifts during freeze/thaw cycling. Sodium chloride has a very low glass transition temperature, making it very difficult to work with during lyophilization. Polysorbates may be necessary for product stability and drug in vials may perform very well in the presence of 0.005% polysorbate. However, if the final product container/closure system is intended to be a pre-filled syringe, polysorbate levels may need to be increased to maintain product stability in the presence of the silicone oil/water interface that is created due to coating of rubber syringe plunger tips for ease of injection. At the same time, adding too much polysorbate to product in a pre-filled syringe can lead to additional extraction of silicone oil from the syringe surfaces, leading to opalescence of the product.

CFAs are best defined through the use of DOE or other statistically designed studies. This allows for the determination of not just the impact of a single parameter, but also the determination of interactions between excipients/parameters and the overall impact to product quality.

Risk Assessments

Risk assessments are performed in formulation development for multiple reasons. The first may be to justify changes to product container/closure, such as the above example of switching from a vial to a pre-filled syringe. Risk assessments may also be applied for the justification of formulation changes/improvements across the product development lifecycle, or in comparison to the RLD.

Design Space

The design space for a formulation is also defined through formulation development studies. Matrix design studies can be used for the investigation of formulation. However, DOE studies offer significant advantage as you are often working to evaluate multiple parameters (pH, ionic strength, stabilizers, etc) that may influence one another. The determination of the formulation design space allows for application of appropriate levels of process control depending on the criticality of a given excipient. For example, the exact concentration of sucrose may not have a significant impact on product stability, while minor variations in pH may have a major impact on degradation rates.

Control Strategy

Much of the formulation control strategy related to product formulation is centered around manufacturing controls and acceptable excipient ranges. However, another major component to the control strategy for product formulation is control of raw materials. Formulation is often the final step in the manufacturing step. As such, the impurities that enter the product during formulation have no opportunity for removal during purification. Therefore, selection of well characterized raw materials (e.g., compendial excipients, sugars with low endotoxin levels, polysorbates with low peroxide levels, etc.) may be critical to product quality.

Formulation Capability and Continual Improvement

The capable formulation is one that has a well-defined design space, is robust, and will maximize product stability. It may be determined for a biosimilar that the RLD formulation is the most stable. In this case, opportunities for continual improvement of the formulation are limited. Even in cases where a novel formulation is desired for a biosimilar, the goal is to have the final formulation locked prior to initiation of phase 3 studies. Changes to formulation are considered a major change to the product, as they do have the potential for altering drug metabolism and pharmacokinetic

(DMPK) profiles. While QbD concepts are applicable to formulation development, the timelines are often shifted to earlier stages compared to those applied for process development.

Prior Knowledge

Prior knowledge is applied to formulation development in different ways. In some cases, companies may choose a platform formulation a platform formulation study design as a starting point for formulation development. Literature studies on similar molecules will provide hints as to which excipients are most likely to stabilize a product. Knowledge of specific properties of the molecule, such as the isoelectric point, the number of surface exposed methionine residues, or the number of free cysteines, may also guide which parameters and excipients are evaluated during formulation. Knowledge of the originator formulation is of course referenced for comparative purposes.

Conclusion

This chapter discussed both the QbD framework and the application of QbD to biosimilar development. While QbD is a goal for the development of any product, its value becomes even more significant for biosimilar development due to shortened development timelines. Historical events related to safety of follow-on biologics, including those manufactured with shared knowledge between companies (e.g., the association of pure red cell aplasia in patients receiving erythropoietin manufactured by Johnson & Johnson instead of by Amgen) have demonstrated that a higher degree of patient risk may exist for biosimilars compared to generics. Utilization of QbD concepts such as described in this paper, especially the application of risk evaluations, is necessary to ensure patient safety.

While patient safety is the number one priority, companies also benefit from the application of QbD to their biosimilar product development programs. The establishment of biosimilarity is critical to the reduction of expenses and development time related to the performance of clinical studies. Studies performed for the demonstration of biosimilarity are much more convincing to agency reviewers when they incorporate the QbD principles. Additional benefits include a reduction of manufacturing development timelines and cost through the use of statistically designed studies.

The application of QbD concepts during analytical method development and formulation development is necessary due to the need to submit a higher level of product characterization and understanding at earlier stages. Ideally, analytics and product formulation for a biosimilar may be defined and extensive characterization studies performed to support an IND application in order to initiate first-in-human studies.

Tighter restrictions on the product characteristics for a biosimilar compared to the innovator also necessitate QbD. As was stated earlier, the acceptance criteria for the establishment of biosimilarity may be more restrictive than those applied to the RLD for justification of a change in manufacturing process. Through the application of risk assessments and DOE studies that define critical product and process attributes can a biosimilar be successfully and efficiently developed. The application of a robust control strategy is necessary for assurance of consistency for the biosimilar product across the product lifecycle.

References

Awotwe-Otoo D, Agarabi C, Wu GK, Casey E, Read E, Lute S, et al. Quality by design: impact of formulation variables and their interactions on quality attributes of a lyophilized monoclonal antibody. Int J Pharm. 2012;438(1–2):167–75.

Burdick RK, LeBlond DJ, Pfahler LB, Quiroz J, Sidor L, Vukovinsky K, Zhang L. Statistical applications for chemistry, manufacturing and controls (CMC) in the pharmaceutical industry. In: Gail M, Samet J, Singer B, Tsiatis A, editors. Statistics for biology and health. Cham: Springer; 2017. p. 363–8.

EDQM. European pharmacopoeia. 9th ed. Strasbourg: Council of European Pharmacopoeia Europe; 2017. p. 2471.

Generic Drugs. Quality by design for ANDAs: an example for modified release dosage forms. 2011. https://www.fda.gov/downloads/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/abbreviatednewdrugapplicationandagenerics/ucm286595.pdf.

Goldek M, Liebowitz S, McCarthy R, McNally G, Oksanen C, Schultz T, Sundararajan M, Vorkapich R, Vukovinsky K, Watts C, Millili G. Process robustness – a PQRI white paper. Pharm Eng. 2006;26(6):1–11.

Ha E, Wang W, Wang YJ. Peroxide formation in polysorbate 80 and protein stability. J Pharm Sci. 2002;91(10):2252–64.

Herman AC, Boone TC, Liu HS. Characterization, formulation, and stability of Neupogen® (Filgrastim), a recombinant human granulocyte-colony stimulating factor. In: Pearlman R, Wang J, editors. Formulation, characterization and stability of protein drugs. New York: Plenum; 1996.

Herman AC, Boone TC, Lu HS. Characterization, formulation, and stability of neupogen® (Filgrastim), a recombinant human granulocyte-colony stimulating factor. In: Pearlman R, Wang YJ, editors. Formulation, characterization, and stability of protein drugs: case histories. Pharmaceutical biotechnology, vol. 9. Boston: Springer; 2002.

ICH. ICH harmonized tripartite guideline: specifications: test procedures and acceptance criteria for biotechnological/biological products. Q6B, 1999.

ICH. ICH harmonized tripartite guideline: quality risk management Q9, 2005.

ICH. ICH harmonised tripartite guideline: pharmaceutical development Q8 (R2), 2009.

ICH. ICH harmonized tripartite guideline: development and manufacture of drug substances (chemical entities and biotechnological/biological entities), Q11, 2012.

Little AD. Scatter acorns that oaks may grow. Cambridge: MIT Libraries; 2013. Retrieved 13 November 2013.

MacGregor JF, Bruwer MJ. A framework for the development of design and control spaces. J Pharm Innov. 2008;3:15–22.

Nazzal S, Nutan M, Palamakula A, Shah R, Zaghloul AA, Khan MA. Optimization of a self-nanoemulsified tablet dosage form of ubiquinone using response surface methodology: effect of formulation ingredients. Int J Pharm. 2002;240:103–14.

Sangshetti J, Deshpande M, Zaheer Z, Shinde D, Arote R. Quality by design approach: regulatory need. Arab J Chem. 2017;10:S3412–25.

Singh SR, Zhang J, O'Dell C, Hsieh M, Goldstein J, Liu J, Srivastava A. Effect of polysorbate 80 quality of photostability of a monoclonal antibody. AAPS Pharm SciTech. 2012;13(2):422–30.

U.S. Food and Drug Administration CDER. Guidance for industry: PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance. 2004.

U.S. Food and Drug Administration CDER. Guidance for industry: statistical approaches to evaluate analytical similarity. 2017.

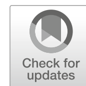
U.S. Pharmacopoeia. 2017 U.S. pharmacopoeia-national formulary [USP 40 NF 35], vol. 1. Rockville: Pharmacopeial Convention, Inc; 2016. p. 4196.

Woodcock J. Concept of pharmaceutical quality. Am Pharm J Rev. 2004;7:10–5.

Yu LX. Pharmaceutical quality by design: product and process development, understanding, and control. Pharm Res. 2008;25:781–91.

Yu LX, Amidon G, Khan M, Hoag S, Polli J, Raju GK, Woodcock J. Understanding pharmaceutical quality by design. AAPS J. 2014;16(4):771–83.

Chapter 10
Drug Product Considerations for Biosimilars



Satish K. Singh, Susanne Jörg, and Hanns-Christian Mahler

Abstract Biosimilars are increasingly important for the expansion of access to life-saving biologics and to reduce the cost of prescription medicine. The design of the drug product for a biosimilar is of utmost importance, as it directly relates to patient efficacy, safety and product quality, even as it defines the similarity of the product to the reference innovator product. There are various components of a drug product: the formulation, the container closure system and/or device, and the drug product manufacturing process that must be considered during any biologic parenteral drug product development. The biosimilar drug product development strategy has to take additional multiple factors into consideration including matching the reference product presentation, intellectual property and the current state of the art in science and technology. This chapter reviews these factors in detail using case-studies from currently approved biosimilars.

Keywords Product · Composition · Formulation · Dosage form · Device · Manufacturing process · Container closure

Introduction

Biologics are an important class of therapeutic drugs that have been very successful in the treatment of severe diseases such as cancer, arthritic diseases and others, and especially with targets that may not be readily addressable with small molecules. Almost eight of the 10 top selling drugs in the world are now biologics and come with high list prices. Biologics account for 28% of total medication expenditure in the US even though they represent less than 1% of all prescriptions (Scott Morton et al. 2016). One of the mechanisms to reduce total cost to the health-care system has been to enable biosimilar products.

S. K. Singh (✉) · S. Jörg · H.-C. Mahler
Lonza Pharma and Biotech, Drug Product Services, Basel, Switzerland
e-mail: Satish.singh@lonza.com; Susanne.joerg@lonza.com; Hanns-Christian.mahler@lonza.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_10

A biosimilar is an off-patent biologic medicinal product that is produced by manufacturers other than the originators/innovators, which is highly similar to the innovator ($\equiv$ reference) product and considered therapeutically equivalent in the reference indication(s). Reference to the innovator product, generally and preferably from the target jurisdiction, is key to the approval process. The foundation to approval is a set of detailed physical and biochemical characterization studies that show a high level of similarity in key characteristics between the biosimilar and innovator. Comparable non-clinical and clinical safety (including immunogenicity), and (occasionally) efficacy profiles represent the next pieces in this totality of evidence approach towards the biosimilar designation. Well-defined regulatory framework have been created for a product to be designated a biosimilar, especially in the developed markets (e.g. CHMP/437/04 2005 in the EU, and the BPCI Act 2009 in the US). These guidance or guidelines lay out the requirements to demonstrate the similar nature of two biological products, and rely significantly on the demonstration of structural and functional similarity by multiple advanced analytical techniques.

Analytical studies are the basis of assessing similarity to the reference products, and a preclinical and clinical study or studies are usually required to demonstrate overall safety, purity and potency as well as comparability in one or more indications for which the reference product is licensed.

Clinical studies are also often required, to ensure that pharmacodynamics and pharmacokinetics in humans are indeed similar to that of the reference product. These human studies also provide data related to the risk for immunogenicity, although the statistical power to detect low rates of immunogenicity in such studies may be poor. Immunogenicity is thus often a residual concern that cannot be easily addressed by in vitro or animal studies either. Pharmacovigilance is also an important consideration for biosimilars, to ensure that there is a level of monitoring in the field. Automatic substitutions of reference products with biosimilars are also not allowed, for example, in the EU, to ensure that patients do not move between product versions during a given treatment regimen and to avoid related risks.

Unlike small molecules, biologics are generally complex, both in the process that they are produced from, and in their inherent molecular structure. The structural complexity almost ensures that the pool of molecules in a batch will have a certain degree of heterogeneity. The biosimilar manufacturers do not have access to key material or information, such as cell/host construct, clone, details of the fermentation/bioreactor conditions and purification process, nor the purified (unformulated) active pharmaceutical ingredient. The quality target profile of the innovator molecule (and thus of the biosimilar product) is therefore derived from analyzing the reference molecule from the commercially available product. For this purpose, (a number of) reference drug product (batches) are procured from the market and subject to an extensive array of analytical methods. Defining the target for developing a biosimilar molecule thus relies on the reference drug product availability, generally of various ages and potentially from different versions of a process (Lamanna et al. 2018).

Apart from producing a "highly similar" molecule, biosimilar development needs to consider the design of the biosimilar drug product, as it directly relates to patient

safety and efficacy, as well as maintenance of quality and similarity to the reference innovator product. The general development objective is that the Quality Target Product Profile (QTPP) must be similar to the reference product, ensuring that the storage and dosing (preparation, administration route, doses etc.) are similar. Thus the drug product presentation must match that of the reference product in dosage form, container/closure, and strength (although not all strengths are required if multiple exists) (U.S. FDA 2017a). This reduces risk for confusion and (medication) errors by the physician, pharmacist and the patient.

There are three key interacting components to the proper and successful development of a parenteral drug product: the formulation, the container/closure system, and the manufacturing process. The same holds true for a biosimilar with additional requirements arising from the need to match the reference in some key attributes. The formulation of a biosimilar does not have to be the same as that of the innovator as long as the type and quantity of degradation is similar (or lower) in comparison to the reference product. Intellectual property considerations can play a significant role in the design of the formulation. The container/closure system significantly impacts various elements of the drug product, including stability, product quality, but also extractable and delivered volume. In the case where the reference product is presented in an administration device (e.g. prefilled syringe, autoinjector or pen injector), the type of a device impacts usability and possibly product quality. Matching the usability profile, especially for home-use products can become a significant challenge, although a "similar" device presentation is not required. The design of the drug product manufacturing process of the reference is often not available or known. While there is no requirement to match the process, the impact of (drug product) process design and process parameters must be well understood, as would be the case for any biologic. The following sections of this chapter discuss these various drug product considerations in greater depth, using examples from approved products.

In the subsequent text, the nomenclature "innovator product" and "reference (medical) product" is used interchangeably.

Development Strategies for Biosimilars Drug Products

Molecule and Drug Substance

As a general concept, biosimilars are required to have the same primary sequence as the originator product. However, given that cell line, fermentation/bioreactor and purification processes are to a large extent different to the originator process, there likely are some differences in some characteristics e.g. glycosylation, charge isomers, sequence variants etc. Modern high resolving analytical methods can detect small differences in structure and sequence, the significance of which has to be understood in the context of their impact on safety and efficacy, and differences may be deemed non-critical or critical to remedy (Schiestl et al. 2011). For example, in the recent filing from Mylan for a biosimilar for trastuzumab (Herceptin), minor

differences were detected in the glycosylation profile (higher total mannose, lower non-glycosylated heavy chain, and slightly higher total sialic acid), although the sites, occupancy and species were same. Mannose and sialic acid content can impact the pharmacokinetics of the molecule, while glycosylation on the heavy chain can impact the effector function. However antibody-dependent cellular cytotoxicity (ADCC) activity using cell-based bioassay and Fc receptor binding kinetics showed no differences and thus the differences were judged to be of no significance. The same assays were used to show that deamidated species did not have any impact on target binding and potency, allowing the lower level of acidic species in the biosimilar to be considered acceptable. The site related to the differences in acidic species level was identified as an (light chain) asparagine in the Her2 binding region (U.S. FDA 2017c).

Process differences can also lead to differences in process residuals and impurities. For example, host-cell proteins (HCP) for a biosimilar are likely to have a different profile than the originator product. In most cases, the process impurities may not be important if reduced to levels where they have no biological impact, but in an early version of Omnitrope, these impurities led to a higher level of anti-drug antibodies (ADAs) compared to the reference (Thakrar et al. 2010). The purification process was subsequently revised to reduce HCP levels, prior to approval.

Drug Product Considerations for Biosimilars

Considerations for Assessment of Similarity

The requirements of similarity begin at the sequence and structure level including post-translational modifications, and extend to purity, charge variants, degradation products, aggregation levels, and of course functional and/or potency. With nearly all attributes, a multiplicity of (preferably orthogonal) methods, are required. Since product-related impurities can include potential degradants, levels of these have to be assessed in comparison to the reference product, keeping in mind the potential age of the reference. For this purpose, securing a supply of reference product over a period of time and if possible of different ages, is a good tactic. These should be enrolled in stability studies side-by-side with the biosimilar product, to follow at least the kinetics of the degradation, even if the starting points are not really a true time zero for the reference. In the worst case, if a degradant is seen that is not present in the reference, detailed characterization would be required to understand its origin, structure and potential safety and clinical impact. If the degradation pathways and products are the same as reference, the rate of degradation (and possibly the absolute amounts) should be assessed to show that the biosimilar has a slower or equal rate compared to the reference. Accelerated and forced degradation (acidic, basic, oxidative, photo, thermal) studies are commonly utilized for this purpose.

The range of analytical methods employed in extended characterization and for stability is critical to the success. A recent white paper provides good insight and case studies for the rational selection, criticality assessment, and ranking (tier levels)

of quality attributes and test methods to determine analytical similarity between biosimilars and their reference product (Vandekerckhove et al. 2018).

The strategy for performing the above side-by-side assessment requires some thought since the matrix in which the active agent is dissolved, can have a significant impact on the outcomes. The strategy is therefore determined by the objective of the exercise. When studying or comparing the molecules themselves, e.g. for structural comparisons by FTIR, CD or DSC, any (solution) compositional differences may have an impact. In this case, removing the matrix effect by buffer exchanging into a common matrix will be required. The same would hold for forced degradation studies looking at the intrinsic stability of the molecule produced by the reference vs the biosimilar. On the other hand, shelf-life determining stability studies have to be conducted in the intended matrix, while comparing the rate of degradation over time. In this case, the formulation matrix between the innovator and the biosimilar could be different.

Formulations for Biosimilars

A general formulation of a biologic drug product contains the active pharmaceutical ingredient (API) i.e. the protein molecule in an aqueous solution with excipients that provide buffering and maintenance of pH (buffers), with excipients that provide stabilization and tonicity (sugars, surfactants, amino acids, antioxidants, chelators). Multi-use products will generally include anti-microbial preservatives.

An overview of products approved (or rejected) in the EU is provided in Table 10.1. Biosimilar products approved in the US are shown in Table 10.2. Table 10.2 also shows the comparison of the biosimilar versus its reference product composition (formulation) and other product details. It is apparent that there are two general scenarios: (a) the biosimilar has the same formulation as its reference (medical) product, or (b) there may be one or more changes to the formulation in the biosimilar compared to its reference product.

At the outset, it may appear preferable to use the same formulation as the innovator for the biosimilar since detailed composition information is available in the US package insert. This could minimize risk or concerns that there may arise differences in formulation-driven degradation profile or impurities and thus, impact the clinical efficacy or safety. [It must be noted however that even if the composition of the biosimilar and reference product are the same, it is not assured that the performance and degradation would be similar. This may happen for example, if different quality standards or suppliers of (critical) excipients are used]. However, there are likely to be scenarios, where the modification of the formulation in the biosimilar (compared to the reference product) is required such as when a blocking formulation patent is still in force. Clearly, a review of the intellectual property situation around the molecule and product is the first step in defining the formulation development strategy.

Setting aside intellectual property (IP) drivers, a number of biologics that are being evaluated for biosimilar development may also have suboptimal formulations

Table 10.1 Biosimilar products filed in the EU (including refused or withdrawn) (all information from summary of product characteristics/European Public Assessment Reports) (Accessed April 2018) http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Abasaglar (previously Abasria)	Insulin Glargine	Eli Lilly Regional Operations GmbH	Authorized	3.64 mg/mL	Zinc oxide, hydrochloric acid + sodium hydroxide (for pH adjustment)	Glycerol		Metacresol	Liquid	Subcutaneous
Abseamed	Epoetin Alfa	Medice Arzneimittel Pütter Gmbh &Co. Kg	Authorized	0.0168 mg/mL, 0.084 mg/mL, 0.336 mg/mL	Sodium dihydrogen phosphate dihydrate, disodium phosphate dihydrate, hydrochloric acid + sodium hydroxide (for pH-adjustment)	Glycine, sodium chloride	Polysorbate 80		Liquid	Intravenous, subcutaneous
Accofil	Filgrastim	Accord Healthcare Ltd.	Authorized	0.6 mg/mL	Acetic acid, sodium hydroxide	Sorbitol (E420)	Polysorbate 80		Liquid	Intravenous, subcutaneous

Alpheon	Recombinant Human Interferon Alfa-2a	Biopartners Gmh	Refused	3×10^6 IU/mL (label claim 6×10^6 IU/mL)		during develop-ment lyophilised drug contained HSA, in liquid formulation replaced by another stabilizer			Liquid	
Amgevita	Adalimumab	Amgen Europe	Authorized	50 mg/mL	Glacial acetic acid, sodium hydroxide	Sucrose	Polysorbate 80		Liquid	Subcutaneous
Bemfola	Follitropin Alfa	Gedeon Richter Plc.	Authorized	0.044 mg/mL	Disodium hydrogen phosphate dihydrate, sodium dihydrogen, phosphate dihydrate, phosphoric acid pH 6.7–7.3	Sucrose, methionine	Poloxamer 188		Liquid	Subcutaneous
Benepali	Etanercept	Samsung Bioepis UK Limited (Sbuk)	Authorized	25 mg syringe	Sodium dihydrogen phosphate monohydrate, disodium hydrogen phosphate heptahydrate	Sucrose, sodium chloride			Liquid	Subcutaneous

(continued)

Table 10.1 (continued)

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Binocrit	Epoetin Alfa	Sandoz Gmbh	Authorized	0.0168 mg/mL—0.336 mg/mL	Sodium dihydrogen phosphate dihydrate, disodium phosphate dihydrate, hydrochloric acid + sodium hydroxide (for pH-adjustment)	Glycine, sodium chloride	Polysorbate 80		Liquid	Intravenous, subcutaneous
Biograstim	filgrastim	Abz-Pharma Gmbh	Withdrawn	0.6 mg/mL	Sodium hydroxide, glacial acetic acid	Sorbitol (E420)	Polysorbate 80		Liquid	Intravenous, subcutaneous
Blitzima	Rituximab	Celltrion	Authorized	10 mg/mL	tri-sodium citrate dihydrate	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution
Cyltezo	adalimumab	Boehringer Ingelheim International GmbH	Authorized	50 mg/mL	Sodium acetate trihydrate, glacial acetic acid	α,α-trehalose dihydrate	Polysorbate 80		Liquid	Subcutaneous

Epoetin Alfa Hexal	Epoetin Alfa	Hexal Ag	Authorized	0.016 mg/mL, 0.084 mg/mL, 0.336 mg/mL	Sodium dihydrogen phosphate dihydrate, disodium phosphate dihydrate, sodium chloride hydrochloric acid + sodium hydroxide (for pH-adjustment)	Glycine	Polysorbate 80		Liquid	Intravenous, subcutaneous
Erelzi	Etanercept	Sandoz GmbH	Authorized	20 mg, 50 mg/pre-filled syringe, pen	Citric acid anhydrous, sodium citrate dihydrate, sodium hydroxide, hydrochloric acid	sucrose, L-lysine hydrochlo-ride, sodium chloride			Liquid	Subcutaneous
Filgrastim hexal	Filgrastim	Hexal AG	Authorized	0.6 mg/mL, 0.96 mg/mL	Glutamic acid	Sorbitol	Polysorbate 80		Liquid	Intravenous, subcutaneous
Filgrastim ratiopharm	Filgrastim	Ratiopharm Gmbh	Withdrawn	0.6 mg/mL	Sodium hydroxide, glacial acetic acid	Sorbitol	Polysorbate 80		Liquid	Intravenous, subcutaneous

(continued)

Table 10.1 (continued)

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Flixabi	Infliximab	Samsung Bioepis UK Limited (SBUK)	Authorized	10 mg/mL after reconstruction	Monobasic sodium phosphate monohydrate, dibasic sodium phosphate heptahydrate	Sucrose	Polysorbate 80		Lyo	Intravenous use after reconstruction and dilution
Grastofil	Filgrastim	Apotex Europe Bv	Authorized	0.6 mg/mL	Acetic acid glacial, Sodium hydroxide	Sorbitol (E420)	Polysorbate 80		Liquid	Intravenous, subcutaneous
Imraldi	Adalimumab	Samsung Bioepis UK Limited (Sbuk)	Authorized	50 mg/mL	Sodium citrate, citric acid monohydrate, histidine, histidine hydrochloride monohydrate	sorbitol	Polysorbate 20		Liquid	Subcutaneous
Inflectra	Infliximab	Hospira UK Limited	Authorized	10 mg/mL after reconstruction	Sodium dihydrogen phosphate monohydrate, Disodium phosphate dihydrate	Sucrose	Polysorbate 80		Lyo	Intravenous use after reconstitution and dilution

Inhixa	Enoxaparin Sodium	Techdow Europe AB	Authorized	100 mg/mL	–	–	–		Liquid	Subcutaneous, intravenous, extracorporeal use (in the dialysis circuit)
Lusduna	Insulin Glargine	Merck Sharp & Dohme Limited	Authorized	3.64 mg/mL	Zinc chloride, hydrochloric acid, sodium hydroxide (for pH-adjustment)	Glycerol		Metacresol	Liquid	Subcutaneous
Movymia	Teriparatide	STADA Arzneimittel AG	Authorized	0.25 mg/mL	Glacial acetic acid, sodium acetate trihydrate, hydrochloric acid (for pH adjustment) and sodium hydroxide (for pH adjustment), pH 3.8–4.5	Mannitol		Metacresol	Liquid	Subcutaneous
Mvasi	Bevacizumab	Amgen Europe B.V.	Authorized	25 mg/mL	Sodium phosphate	α,α-trehalose dihydrate	Polysorbate 20		Liquid	Intravenous use after dilution

(continued)

Table 10.1 (continued)

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Nivestim	Filgrastim	Hospira UK Ltd	Authorized	0.6 mg/mL, 0.96 mg/mL	Acetic acid, glacial, Sodium hydroxide	Sorbitol (E420)	Polysorbate 80		Liquid	Dilution in 5% glucose solution possible
Omnitrope	Somatropin	Sandoz GmbH	Authorized	1.3 mg/mL	Disodium hydrogen phosphate heptahydrate, sodium dihydrogen phosphate dihydrate	Glycine			Lyo	Subcutaneous use after reconstitution
Ontruzant	Trastuzumab	Samsung Bioepis UK Limited (SBUK)	Authorized	21 mg/mL after reconstruction	L-histidine hydrochloride monohydrate, L-histidine	α,α-trehalose dihydrate	Polysorbate 20		Lyo	Intravenous after reconstitution and dilution
Ovaleap	Follitropin Alfa	Teva Pharma B.V.	Authorized	0.044 mg/mL	Sodium dihydrogen phosphate dihydrate, sodium hydroxide (2M) (for pH adjustment), pH 6.8–7.2	Mannitol, methionine	Polysorbate 20	Benzyl alcohol, benzalkonium chloride	Liquid	Subcutaneous

Ratiograstim	Filgrastim	Ratiopharm GmbH	Authorized	0.6 mg/mL	Acetic acid, glacial, Sodium hydroxide	50 mg/mL Sorbitol	Polysorbate 80		Liquid	Intravenous, subcutaneous
Remsima	Infliximab	Celltrion Healthcare Hungary Kft.	Authorized	10 mg/mL after reconstruction	Sodium dihydrogen phosphate monohydrate, disodium phosphate dihydrate	Sucrose	Polysorbate 80		Lyo	Intravenous use after reconstitution and dilution
Retacrit	Epoetin Zeta	Hospira UK Limited	Authorized	3333 IU/mL, 10,000 IU/mL, 40,000 IU/mL	Disodium phosphate dihydrate, Sodium dihydrogen phosphate dihydrate, Sodium hydroxide (pH adjuster) hydrochloric acid (pH adjuster)	Glycine, leucine, isoleucine, threonine, glutamic acid, pheny-lalanine, sodium chloride, calcium chloride dihydrate	Polysorbate 20		Liquid	Intravenous, subcutaneous

(continued)

Table 10.1 (continued)

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Ritemvia	Rituximab	Celltrion Healthcare Hungary Kft.	Authorized	10 mg/mL	tri-sodium citrate dihydrate	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution
Rituzena (prev. Tuxella)	Rituximab	Celltrion Healthcare Hungary Kft.	Authorized	10 mg/mL	tri-sodium citrate dihydrate	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution
Rixathon	Rituximab	Sandoz Gmbh	Authorized	10 mg/mL	Sodium citrate, sodium hydroxide, hydrochloric acid	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution
Riximyo	Rituximab	Sandoz Gmbh	Authorized	10 mg/mL	Sodium citrate, sodium hydroxide, hydrochloric acid	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution

Silapo	Epoetin Zeta	Stada Arzneimittel AG	Authorized	3333 IU/mL	Disodium phosphate dihydrate, sodium hydroxide + hydrochloric acid (pH adjustment), sodium dihydrogen phosphate dihydrate	Glycine, leucine, isoleucine, threonine, glutamic acid, pheny-lalanine, sodium chloride, calcium chloride dihydrate	Polysorbate 20		Liquid	Intravenous, subcutaneous
Solumarv	Insulin human	Marvel Lifesciences Ltd	Refused	100 IU/mL	Hydrochloric acid and sodium hydroxide	16 mg/mL glycerol		3 mg/mL metacresol	Liquid	

(continued)

Table 10.1 (continued)

Biosimilar proprietary name	Common name	Owner	Status	Concentration	Buffer system	Stabilizer/ tonicifier	Surfactant	Preservative	Form	Administration
Solymbic	Adalimumab	Amgen Europe B.V.	Authorized	50 mg/mL	Glacial acetic acid, sodium hydroxide (for pH adjustment)	Sucrose	Polysorbate 80		Liquid	Subcutaneous
Terrosa	Teriparatide	Gedeon Richter Plc.	Authorized	0.25 mg/mL	Glacial acetic acid, sodium acetate trihydrate, hydrochloric acid + sodium hydroxide (for pH adjustment), pH 3.8–4.5	Mannitol		Metacresol	Liquid	Subcutaneous
Tevagrastim	Filgrastim	Teva GmbH	Authorized	0.6 mg/mL	Acetic acid, glacial, sodium hydroxide	Sorbitol (E420)	Polysorbate 80		Liquid	Intravenous, subcutaneous

Thorinane	Enoxaparin Sodium	Pharmathen S.A.	Authorized	100 mg/mL	–	–	–		Liquid	Intravenous, subcutaneous
Truxima	Rituximab	Celltrion Healthcare Hungary Kft.	Authorized	10 mg/mL	tri-sodium citrate dihydrate	Sodium chloride	Polysorbate 80		Liquid	Intravenous use after dilution
Valtropin	Somatropin	BioPartners GmbH	Withdrawn	3.3 mg/mL	Sodium phosphate monobasic, sodium phosphate dibasic, sodium hydroxide + hydrochloric acid pH adjustment,	Mannitol, glycerine		Metacresol (solvent)	Lyo	Subcutaneous
Zarzio	Filgrastim	Sandoz GmbH	Authorized	0.6 mg/mL, 0.96 mg/mL	Glutamic acid	Sorbitol (E420)	Polysorbate 80		Liquid	Intravenous, subcutaneous

Table 10.2 Comparison of biosimilar products approved in the US with the reference product

Biosimilar product (common name) [manufacturer]	Biosimilar product details	Reference product (common name) (manufacturer)	Reference product details	Comments
Zarxio® (filgrastim-sndz) [Sandoz]	• Glutamate buffer, Sorbitol, Polysorbate 80 pH 4.4 • 300 mcg/0.5 mL and 480 mcg/0.8 mL in PFS with needle guard	Neupogen® (filgrastim) [Amgen]	• Acetate buffer, sorbitol, polysorbate 80 pH 4.0 • 300 mcg/mL and 480 mcg/1.6 mL in single-dose vials • 300 mcg/0.5 mL and 480 mcg/0.8 mL in PFS with Needle Guard (27G, 1/2 inch needle)	Different buffer and pH due to intellectual property considerations
Inflectra® (infliximab-dyyb) [Celltrion]	• Monobasic sodium phosphate monohydrate, dibasic sodium phosphate dihydrate, Sucrose, polysorbate 80 pH 7.2 • Lyophilized, 100 mg in a 20-mL single-use vial, Reconstitute with 10 mL sWFI • Store at 2–8 °C (36–46 °F)	Remicade® (infliximab) [Janssen]	• Monobasic sodium phosphate monohydrate, dibasic sodium phosphate dihydrate, Sucrose, Polysorbate 80 pH 7.2 • Lyophilized, 100 mg in a 20-mL vial, reconstitute with 10 mL sWFI • Store at 2–8 °C (36–46 °F) unopened vials may also be stored up to maximum of 30 °C (86 °F) for a single period of up to 6 months (not exceeding original expiry date)	Same composition as reference product
Renflexis® (infliximab-abda) [Samsung Bioepis]	• Monobasic sodium phosphate monohydrate, dibasic sodium phosphate heptahydrate, Sucrose, Polysorbate 80 pH ~6 • Lyophilized, 100 mg in a 20-mL single-use vial, Reconstitute with 10 mL sWFI • Store at 2–8 °C (36–46 °F)			Same composition as reference except the dibasic sodium phosphate which is a heptahydrate instead of dihydrate form in reference product

Ixifi™ (Infliximab-qbtx) [Pfizer]	• Disodium succinate hexahydrate, succinic acid, sucrose, polysorbate 80 pH ~6 • Lyophilized, 100 mg in a 15-mL single-use vial, reconstitute with 10 mL sWFI • Store at 2–8 °C (36–46 °F) unopened vials may also be stored up to maximum of 30 °C (86 °F) for a single period of up to 6 months (not exceeding original expiry date)			Different buffer
Erelzi® (etanercept-szzs) [Sandoz]	• Citric acid, sodium citrate, sodium chloride, sucrose, L-Lysine HCl pH 6.3 ± 0.2 • 25 mg/0.5 mL and 50 mg/mL in PFS with needle guard (27G, 1/2 inch needle) • 50 mg/mL in prefilled Pen • Store at 2–8 °C (36–46 °F) storage of individual syringes or pens at room temperature (20–25 °C, 68–77 °F) for a maximum single period of 28 days is permissible	Enbrel® (etanercept) [Amgen]	• Sodium chloride, Sucrose, L-Arginine HCl pH 6.3 ± 0.2 • 25 mg/0.5 mL and 50 mg/mL in PFS with Needle Guard (27G, 1/2 inch needle) • 50 mg/mL in prefilled Autoinjector • 50 mg/mL in prefilled cartridge for use with autoinjector • Store at 2–8 °C (36–46 °F) storage of individual syringes or autoinjectors at room temperature (20–25 °C, 68–77 °F) for a maximum single period of 14 days is permissible • Tromethamine, mannitol, sucrose • Lyophilized, 25 mg multi-dose vial, reconstitution solution [1 mL bWFI (0.9% benzyl alcohol)] supplied in PFS, final pH 7.4 ± 0.3 • Also included—needle (27G, 1/2 inch) and vial adapter	Erelzi® formulation contains citrate and lysine instead of phosphate and arginine in Enbrel formulation due to intellectual property considerations

(continued)

Table 10.2 (continued)

Biosimilar product (common name) [manufacturer]	Biosimilar product details	Reference product (common name) (manufacturer)	Reference product details	Comments
Amjevita® (adalimumab-atto) [Amgen]	• Glacial acetic acid, (NaOH), sucrose, polysorbate 80 pH 5.2 • 40 mg/0.8 mL and 20 mg/0.4 mL in PFS 40 mg/0.8 mL in autoinjector	Humira® (adalimumab) [AbbVie]	• Original formulation: monobasic sodium phosphate dihydrate, dibasic sodium phosphate dihydrate, sodium citrate, citric acid monohydrate, sodium chloride, mannitol, polysorbate 80, pH ~5.2—for dosages 40 mg/0.8 mL, 20 mg/0.4 mL, 10 mg/0.2 mL • Revised formulation: mannitol, polysorbate 80 pH ~5.2—for high concentration dosages 80 mg/0.8 mL, 40 mg/0.4 mL, 20 mg/0.2 mL, 10 mg/0.1 mL • 80 mg/0.8 mL, 40 mg/0.8 mL, 40 mg/0.4 mL, 20 mg/0.4 mL, 20 mg/0.2 mL, 10 mg/0.2 mL, 10 mg/0.1 mL in PFS (27G, 1/2 inch thin-wall needles for high concentration doses) • 80 mg/0.8 mL, 40 mg/0.8 mL, 40 mg/0.4 mL in prefilled pen (27G, 1/2 inch thin-wall needles for high concentration doses) • 40 mg/0.8 mL in single-use vial • Store at 2–8 °C (36–46 °F) may be stored at room temperature up to a maximum of 77 °F (25 °C) for a period of 14 days	Different buffer and stabilizer

Cyltezo™ (adalimumab-abdm) [Boehringer Ingelheim]	• Glacial acetic acid, sodium acetate trihydrate, trehalose dihydrate, polysorbate 80 pH ~5.2 • 40 mg/0.8 mL in PFS	Different buffer and stabilizer		
Mvasi™ (bevacizumab-awwb) [Amgen]	• Sodium phosphate monobasic monohydrate, sodium phosphate dibasic anhydrous, α,α-Trehalose dihydrate, polysorbate 20 pH 6.2 • 100 mg/4 mL in single-use vial 400 mg/16 mL in single-use vial	Avastin® (bevacizumab) [Genentech]	• Sodium phosphate monobasic monohydrate, Sodium phosphate dibasic anhydrous, α,α-Trehalose dihydrate, polysorbate 20 pH 6.2 • 100 mg/4 mL in single-use vial 400 mg/16 mL in single-use vial	Same composition as reference product
Ogivri™ (trastuzumab-dkst) [Mylan]	• Histidine, Histidine hydrochloride monohydrate, Sorbitol, PEG3350 pH ~6 • Lyophilized, 420 mg per multi-dose vial, reconstitute with 20 mL bWFI (1.1% benzyl alcohol) or 20 mL sWFI, bWFI is supplied in carton	Herceptin® (trastuzumab) [Genentech]	• Histidine, Histidine hydrochloride monohydrate, α,α-Trehalose dihydrate, polysorbate 20 pH ~6 • Lyophilized, 150 mg per single-use vial, reconstitute with 7.4 mL sWFI (not supplied) • 420 mg per multi-dose vial, reconstitute with 20 mL bWFI (1.1% benzyl alcohol) or 20 mL sWFI, bWFI is supplied in carton	Different stabilizers due to intellectual property considerations

(continued)

Table 10.2 (continued)

Biosimilar product (common name) [manufacturer]	Biosimilar product details	Reference product (common name) (manufacturer)	Reference product details	Comments
Retacrit (epoetin alfa-epbx) [Hospira/Pfizer]	• Sodium phosphate dibasic anhydrous, Sodium phosphate monobasic monohydrate, $CaCl_2 \cdot 2H_2O$, glycine, isoleucine, leucine, L-glutamic acid, phenylalanine, threonine, NaCl, Polysorbate 20; pH not listed (but is different per FDA briefing package) • Unit dose vials 2000, 3000, 4000, 10,000, 40,000 units/mL	Epogen/Procrit (epoetin alfa) [Amgen] [Janssen]	• Single dose: citric acid, sodium citrate, NaCl, human serum albumin pH 6.9 ± 0.3 • Multi-dose: citric acid, sodium citrate, NaCl, human serum albumin, benzyl alcohol pH 6.1 ± 0.3 • 2000, 3000, 4000, 10,000 units/mL in single-dose vials • 20,000 units/2 mL and 20,000 units/mL in multi-dose vials	Different buffer and stabilizers
Fulphila (pegfilgrastim-jmdb) [Mylan]	• (Sodium) acetate, sorbitol, polysorbate 20 pH 4.0 • 6 mg/0.6 mL in PFS (29G, 1/2 inch with needle guard) • Store between 2 and 8 °C, discard syringes stored at room temperature for more than 72 h	Neulasta (pegfilgrastim) [Amgen]	• (Sodium) acetate, sorbitol, polysorbate 20 pH 4.0 • 6 mg/0.6 mL in PFS (27G, 1/2 inch with Needle guard) • On-body Injector to be used with 6 mg/0.6 mL PFS (27G, 1/2 inch) supplied together • Store between 2 and 8 °C, discard syringes stored at room temperature for more than 48 h	Same composition as reference product

All information from US FDA Drugs@FDA website https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm; and from Prescribing Information/Labels. Accessed June 2018

that could be readily improved with current knowledge, analytical techniques and experience. The objective is to have as good if not better stability. While not entirely clear, it is possible that a longer shelf-life could be approved if the stability is better for the biosimilar. On the other hand, there are clear examples of differences in "in-use" periods (discussed later). An improved formulation may also be warranted in cases where immunogenicity is a specific concern due to an increased awareness and understanding of the potential risk posed by (subvisible) particulate matter and aggregation (Carpenter et al. 2009; Rosenberg 2006; Singh et al. 2010).

One of the key formulation development decisions is the pH (and buffer system). Screening studies may find that a pH different than that of the reference is optimal. The decision to go with a different pH will require consideration of the differences in (rates of) degradation pathways and products due to the pH. For proteins, differences in pH and to some extent the buffer species, can impact rates of deamidation, fragmentation and aggregation (Manning et al. 2010). Detailed analysis of the chemical degradants and sites of modification can be obtained while aggregates are only defined by their size [e.g. oligomerization state by size exclusion chromatography (SEC) or analytical ultra-centrifugation]. A risk-based decision on pH selection would be required by relating the relative risks posed by the degradants to safety and efficacy. From the examples in Table 10.2, it can be seen that in three cases, the pH differs between reference and biosimilar—for filgrastim (Zarxio and Neupogen), for infliximab (Ixifi and Remicade), and for epoetin alfa (Retacrit and Epogen/Procrit). Zarxio and Neupogen are liquid products with pHs 4.4 (glutamate buffer) and 4.0 (acetate buffer) respectively, the difference being driven by IP reasons. However, stability data by RP-HPLC (oxidized and deamidated/norleucine variants) and SEC (aggregates) showed that there was no difference in degradation profiles (Sandoz 2015), likely because the difference in pH is relatively small. Higher order structure of the reference and biosimilar molecules (after exchanging into the biosimilar and the reference pH/buffer systems) were also compared using natural isotope abundance 2D NMR (^{1}H-^{15}N HSQC), to show that the structures were similar when in similar formulations. [Spectra were also recorded with the biosimilar molecule in different pHs 3.0, 4.0, 4.4 to show the ability to detect changes in chemical shift of N-H signal due to different environment] (U.S. FDA 2015a). In the case of the infliximab products, the pH difference is much larger (6.0 in succinate buffer compared to 7.2 in a phosphate buffer for Ixifli and Remicade respectively). However, the products are lyophilized and the impact of pH difference during the short holds in liquid state (drug substance, pre-lyophilization and reconstituted solution during use) is probably not significant. It is quite likely that the drug substances in both cases are stored frozen so that the difference in pH does not lead to a divergence over time in quality characteristics after release (Singh 2007; Singh and Nema 2010). Plus, it is possible that the phosphate buffer in Remicade results in the well-known phenomenon of drop in pH on freezing to −70 °C, bringing it inadvertently closer to pH 6. Unfortunately, the pH of Retacrit is not disclosed in the prescribing information but the FDA briefing package states that the pH is different than that of the Reference. Retacrit is based on a phosphate

buffer so the pH is likely close to the pH 6.9 ± 0.3 of Epogen/Procrit (U.S. FDA 2017d).

Choice of structural stabilizers such as sucrose, trehalose or sorbitol will also require consideration of compatibility with the storage aspects of the drug substance. Sucrose is the most accommodating while trehalose and to a lesser extent sorbitol carry risks for frozen-state storage (Piedmonte et al. 2007; Singh 2011, 2018). Since there is no regulatory requirement to match drug substance degradation profiles or storage conditions or shelf-lives, the decision of structural stabilizer will be driven by drug product stability, as well as operational reasons. Comparing the product formulations in Table 10.2 shows several variations although sucrose, when used, is generally present in both.

Surfactants are commonly added to biologics products to stabilize against interfacial stresses the protein may encounter during production, storage, shipping, use etc. and are often considered critical excipients. Polysorbate 20 and 80 are the most commonly used surfactants. If present in the reference, the decision to also add to the biosimilar is simple (barring IP concerns). The choice between polysorbate 80 or 20 is often driven by current practices in the company, but staying with the version in the reference is a good strategy to avoid differences in polysorbate-induced degradants/impurities (Kishore et al. 2011a, b). On the other hand, when a surfactant is not present in the reference, the decision to add or not add to the biosimilar could be more difficult, especially if an advantage of the addition is seen in biosimilar formulation studies (e.g. freeze/thaw or agitation). However, Table 10.2 shows that in the one case of a liquid product where there is no surfactant in the reference (Enbrel), there is none in the biosimilar also (Erelzi). The Herceptin (trastuzumab) biosimilar Ogivri has a significantly different formulation and no surfactant compared to polysorbate 20 present in the reference. In this case though, the lyophilized product presentation significantly reduces the risk of not having a surfactant in the product. For the corresponding drug substance (likely stored frozen), lack of surfactant could be a risk for aggregation/particulate formation but is possibility mitigated by the addition of PEG3350 (Singh and Nema 2010). Frozen state storage of sorbitol-containing solutions must ensure that the potential for crystallization of a pseudo-polymorph of sorbitol is avoided (Piedmonte et al. 2007, 2015).

A review of some approved drug products shows that differences in formulations will enable different (presumably longer) shelf lives. Although the shelf-life of the Humira products is not readily available, Amjevita syringes have 30 months and Cyltezo syringes have 24 months at 2–8 °C (U.S. FDA 2016a, 2017e). Similarly, Renflixis has 30 months, Inflectra has 51 months, and Ixifli has 42 months at 2–8 °C (U.S. FDA 2016b, 2017f, 2017g). Renflixis and Inflectra labels also do not allow storage at 30 °C for up to 6 months, unlike the labels for Ixifi and the reference Remicade. The current approved shelf-lives may simply be a function of the amount of data provided and future supplements are likely to result in these being updated. Interestingly, the Erelzi prescribing information allows storage at room temperature (20–25 °C) for up to 28 days (U.S. FDA 2016c), while the corresponding Enbrel label allows 14 days.

For biosimilar development, it is important to understand that a different (improved) stability of the biosimilar must be due to the improved formulation and not due to differences in the molecule. The chemistry review of Amjevita underscores this as the data shows that the biosimilar is substantially more stable than Humira as seen in a forced degradation study at 50 °C. However, this difference in stability is explained by the differences in formulations because the ABP501 (the Amjevita molecule) has a higher rate of aggregation when formulated in the Humira formulation. The Reviewer concludes that the improved stability of Amjevita over Humira is due to the formulation and not due to intrinsic differences in the molecule (U.S. FDA 2016d).

Among the standard drug product quality parameters which are often obligatory critical quality attributes such as osmolality and pH, there is no expectation of being similar or identical to reference (but should be consistent within their defined range). Appearance (visible particles/color/clarity) are expected to be largely similar assessed at Tier 3 qualitative comparison. Deliverable volume and protein concentration are required to match over at least a Tier 2 quality range (see e.g. Amjevita Chemistry Review (U.S. FDA 2016d).

Differences in formulation selected can have an impact on various pre-clinical and clinical study designs as was the case for Zarxio, where supportive toxicology and toxicokinetic studies were conducted using the Zarxio molecule in the Neupogen formulation. In pivotal safety studies, matching dilution buffers were prepared such that each product was diluted in its own buffer in order to make the study most predictive of clinical outcome (Sandoz 2015). Note that the FDA guidance recommends that if differences in manufacturing (e.g., impurities or excipients) between the proposed biosimilar product and the reference product may result in different immunogenicity, measurement of anti-protein antibody response in animals may be important for assessing patient safety (U.S. FDA 2015b). Again, in the case of Zarxio, such a comparative immunogenicity assessment was completed. The objective was not the assessment of immunogenicity per se, but to assess the likelihood for relative differences in immunogenicity (Sandoz 2015).

A final aspect of formulation development is the consideration of in-use studies to simulate clinical administration and to ensure that the intended dose can be delivered with the appropriate quality (Ricci et al. 2015). This is especially important for intravenous infusion products since a dilution step in a suitable diluent is often required. A review of the labels shows that the Dosage and Administration Instructions are equivalent between biosimilar and reference for such products. [For the subcutaneous injection products, the in-use stability data in primary container/closure system is provided as part of the stability studies, and may lead to slightly different label text, as discussed above]. Examination of the BLA reviews of the biosimilars did not show examples of any specific studies except a brief mention in the Inflectra Chemistry Review as part of the review of additional particulates data (U.S. FDA 2016e). However, it is likely that such in-use stability studies are performed and the data provided as part of the Compatibility section in section P.2 of the BLA. The authors further believe that these studies are mandatory if the formulation of the biosimilar is different than that of the reference. Along the

same lines, the FDA raised the need to perform microbiological growth potential studies after container penetration and after preparation of infusion solutions (Lolas and Metcalfe 2011; Metcalfe 2009). Considering that this is a relatively recent requirement, it is unlikely that such a study had been performed by the reference. It is unclear to the authors, if the biosimilars are being held to this requirement, or that the label text for in-use duration is simply direct ported over from the reference and applied (likely).

In summary, formulation development strategy for a biosimilar requires a consideration of IP followed by a consideration of what is best for the molecule compared to the reference formulation. It is important that the molecule itself is intrinsically similar; improved intrinsic (structural, conformational or colloidal) stability may lead to it being classified as a "biobetter". Improved stability due to improved formulation is acceptable as long as the molecule is similar. Following the same logic, one may try to presume that reduced stability due to a poorer formulation should be acceptable as long as the molecule is similar—however, we would strongly recommend against this as an excuse for poor formulation development as the risk is too high.

Dosage Forms for Biosimilars

In theory, the dosage form may also differ between biosimilar and originator product, i.e. whether a liquid or freeze-dried formulation is being used for the biosimilar. A change of this type is usually not advisable in order to minimize risk for types or levels of degradants to occur during shelf-life of the biosimilar, that may question the comparability and ability to reference.

Among the current approved products in the US, there are no differences in dosage forms although the biosimilar may not have all the variations or presentations (strengths, dosage forms) available as the reference. For example, Erelzi does not have a multi-dose lyophilized vial or a (pre-filled cartridge + reusable autoinjector) presentation similar to Enbrel. Humira comes in a number of dose strengths in syringes, pens and also a vial. Not all these doses are similarly available in Amjevita or Cyltezo. This is likely simply due to the biosimilar companies making the decision to come into the market with one or few key doses and potentially add others over time based on business drivers. This can in certain cases limit the indications or patient groups that can use the biosimilar. Amjevita, in its label does not cover juvenile idiopathic arthritis (JIA) patients below 15 kg of weight since the low dose pre-filled syringe (10 mg) is not available and dosing half the volume from the 20 mg syringe is not indicated. Similarly, Cyltezo is not indicated for JIA patients below 30 kg of weight since it is available in only one strength (40 mg syringe). It must be noted however that Humira has orphan drug marketing exclusivity for JIA patients aged 2 years to less than 4 years until 30 Sep 2021 (U.S. FDA 2018a), and the lack of low dose strength variations of Amjevita and Cyltezo is probably related to this situation. Erelzi cannot be dosed to JIA patients below 63 kg since there is no (multidose lyophilized) vial product (akin to Enbrel) to enable weight–based dosing.

Development of the dosage form for a biosimilar requires a decision on container/closure (c/c) and overfill to allow label equivalency to be met. (A specific discussion on the quality aspects of the c/c is covered later). Among the liquid in vial products, there is not sufficient information to assess if there are any differences in sizes of c/c used or in overfill amounts. Information about the overfill amount in the 150 mg Herceptin vial is provided in the label and presumably helps to set a target for the biosimilars. Availability of this information is unusual though, and in most cases multiple batches of the reference product have to be thoroughly analyzed to assess the overfill level in relation to labelled dose. [This is easier for PFS presentations although the same principle applies]. The US FDA expects the concentration, deliverable volume and thus the dose to be matched to the reference. Since the dose is calculated based on concentration, assessment of concentration and therefore the determination/selection of the extinction coefficient is an extremely critical exercise and subject to significant scrutiny in the BLA review process. A good understanding of the overfill is particularly critical for lyophilized products since reconstitution with labelled amount of diluent is performed to obtain a product solution purportedly at "target concentration". This information can be used to assess extinction coefficient or if an independent estimate of extinction coefficient has been made, the fill volume is back-calculated. Since fill volumes are always variable due to inherent variability in filling process, a significant degree of uncertainty is built into all these estimates.

The infliximab products, Inflectra, and Renflexis are lyophilized in 20-mL sized vials, similar to Remicade, while Ixifi is provided in a 15-mL vial (per their prescribing information). They are all indicated for reconstitution with 10 mL of sWFI. Since the 15R and 20R vials have different diameters, it is possible that the final overfills may be different, as long as the labelled 100 mg can be withdrawn in 10 mL. A related question for lyophilized products is the choice of pre-lyo concentration of the bulk drug product (BDP) filled into the vial prior to lyophilization. In principle, a higher concentration solution can be lyophilized as long as it has the same concentration as the reference after reconstitution. Furthermore, it is also important that the reconstitution diluent volume used is the same between reference and biosimilar to prevent confusion and errors. Therefore, using a different pre-lyo concentration of BDP for lyophilization can result in a difference in contribution of cake volume to the reconstituted product volume. This can lead to difficulties in matching final product concentration after reconstitution (as well as potentially the extractable volume/dose), when the underlying uncertainties of overfill amounts and extinction coefficients are factored in.

Container/Closure Systems and Devices for Biosimilars

The reference product usually sets the baseline and standard for the dosage form and presentation. For example, if an reference product is presented in a ready-to-use, prefilled syringe, it is usually advisable from a variety of standpoints to (at least) match the standard of the originator. However, various scenarios are possible (see Table 10.3).

Table 10.3 Scenarios for choice of primary packaging and device for biosimilar vs reference products

Biosimilar	Reference	Comment
Vial/stopper	Vial/stopper	Vial and stopper supplier and quality may differ.
Syringe (e.g., SIN-PFS)[a] or syringe in autoinjector	Vial	Expect impact from the syringe in the biosimilar, e.g., silicone oil, tungsten Requires extensive design verification of the biosimilar
Vial	Syringe (e.g., SIN-PFS) or Syringe in autoinjector	Expect other (lesser) process residuals in the vial compared to syringe. However, biosimilar could be less competitive
Syringe (e.g., SIN-PFS) or Syringe in autoinjector	Syringe (e.g., SIN-PFS) or Syringe in autoinjector	Syringe supplier and quality may differ Autoinjector very likely of different brand and type Requires extensive design verification of the biosimilar

[a]Staked-in needle—prefilled syringe

The formulation scientist for a biosimilar will have to address the question of whether the primary packaging should be selected to "match" the reference (assuming ability to use modern techniques to characterize and identify with sufficient confidence), or is the choice driven by the state-of-art and knowledge of what is the most appropriate current system. In other words, the container closure system of the reference may be somewhat outdated, given that it has likely been approved several years ago. This could be the case for the glass and certainly true for the rubber components (stopper, plunger, needle shield etc.). It is the authors' opinion that the biosimilar development organization must aim to use the highest quality c/c systems for their product from the outset.

For vials and syringes, the current standard is Type 1 glass. The vial supplier, glass type, and vial forming (process) parameters are likely to differ from the reference. While there is no specific reason that this would be a drawback, a good understanding of the impact of glass on product (and vice versa) is important. For example, it is known that different types of hydrolytic glass may show differences with regards to (a) leachables (Jahn 2018), (b) delamination risk (Ditter et al. 2018a, b), and (c) fogging, a commonly observed defect in lyophilized products (Abdul-Fattah et al. 2013). The latter phenomenon can be particularly vexing, since it may and often is dismissed as a cosmetic defect. However, if the product reaches under the stopper sealing surfaces, it may become a critical vial seal integrity defect. Proper selection of the components is therefore prudent to avoid this situation.

For rubber components, developments in technology have resulted in cleaner (from an extractables and leachables perspective) materials. These higher quality

rubbers with barrier coatings on the product contact side are the current norm and should be used in the biosimilar product, irrespective of the components used by the reference.

Prefilled syringes (PFS) were originally regulated as a drug product based on the primary mode of action and are now considered combination products (drug/device) under 21 CFR Part 4 (valid since 2013). Their development has to be under design controls and accordingly documented in a Design History File (U.S. FDA 2017b). Autoinjector and pen devices also obviously fall under this rule and represent an added regulatory complexity that may not have existed at the time of the approval of the reference (although the rule may apply retrospectively in case of an CMC update). Dimitrova et al. (Dimitrova et al. 2018) provide a good overview of the requirements for this and other aspects of the development of a PFS product.

Development of a PFS requires consideration of the impact of silicone oil and potentially tungsten on the quality of the product. As part of the selection and design criteria for the syringe, levels of silicone oil needed for lubrication and plunger movement, and residual levels of tungsten would have to be specified. Spiking studies with silicone oil and tungsten pin extract may be used to define these design levels to select the components. This should be followed with side-by-side storage stability studies between biosimilar and reference to compare impact of the PFS on product quality over long-term storage. The review of this section of the BLAs for biosimilars shows that the syringe design and performance are assessed independently of the reference (see e.g. Amjevita and Erelzi Other Reviews (U.S. FDA 2016f, g). However, the importance of control of silicone oil and tungsten can be illustrated through two examples discussed below.

While silicone oil in the syringe aids the lubrication, it also can slough off and form particulates, generally subvisible, in the product. Amgen chose to compare the total subvisible particulate count as measured by light obscuration in the Amjevita product batches with Humira batches using Tier 3 qualitative comparison criteria. Additionally, using flow imaging analysis (by MFI), they stratified the $\geq$5 μm subvisible particles population into spherical (arising from silicone oil) and non-spherical (proteinaceous origin) for both products. The $\geq$5 μm total particle counts and $\geq$5 μm non-spherical particle counts were compared between the two products using Tier 2 quality range criteria (U.S. FDA 2016d). Equivalence of the non-spherical particle counts would be considered important due to the potential impact of poteinaceous particles on immunogenicity of the product (Carpenter et al. 2009). (See Singh and Mahler (Singh and Mahler 2018), for a review of how subvisible particles have been generally addressed in a number of biosimilars). [The FDA considered these methods for Tier 3 qualitative comparison for Amjevita (U.S. FDA 2016d), but additional data requirements were requested from Celltrion for Inflectra due to initial concerns raised from higher immunogenicity rates observed with CT-P13 compared to Remicade (US) in a small single-dose trial (U.S. FDA 2016e). Celltrion evaluated the subvisible particles data using a Tier 1 equivalence test (Celltrion 2016), although the FDA seems to have only assessed the data as Tier 3 qualitative comparison (U.S. FDA 2016e).]

Tungsten pins are used during syringe manufacture to create the groove into which the needle is attached for staked-in needle PFS. Residues of tungsten may be left behind and can be leached into the product causing oxidation and aggregation of the protein (Liu et al. 2010). In a clinical trial of biosimilar epoetin alfa (HX575, Binocrit by Sandoz), two cases of neutralizing antibodies were reported. Since the presence of neutralizing antibodies against epoetin carries the very high risk of development of Pure Red Cell Aplasia in which the antibodies cross-react against endogenous erythropoietin, the cause of the formation of antibodies was subject to an extensive investigation. The likely cause was identified as being the presence of soluble tungsten in the syringes (Rubic-Schneider et al. 2017; Seidl et al. 2012). The authors recommend that high quality syringes be used for biologics in general, with stringent quality control criteria on the silicone oil and residual tungsten levels. Reference products will generally have been approved years ago when the state of knowledge and quality criteria of c/c systems were generally not as rigorous. However, for a biosimilar, investment in current state-of-the-art c/c systems makes good technical, strategic and business sense.

A number of the biosimilars and references are approved with prefilled syringes as the c/c. Here, one can see the impact of selected PFS design, especially due to the recent requirement of needle-stick injury prevention features (U.S. Department of Labor 2011; European Agency for Safety and Health at Work 2010). Zarxio has been approved in a PFS with a safety feature that interferes with the visibility of some graduation markings and thus prevents the lowest doses (0.1 mL and 0.2 mL) from being delivered accurately. Thus the label recommends to not administer the lowest doses directly from the syringe. The Neupogen PFS also has a safety feature but of a different design that does not have this limitation (per prescribing information). [This issue obviously does not arise when the entire syringe has to be dosed—for example with Erelzi PFS which has a needle-stick protection feature in contrast with Enbrel PFS which does not]. Furthermore, filgrastim is also dosed by the IV route after dilution in 5% dextrose. Neupogen is available in a vial from which the dose required for dilution is extracted (based on weight of patient). Since Zarxio is not available in a vial, the ability to prepare such an IV dose by extracting (=expressing) solution from the Zarxio syringe was an important review question. This use of the syringe in this fashion was approved only after review from the Division of Medication Error Prevention and Analysis at the FDA (U.S. FDA 2015c).

Another aspect of PFS selection is the choice of needle (in case of staked-in-needles (SIN) which is currently the situation for all products approved with PFS). Humira with its variety of product presentations uses two version of PFSs. Older (formulation) versions with (gauge not disclosed) 1/2 inch needle with needle cover that may contain latex, and newer (high concentration formulation) versions with thin-walled (gauge not disclosed) 1/2 inch needle with (black) needle covers that do not contain latex. The Amjevita PFS uses a 29G (presumably to match the needle gauge in Humira) 1/2 inch needle, but the needle cover may contain latex. Cyltezo PFS comes with a standard 27G 1/2 inch needle with a needle cover that may contain latex (see Table 10.2). Clearly, having a matching PFS presentation seems to be important, but the features of the PFS need not be matched

for regulatory approval. Differences in c/c are thus likely a calculated business risk and not a technical risk. This is also clear when it comes to devices to aid administration of injection. Amjevita uses an autoinjector (27G 1/2 inch needle SIN-PFS) while Humira uses a pen (needles as described above) and Cyltezo currently has none. Erlezi uses a pen while Enbrel has an autoinjector as well as a reusable autoinjector that is loaded with a prefilled cartridge (see Table 10.2). Instructions for Use (IFU) thus have to be developed and validated independently by the biosimilar organization for the device in question, and may or may not bear any resemblance to the instructions for use of the reference (apart from site of injection). This difference in administration aid does carry the risk of not getting approval for interchangeability since the IFUs between a pen and an autoinjector can be quite different and carry the risk of confusion for the patient if / when they are switched.

A good example of the type of testing required for devices (PFS and autoinjector) including human factors studies, is available from the Amjevita BLA reviews—other reviews (U.S. FDA 2016g). These tests are not done in a comparative (to reference) fashion but are the same as would be required for any new PFS or device product. Functionality tests, along with expelled product quality tests, have to be carried out on stability also, which implies that the program must plan to manufacture an adequate number of devices at least during validation. However, in all the above cases of the injection administration device being approved with the biosimilar, the biosimilar company has leveraged their in-house platform device. Thus Amjevita uses the same SureClick autoinjector platform as Enbrel, both from Amgen. Similarly, the Erlezi SensorReady pen is the same as used for Cosentyx by Novartis/Sandoz. This can simplify the development and approval since a large part of the design history file can be leveraged. Human factors studies may also be leveraged if patient population groups are similar (e.g. for Enbrel and Cosentyx) (U.S. FDA 2016f). In these situations, the FDA expects that the IFU (of the biosimilar) would closely follow the IFU of the (platform) device being leveraged, with only the product specific information being different.

Finally, as part of the container/closure development, an independent extractable and leachables program is required and would be similar to a standard program since the requirements are the same whether it is an innovator or a biosimilar. The safety of any leachables would be evaluated independently of the innovator, using current guidance from the regulators and the pharmacopeia (Jahn 2018).

Drug Product Manufacturing Process

Biologics drug products are almost universally manufactured using aseptic processes since terminal sterilization is not possible. Sterile filtration is the key step to sterilize the final product, although the whole process has to be designed towards ensuring a clean and sterile product. Furthermore, the drug product (DP) manufacturing process cannot improve upon the quality of the active molecule received from downstream processing, but a poorly designed DP process can

Table 10.4 Typical drug product manufacturing process unit operations

Unit operation	Factors for consideration	Risks
Drug substance thawing	Drug substance container, scale, freezing and thawing process Drug Substance formulation (fully formulated in comparison to DP or only partly; concentration *vis a vis* DP)	Generation of aggregates and particles
Pooling/homogenization/mixing	Number of containers/volumes; pooling process	Generation of aggregates and particles
Final formulation (compounding)	Not needed when DS is fully formulated. Otherwise, need to add excipients to final concentrations using compounding buffer	Generation of aggregates and particles
Sterile filtration	Choice of filter membrane type Choice of filtration mode and process parameters (e.g. pump, pressure; rate)	Adsorption of protein and critical excipients such as surfactants and preservatives Generation of aggregates and particles
Filling	Filling pump type Pump parameters Tubing used	Adsorption of critical excipients such as surfactants and preservatives Generation of aggregates and particles Target overfill and overfill ranges (impact on extractable volume as well as cost of goods)
Lyophilization (if relevant)	Lyophilizer design Lyophilization cycle	Product quality including cake appearance and related defects, e.g., fogging, (partial) collapse etc. Residual moisture Moisture transfer from stoppers
Stoppering and capping	Capping type Capping parameters Container closure system	Container closure integrity

certainly degrade the quality. Thus, a DP process has to be designed with quality in mind. There are of course no requirements for the biosimilar to match the reference product process since the specifics of the reference process are anyway unknown.

A general listing of the unit operations for DP manufacture is provided in Table 10.4 with potential risk factors for impact on product quality. The biosimilar development scientist needs to study the various unit operations in the context of

the chosen commercial facility and determine the proven acceptable ranges and normal operating ranges for the product in question. Some of the key decisions and factors include the batch size and facility design (e.g. conventional cleanroom vs. Restricted Access Barrier Systems, vs. isolator; single-use disposables versus stainless steel; (rotary) piston pumps vs. time-pressure fillers vs. peristaltic pumps; vial and stopper washing and depyrogenation/sterilization vs. ready-to-sterilize vs. ready-to-use components etc.). Products with multiple dose strengths will likely entail compounding during DP manufacture while others may use fully compounded / ready-to-fill DS. We provide here a few references that may be of value when performing a risk assessment of the unit operations to determine which process development studies to perform. Use of nested vials is becoming increasingly popular but sterilization by ethylene oxide can result in residues that can impact the quality of the protein (Chen et al. 2015). Different types of pumps can variously impact the level of aggregation and particulate formation with the peristaltic being the most gentle and convenient with disposable product contact parts (Nayak et al. 2011; Tyagi et al. 2009). Improved pump head design and drives have also improved the accuracy of fill by peristaltic pumps significantly (Lambert 2008). Isolators are increasingly coming into use for aseptic processing, especially in new facilities. The interior of these isolators is commonly decontaminated / sterilized by vaporized hydrogen peroxide (Agalloco 2018; Hopkins 2018). Residues of vaporized hydrogen peroxide in isolators can result in oxidation of proteins if proper deaeration procedures and control limits are not established (Cheng et al. 2016; Wang et al. 2004). Concentration of key excipients has to be monitored during process design to ensure that the target levels are achieved. Preservatives and surfactants are adsorbed by filters and by flexible tubing (Bahal and Romansky 2001a; Bahal and Romansky 2001b; Saller et al. 2017) requiring proper discard and purge procedures. The silastic tubing used in peristaltic pumps can also shed particles depending on the brand/type as well as pumping characteristics (Saller et al. 2015).

Other Considerations

Biologics drug substances are generally stored in the frozen state to disconnect the use period of the drug substance from the shelf-life of the drug product (Singh and Nema 2010). The formulation of the drug substance is commonly based on the composition defined for the drug product with the potential difference being that the concentration of protein may be higher in the DS than in the DP. The formulation development of the DP thus should take the need for frozen state storage of DS in consideration. In the case of biosimilars, matching the formulation of the reference may or may not lead to the best composition for frozen state storage,

in light of current state of knowledge (e.g. the use of phosphate buffers, sorbitol, trehalose) (Singh and Nema 2010). However, deep freezing (below −40 °C) into the glassy state generally ameliorates most of these limitations by slowing down all reactions, thus allowing freedom to explore in the formulation space. Since there is no requirement to match the stability profiles of the drug substances, the most appropriate storage system (container, technology, temperature) should be selected for the biosimilar based upon operational requirements. Good reviews of the science and technology of frozen state storage of proteins have been published and are recommended by the authors (Singh and Nema 2010; Kolhe and Goswami 2018).

A related question arises on the handling and storage of reference product material. A good biosimilar program requires a good sourcing program for the reference drug product. Typically biosimilar development companies buy several lots of the reference product over several years, and generally from multiple markets (e.g. US, EU, Canada etc.). This is usually an expensive undertaking. Extensive analysis of the reference product is typically done for setting QTPP (goal posts) for biosimilar development, within the labelled shelf-life period and/or at expiration. However, there may arise a question about preserving a quantity of the innvoator drug product past its expiry for certain long-term goals, e.g., as an analytical reference while new methods are developed, or for some clinical pharmacological assay needs. Freezing of DP should be undertaken with caution. Most biologics DP state on the label “Do not freeze”. However, for developmental purposes, a freezing program can be undertaken. Freezing lyophilized products in the lyophilized form is generally risk-free. Freezing liquid products safely depends on a number of factors. The first most obvious one is the fill volume in the original container. High fill volumes carry the risk for vial or PFS breakage on freezing. The product solution may be dispensed into new low-fill containers to avoid this risk. The composition of the reference product may not be optimum for freezing (as discussed above), but deep freezing (below −40 °C) will generally work to keep the product stable regardless. Finally, the thawed product should be considered as suitable and representative for analysis of chemical modifications only, since the aggregation and particulate matter status of the product may be changed by the freezing and thawing itself. (Obviously, this does not hold for when the freeze/thaw cycling studies are performed as part of stress testing for formulation development).

Summary

The biosimilar drug product development strategy begins with an analysis of the intellectual property background and then proceeds in a systematic manner with the QTPP being defined by an extensive analysis of the reference product. While there are no requirements to match the formulation of the reference, the biosimilar product must have (some of) the same strength(s) and dosage form(s) as the reference. Greater differences are seen on the device front where the biosimilars do not always match the reference. This may be an issue for interchangeability. The drug product

manufacturing process for the biosimilar is designed independently and is intended to produce a quality and safe product.

Note: The US FDA has withdrawn the draft guidance on statistical approaches to evaluate similarity as of 21 June 2018 (U.S. FDA 2018b).

References

Abdul-Fattah AM, Oeschger R, Roehl H, Bauer Dauphin I, Worgull M, Kallmeyer G, Mahler HC. Investigating factors leading to fogging of glass vials in lyophilized drug products. Eur J Pharm Biopharm. 2013;85(2):314–26.

Agalloco J. Open letter in reponse to MHRA blog on Vapour Hydrogen Peroxide fragility. 2018. https://www.linkedin.com/pulse/open-letter-response-mhra-blog-vapour-hydrogen-jim-agalloco/. Accessed June 2018.

Bahal SM, Romansky JM. Sorption of parabens by fluoropolymer or fluoropolymer laminated flexible tubing. PDA J Pharm Sci Technol. 2001a;55(6):417–21.

Bahal SM, Romansky JM. Sorption of parabens by flexible tubings. Pharm Dev Technol. 2001b;6(3):431–40.

Carpenter JF, Randolph TW, Jiskoot W, Crommelin DJ, Middaugh CR, Winter G, Fan YX, Kirshner S, Verthelyi D, Kozlowski S, Clouse KA, Swann PG, Rosenberg A, Cherney B. Overlooking subvisible particles in therapeutic protein products: gaps that may compromise product quality. J Pharm Sci. 2009;98(4):1201–5.

Celltrion. CT-P13 (infliximab biosimilar). Briefing Document for the Arthritis Advisory Committee. 2016. https://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/arthritisadvisorycommittee/ucm484860.pdf. Accessed June 2018.

Chen L, Sloey C, Zhang Z, Bondarenko PV, Kim H, Ren D, Kanapuram S. Chemical modifications of therapeutic proteins induced by residual ethylene oxide. J Pharm Sci. 2015;104(2):731–9.

Cheng W, Zheng X, Yang M. Hydrogen peroxide induced protein oxidation during storage and lyophilization process. J Pharm Sci. 2016;105(6):1837–42.

Dimitrova MN, Bee JS, Lu L, Fernandez JF. Development of prefilled syringe combination products for biologics. In: Warne NW, Mahler H-C, editors. Challenges in protein product development. AAPS/Springer; 2018. p. 203–24.

Ditter D, Mahler HC, Gohlke L, Nieto A, Roehl H, Huwyler J, Wahl M, Allmendinger A. Impact of vial washing and depyrogenation on surface properties and delamination risk of glass vials. Pharm Res. 2018a;35(7):146.

Ditter D, Nieto A, Mahler HC, Roehl H, Wahl M, Huwyler J, Allmendinger A. Evaluation of glass delamination risk in pharmaceutical 10 mL/10R vials. J Pharm Sci. 2018b;107(2):624–37.

European Agency for Safety and Health at Work. Directive 2010/32/EU - Prevention from sharp injuries in the hospital and healthcare sector. 2010. https://osha.europa.eu/en/legislation/directives/council-directive-2010-32-eu-prevention-from-sharp-injuries-in-the-hospital-and-healthcare-sector. Accessed June 2018.

Hopkins A. VHP (vapour hydrogen peroxide) fragility. 2018. https://mhrainspectorate.blog.gov.uk/2018/04/20/vhp-vapour-hydrogen-peroxide-fragility/. Accessed June 2018.

Jahn M. Leachables and extractables: from regulatory expectations to laboratory assessment. In: Warne NW, Mahler HC, editors. Challenges in protein product development. AAPS/Springer; 2018. p. 337–51.

Kishore RS, Kiese S, Fischer S, Pappenberger A, Grauschopf U, Mahler HC. The degradation of polysorbates 20 and 80 and its potential impact on the stability of biotherapeutics. Pharm Res. 2011a;28(5):1194–210.

Kishore RS, Pappenberger A, Dauphin IB, Ross A, Buergi B, Staempfli A, Mahler HC. Degradation of polysorbates 20 and 80: studies on thermal autoxidation and hydrolysis. J Pharm Sci. 2011b;100(2):721–31.

Kolhe P, Goswami S. Bulk protein solution: freeze-thaw process, storage and shipping considerations. In: Warne NW, Mahler HC, editors. Challenges in protein product development. AAPS/Springer; 2018. p. 313–36.

Lamanna WC, Holzmann J, Cohen HP, Guo X, Schweigler M, Stangler T, Seidl A, Schiestl M. Maintaining consistent quality and clinical performance of biopharmaceuticals. Expert Opin Biol Ther. 2018;18(4):369–79.

Lambert P. Dispensing biopharmaceuticals with piston and peristaltic pumps. PharmTech; 2008.

Liu W, Swift R, Torraca G, Nashed-Samuel Y, Wen ZQ, Jiang Y, Vance A, Mire-Sluis A, Freund E, Davis J, Narhi L. Root cause analysis of tungsten-induced protein aggregation in pre-filled syringes. PDA J Pharm Sci Technol. 2010;64(1):11–9.

Lolas AG, Metcalfe JW. Evaluation of the microbial growth potential of pharmaceutical drug products and quality by design. PDA J Pharm Sci Technol. 2011;65(1):63–70.

Manning MC, Chou DK, Murphy BM, Payne RW, Katayama DS. Stability of protein pharmaceuticals: an update. Pharm Res. 2010;27(4):544–75.

Metcalfe JW. Microbiological quality of drug products after penetration of the container system for dose preparation prior to patient administration. Am Pharm Rev. 2009;12(1, Jan/Feb):84–9.

Nayak A, Colandene J, Bradford V, Perkins M. Characterization of subvisible particle formation during the filling pump operation of a monoclonal antibody solution. J Pharm Sci. 2011;100(10):4198–204.

Piedmonte DM, Summers C, McAuley A, Karamujic L, Ratnaswamy G. Sorbitol crystallization can lead to protein aggregation in frozen protein formulations. Pharm Res. 2007;24(1):136–46.

Piedmonte DM, Hair A, Baker P, Brych L, Nagapudi K, Lin H, Cao W, Hershenson S, Ratnaswamy G. Sorbitol crystallization-induced aggregation in frozen mAb formulations. J Pharm Sci. 2015;104(2):686–97.

Ricci MS, Frazier M, Moore J, Cromwell M, Galush WJ, Patel AR, Adler M, Altenburger U, Grauschopf U, Goldbach P, Fast JL, Kramer I, Mahler HC. In-use physicochemical and microbiological stability of biological parenteral products. Am J Health Syst Pharm. 2015;72(5):396–407.

Rosenberg AS. Effects of protein aggregates: an immunologic perspective. AAPS J. 2006;8(3):E501–7.

Rubic-Schneider T, Kuwana M, Christen B, Assenmacher M, Hainzl O, Zimmermann F, Fischer R, Koppenburg V, Chibout SD, Wright TM, Seidl A, Kammuller M. T-cell assays confirm immunogenicity of tungsten-induced erythropoietin aggregates associated with pure red cell aplasia. Blood Adv. 2017;1(6):367–79.

Saller V, Matilainen J, Grauschopf U, Bechtold-Peters K, Mahler HC, Friess W. Particle shedding from peristaltic pump tubing in biopharmaceutical drug product manufacturing. J Pharm Sci. 2015;104(4):1440–50.

Saller V, Matilainen J, Rothkopf C, Serafin D, Bechtold-Peters K, Mahler HC, Friess W. Preservative loss from silicone tubing during filling processes. Eur J Pharm Biopharm. 2017;112:109–18.

Sandoz. FDA Oncologic Drugs Advisory Committee Meeting. Zarxio (filgrastim). Advisory Committee Briefing Materials. 2015. https://wayback.archive-it.org/7993/20170405222933/https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM428782.pdf. Accessed June 2018.

Schiestl M, Stangler T, Torella C, Cepeljnik T, Toll H, Grau R. Acceptable changes in quality attributes of glycosylated biopharmaceuticals. Nat Biotechnol. 2011;29(4):310–2.

Scott Morton F, Stern AD, Stern S. The impact of entry of biosimilars: evidence from Europe; Harvard Business School Working Paper. No. 16-141: June 2016. Revised April 2018. https://www.hbs.edu/faculty/Pages/item.aspx?num=51309. Accessed June 2018

Seidl A, Hainzl O, Richter M, Fischer R, Bohm S, Deutel B, Hartinger M, Windisch J, Casadevall N, London GM, Macdougall I. Tungsten-induced denaturation and aggregation of epoetin alfa during primary packaging as a cause of immunogenicity. Pharm Res. 2012;29(6):1454–67.

Singh SK. Storage considerations as part of the formulation development program for biologics. Am Pharm Rev. 2007;10(2, Mar/Apr):26–33.

Singh SK. Impact of product-related factors on immunogenicity of biotherapeutics. J Pharm Sci. 2011;100(2):354–87.

Singh SK. Sucrose and trehalose in therapeutic protein formulations. In: Warne NW, Mahler HC, editors. Challenges in protein product development. AAPS/Springer; 2018. p. 63–95.

Singh SK, Mahler H-C. Current trends in control and monitoring of subvisible particulate matter in biotherapeutics. BioPharma Asia. 2018;7(1, Jan/Feb):30–7.

Singh SK, Nema S. Freezing and thawing of protein solutions. In: Jameel F, Hershenson S, editors. Formulation and process development strategies for manufacturing biopharmaceuticals. Hoboken: Wiley; 2010. p. 625–75.

Singh SK, Afonina N, Awwad M, Bechtold-Peters K, Blue JT, Chou D, Cromwell M, Krause HJ, Mahler HC, Meyer BK, Narhi L, Nesta DP, Spitznagel T. An industry perspective on the monitoring of subvisible particles as a quality attribute for protein therapeutics. J Pharm Sci. 2010;99(8):3302–21.

Thakrar K, Bodalia P, Grosso A. Assessing the efficacy and safety of omnitrope. Br J Clin Pharm. 2010;2:298–301.

Tyagi AK, Randolph TW, Dong A, Maloney KM, Hitscherich C Jr, Carpenter JF. IgG particle formation during filling pump operation: a case study of heterogeneous nucleation on stainless steel nanoparticles. J Pharm Sci. 2009;98(1):94–104.

U.S. Department of Labor. Occupational exposure to bloodborne pathogens: Needlestick and other sharps injuries. Final rule. OSHA Federal Register Final Rules 2011;66:5317–5325. https://www.osha.gov/SLTC/bloodbornepathogens/recognition.html. Accessed June 2018.

U.S. FDA. FDA Briefing Document. Oncologic Drugs Advisory Committee Meeting. BLA 125553. EP2006, a proposed biosimilar to Neupogen (filgrastim). Sandoz Inc.; 2015a. http://patentdocs.typepad.com/files/briefing-document.pdf. Accessed June 2018.

U.S. FDA. Scientific Considerations in Demonstrating Biosimilarity to Reference Product. Guidance for Industry. 2015b. https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf. Accessed June 2018.

U.S. FDA. CDER. Application Number 125533Orig1s000. Other Review(s). Zarxio (filgrastim-sndz). 2015c. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/125553Orig1s000OtherR.pdf. Accessed June 2018.

U.S. FDA. BLA 761024. Approval Letter Amjevita (adalimumab-otto). 2016a. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2016/761024Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. BLA 125544. Approval Letter Inflectra (infliximab-dyyb). 2016b. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2016/125544Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. BLA 761042. Approval Letter Erelzi (etanercept-szzs). 2016c. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2016/761042Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. CDER Application Number 761024Orig1s000. Chemistry Reviews ABP501 (adalimumab-otto). 2016d. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761024Orig1s000ChemR.pdf. Accessed June 2018.

U.S. FDA. CDER Application Number 125544Orig1s000. Chemistry Reviews CT-P13 (infliximab-dyyb). 2016e. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/125544Orig1s000ChemR.pdf. Accessed June 2018.

U.S. FDA. CDER. Application Number 761042Orig1s000. Other Reviews. Erelzi (etanercept-szzs). 2016f. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761042Orig1s000OtherR.pdf. Accessed June 2018.

U.S. FDA. CDER. Application Number 761024Orig1s000. Other Reviews. ABP501 (adalimumab-otto). 2016g. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761024Orig1s000OtherR.pdf. Accessed June 2018.

U.S. FDA. Considerations in demonstrating interchangeability with a reference product. Guidance for Industry. Draft Guidance. 2017a. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf. Accessed June 2018.

U.S. FDA. Guidance for Industry and FDA Staff: Current Good Manufacturing Practice Requirements for Combination Products. 2017b. https://www.fda.gov/downloads/regulatoryinformation/guidances/ucm429304.pdf. Accessed June 2018.

U.S. FDA. FDA briefing document. Oncologic Drugs Advisory Committee Meeting. BLA 761074. MYL-1401O, a proposed biosimilar to Herceptin (trastuzumab). Mylan Pharmaceuticals. 2017c. https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM566369.pdf. Accessed June 2018.

U.S. FDA. FDA Briefing Document. Oncologics Drug Avisory Commitee Meeting. BLA 125545. "Epoein Hospira", a proposed biosimilar to Epogen/Procrit (epoetin alfa). Hospira Inc.; 2017d. https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM559967.pdf. Accessed June 2018.

U.S. FDA. BLA 760158. Approval Letter Cyltezo (adalimumab-adbm). 2017e. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2017/761058Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. BLA 761054. Approval Letter Renflixis (infliximab-abda). 2017f. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2017/761054Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. BLA 761072. Approval Letter Ixifi (infliximab-qbtx). 2017g. https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2017/761072Orig1s000ltr.pdf. Accessed June 2018.

U.S. FDA. Search orphan drug designations and approvals. 2018a. https://www.accessdata.fda.gov/scripts/opdlisting/oopd/. Accessed June 2018.

U.S. FDA. FDA withdraws draft guidance for industry: Statistical approaches to evaluate analytical similarity. 2018b. https://www.fda.gov/Drugs/DrugSafety/ucm611398.htm. Accessed July 2018.

Vandekerckhove K, Seidl A, Gutka H, Kumar M, Gratzl G, Keire D, Coffey T, Kuehne H. Rational selection, criticality assessment, and tiering of quality attributes and test methods for analytical similarity evaluation of biosimilars. AAPS J. 2018;20(4):68.

Wang W, Cui TY, Wang YJ, Martin-Moe S. Oxidation of protein by vaporized sanitizing agents. PDA J Pharm Sci Technol. 2004;58(3):121–9.

Part IV
Analytical Similarity Considerations for Biosimilars (Protein characterization, Biophysical Tools, Biological Assays, Statistical Methods)

Chapter 11
Principles of Analytical Similarity Assessment

Kristof Vandekerckhove and Russell Reeve

Abstract Detailed evaluation of the similarity in structural and functional properties between a proposed biosimilar product and the reference product necessitates a carefully designed analytical study program. Although regulatory agencies, such as EMA and US FDA, have published guidance documents outlining the requirements for analytical similarity assessment, the implications are often not fully understood by biosimilar developers. This chapter discusses important considerations for all aspects of the design of the analytical similarity assessment, including the selection of the test materials; the product characteristics to be compared, and the associated analytical testing methods and plan; suitability of analytical procedures, in design and performance; processing and interpretation of analytical test data; and the methods for assessment of analytical study results. The design of comparative forced degradation studies, intended to compare the pattern and kinetics of product degradation, is also discussed. Differences in regulatory expectations between EU and USA are identified, together with their implications for the conception of a study program intended to support product approval in both jurisdictions.

Keywords Analytical similarity · Analytical methods · Biosimilar · Quality attributes · Forced degradation · Tiering · Criticality assessment · Process shift · Design of experiments · Statistics

Introduction

Analytical assessment of similarity represents the first step in the sequence of pivotal studies for a proposed biosimilar product. The current state of the art of analytical technology allows sponsors to elucidate the structure and biological

K. Vandekerckhove (✉)
Strategic Drug Development, IQVIA, Vilvoorde, Belgium
e-mail: kvdkerckhove@telenet.be

R. Reeve
Advisory Analytics, IQVIA, Durham, NC, USA

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_11

function of therapeutic proteins in great detail, thereby offering a high probability of detecting existing differences between two biological products. For this reason, both the European Medicines Agency (EMA) (EMA 2014a) and the United States Food and Drug Administration (U.S. FDA) (U.S. FDA 2015a, b) consider the analytical comparison to be the foundation of biosimilar development, with further studies primarily aimed at determining the clinical relevance of observed analytical differences. *In vivo* studies also serve to provide additional reassurance of similarity in therapeutic performance, as not all relevant differences may be detected in the analytical study program.

The extent of analytical characterization can have an impact on subsequent product development: in their “Guidance for Industry: Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product” (U.S. FDA 2016), U.S. FDA discusses an explicit link between the degree of analytical similarity that is achieved, and the extent of further development required to address residual uncertainty in the determination of similarity. Similarly, EMA states in their “Guideline on similar biological medicinal products” (EMA 2014a) that *“the extent and nature of the non-clinical in vivo studies and clinical studies to be performed depend on the level of evidence obtained in the previous step(s) including the robustness of the physicochemical, biological and non-clinical in vitro data”*.

This preamble demonstrates that careful design of the analytical similarity study program is an essential condition for a successful biosimilar development program. All constituents of the analytical study protocols should be meticulously developed in order to meet the ultimate objective of analytical similarity assessment: the detection of all relevant differences between the proposed biosimilar and the chosen reference product.

Choice of Test Materials

General Considerations

Introduction

Biological medicines, unlike their synthetic counterparts, are a heterogeneous mixture of structural variations of a polypeptide and other (desired or undesired) substances, which is the unavoidable result of their production in living organisms. The composition of this mixture varies between individual manufacturing runs, whereby the level of variation is maintained within acceptable limits by appropriate controls of critical process parameters. The manufacturing process does not remain static over time: process changes are a common feature in the pharmaceutical industry, either triggered by necessity (e.g. a new cell bank, cessation of supply of a process material), business motives (e.g. a change in manufacturer), or the desire to further optimize the process (e.g. improved process yield, process simplification). These process changes further augment the variability between production lots.

Even in the event that no changes are made to the process over time, individual product attributes may drift over time. Intense process monitoring and mandatory Annual Product Quality Reviews serve to detect such process drift early-on, but are restricted to predefined (analytical) checks. Post-production, further changes to the chemical composition of the product are possible due to a wide array of possible degradation mechanisms, which may be further complicated by reactions between the formed degradation products.

On top of the variability between drug substance lots caused by the various phenomena described above, incremental variability is introduced by the drug product manufacturing process. If different presentations of the product are licensed, such as different strengths, different container closure systems, or different fill volumes, these may also contribute to the overall variability of the product, due to variability in the chemical micro-environment and exposure to physical stress.

The EU guideline on quality issues for biosimilar products specifies that the objective of the analytical study is to demonstrate similarity of the proposed biosimilar product and the reference product at the level of the finished product (EMA 2014b). The same preference is also found in the applicable U.S. FDA guidance document, stating that the finished product should be analyzed whenever possible (U.S. FDA 2015b). This seemingly straightforward expectation is far more intricate as it may appear at first glance: the finished product may be a collection of different licensed presentations, each of which in turn represents a population of variations in physicochemical and biological properties (for the reasons discussed above). A meaningful comparison of the proposed biosimilar and reference product therefore requires appropriate selection of the test materials, covering the degree of variation in each product with reasonable certainty.

Choice of Presentations

In case different presentations are licensed for the reference product, a sponsor may decide to develop all or only a subset of these presentations for the proposed biosimilar. Such a decision is usually made early in development, and must in any case be final before initiation of the pivotal analytical similarity studies. If only a limited number of presentations are proposed, the most conservative approach is to include all presentations in the analytical study program. When the number of presentations is too large, rational selection of a subset of all proposed strengths for analytical similarity assessment becomes necessary.

A first step in the selection of presentations from the range of proposed biosimilar presentations constitutes of mapping the variables that differ between the presentations. Typical variables are container closure system, fill volume, strength (expressed as protein content or biological activity units), composition, and pharmaceutical form. Selection can then be based on bracketing, as described in ICH guideline Q1D (ICH 2002), meaning that those presentations representing the extremes of each drug product presentation factor are selected for testing.

Drug product presentation selection may however be complicated by several factors. The same presentations may not be licensed in each jurisdiction of interest. Designing a single analytical study program satisfying the data requirements in each region of interest requires the choice of product presentations licensed in all targeted regions, which may be more or less incompatible with the bracketing-based selection. Technical limitations must also be considered: certain formulation factors, such as low protein strength or product composition of specific presentations may interfere with the intended analyses, and justify different choices. Finally, regulatory agencies advise to test samples of the actual batches used in clinical studies in the analytical study program; the presentation(s) selected for clinical evaluation can differ from the 'worst-case' selection discussed above.

It may also be wise to consider other variables for the selection of presentations, particularly if the same selection is also proposed for a reduced design of other product development studies, such as process validation and stability studies. Variables of interest are those (process-related) variables that are fixed within a presentation, but different between presentations. A typical example is batch size, particularly if the commercial demand is known to vary between presentations.

Table 11.1 provides an example of selection of appropriate presentations for analytical similarity studies for a proposed biosimilar of Aranesp®.

Testing of less than the complete set of presentations that is proposed for registration inevitably leaves some degree of uncertainty. This is particularly true for the reference product: testing the same presentations that were selected for the proposed biosimilar assumes that all relevant variables are identical between the products. The validity of this hypothesis cannot be verified, given the inaccessibility of reference product manufacturing details. Sponsors must therefore be vigilant for any signals of qualitative or quantitative differences between presentations, and consider the need of testing additional presentations if such differences are suspected.

Product Age

The quality of therapeutic proteins tends to deteriorate with age, due to a highly complex and multifactorial degradation process. The selected product lots of the proposed biosimilar and reference product must therefore ideally be of comparable age. Meeting this condition requires careful planning: in conventional development practice, the lots of the product under development are often tested closely after production. In contrast, the lots of reference product are usually procured from the open market, and have a certain age as a result of the different steps that must be completed before entering the commercial distribution chain. Without specific attention to product age, analytical similarity studies may compare relatively fresh lots of the proposed biosimilar versus reference product lots of older age. The relevance of such comparison is questionable, in particular for quality attributes that are sensitive to product degradation.

Availability of proposed biosimilar lots of ages spanning the proposed product shelf life can be difficult to achieve in practice. The most important constraint is

Table 11.1 Selection of presentations for biosimilar of Aranesp (case example)

Extremes: 10 µg, 500 µg	*Extremes: 0.3 mL, 1 mL*	*Extremes: 25 µg/mL, 500 µg/mL*	*Extremes: 2,500 units, 8,000 units*	*Sponsor will only develop PFS*	*Prefilled pens will not be developed*	*Vials will not be developed*	*Licensed presentation*	*Licensed presentation*
Dose (µg)	**Volume (mL)**	**Concentration (µg/mL)**	**Batch size**	**CCS[a] 1**	**CCS[a] 2 (EU only)**	**CCS[a] 3**	**EU[c]**	**USA[d]**
10	0.4	**25**	5,000	**PFS**	PF Pen	-	**x**	**x**
15	0.375	40	**2,500**	**PFS**	PF Pen	-	x	-
20	0.5	40	7,000	**PFS**	PF Pen		x	-
25	1	25		-	-	Vial	x	x
25	0.42	*60*[b]	4,000	**PFS**	-	-	-	x
30	**0.3**	100	7,000	**PFS**	PF Pen	-	x	-
40	0.4	100	**8,000**	**PFS**	PF Pen	-	**x**	**x**
40	1	40		-	-	Vial	x	x
50	0.5	100	5,000	**PFS**	PF Pen	-	x	-
60	**0.3**	200	**8,000**	**PFS**	PF Pen	-	**x**	**x**
60	1	60		-	-	Vial	x	x
80	0.4	200	5,000	**PFS**	PF Pen	-	x	-
100	0.5	200	5,000	**PFS**	PF Pen	-	**x**	**x**
100	1	100		-	-	Vial	x	x
130	0.65	200	**2,500**	**PFS**	PF Pen	-	x	-
150	**0.3**	**500**	5,000	**PFS**	PF Pen	-	**x**	**x**
150	0.75	200		-	-	Vial	-	x
200	1	200		-	-	Vial	x	x
200	0.4	**500**	**2,500**	**PFS**	-	-	-	x
300	0.6	**500**	**2,500**	**PFS**	PF Pen	-	**x**	**x**
300	1	300		-	-	Vial	x	x
500	**1**	**500**	**2,500**	**PFS**	PF Pen	-	**x**	**x**
500	1	500		-	-	Vial	-	x

Red bolded = value of interest for selection. Red-dotted boxes: selected presentations for analytical similarity, stability, drug product process validation

[a]*CCS* container closure system

[b]The US label does not mention the product concentration. For the 25 µg PFS, a volume of 0.42 mL is mentioned in the label. By calculation, this entails a protein concentration of 59.524 µg/mL. It is assumed that a protein concentration of 60 µg/mL is used

[c]EMA 2018

[d]U.S. FDA 2018

time: successful biosimilar development is time-sensitive, therefore it is difficult to financially justify the spread of biosimilar lot production across several years. In addition, process changes are common during product development, whereas all lots used in the analytical similarity study program should ideally originate from the final manufacturing process (see section "Biosimilar Product" for further discussion). The most common resolution of this problem is by adoption of a staged analytical similarity assessment program: before the start of clinical development, the proposed biosimilar and the reference product are analytically compared in a first characterization study. At this time, only a limited number of lots of the proposed biosimilar are normally available, at the start of their shelf life. The results of this first study must provide sufficient evidence to justify the start of clinical studies, but are inadequate to support product registration as a biosimilar. A second and conclusive (pivotal) evaluation of analytical similarity is planned closer to the time of registration, and includes lots manufactured at different times. The time required for the conduct of the clinical studies is then used for aging of product batches, stored under the recommended conditions. This approach is not in strict compliance with the recommendations of EMA (EMA 2014a) and U.S. FDA (U.S. FDA 2015a, 2016), favoring a step-wise development approach that starts with comprehensive analytical characterization (meaning: all potentially relevant analytical differences must be identified before proceeding with in vivo studies). Furthermore, the sponsor must be willing to accept a degree of developmental risk, as previously undetected analytical differences may be revealed during the late-stage comprehensive analytical studies. The risk that is specifically related to lack of knowledge of product aging in the early analytical study can be mitigated by conducting comparative forced degradation studies (see section "Forced Degradation" for further discussion).

Analytical Feasibility

Regulatory Agencies recommend analysis of the finished product, as this is the material that is actually administered to patients (see section "Introduction"). The finished product may however not be the optimal test matrix for similarity assessment, particularly if the composition interferes with the analysis; examples of interfering factors are product concentration, excipients and container closure system-derived substances (e.g. leachables, silicone oil in prefilled syringes). In such instances, an appropriate sample pre-treatment protocol must be developed that addresses the (suspected or known) analytical interference. Generally, this means that samples must be purified and/or concentrated, using a technique that takes due account of the properties of the product, the target analysis, and the objective of sample pre-treatment. Any manipulation of the sample bears the risk of artificial changes to its properties and thus a biased test result. For that reason, the effect of the sample treatment procedure must be experimentally evaluated. This can be most easily done by comparing the test result of a sample that does not contain the interfering agent (usually obtained by taking an in-process sample of the biosimilar drug substance upstream of the operation causing the analytical interference, further

processed by an interference-free small scale version of the process if required) versus the result from a drug product sample treated by the proposed procedure. If the results reveal a difference between the samples, the sample treatment procedure must be further optimized or alternative options explored, including consideration of alternative analytical tools not affected by the interference.

Statistical Considerations

The analytical studies serve to determine the similarity of the proposed biosimilar and the reference product, which, as explained before (section "Introduction"), are each the end-result of a large number of input factors with their associated variability. A meaningful comparison of the two populations requires the use of inferential statistical methodology, where possible; this is discussed in further detail in section "Assessment of Analytical Similarity Data". The use of such methodology, in turn, has implications for the selection of the test materials. First, the unit of observation (or: statistical unit) must be defined, which is defined as the entity on which information is collected and is used as the basis for statistical analysis (Eurostat). The choice should in theory be based on a thorough understanding of the different sources of variation within the population, and must support a reliable estimate of population distribution parameters to be used for statistical comparison. Intuitively, the unit of observation is the unit that is sampled from a population, whereby it is the objective to make inferences about the population of those units. In other words, a collection of observation units will be analyzed, in order to infer something about the larger population of those same units of observation.

In practice, a biosimilar developer has only incomplete knowledge of the sources of variability, in particular for the reference product, and is furthermore limited in his choice of the unit of observation by the information that is available on reference product samples. For that reason, the conventional unit of observation for similarity studies is the production batch. Note that subsampling may be of interest in certain situations, for example in the case of quality attributes that are determined by analytical methodology associated with important measurement error (e.g. most biological activity assays): in such cases, the lot mean (which is the statistical unit) can be computed from the different subsample assay values.

Reliable estimation of population distribution parameters, as is required for inferential statistical analysis, is only possible if the collected samples are representative for the underlying data generating processes (EMA 2017a). This is best achieved by random sampling, but such a strategy is difficult or impossible to implement in the context of similarity studies: only a limited number of batches of the proposed biosimilar are usually available, produced consecutively. Similar practical obstacles, such as a restriction in sources and timing for procurement, might cause non-random sampling of reference product batches. Instead, measures must be taken to purposefully select samples that are believed to be representative for the process. Relevant sampling considerations are discussed in sections "Reference Product" and "Biosimilar Product".

Reference Product

Besides the general considerations that are discussed in section "General Considerations", there are a number of additional criteria that apply specifically to the selection of the reference product. A first important consideration stems from legislative provisions in the different countries that are targeted for future registration. EU Directive 2001/83/EC (as amended) (European Parliament and Council 2001) defines 'reference medicinal product' as a medicinal product authorized in the EEA on the basis of a complete dossier. In practice, this means that the reference product must be released for commercial use in the EEA. Similarly, U.S. legislation requires the use of a single reference product that previously has been licensed by FDA (U.S. FDA 2015a). In both jurisdictions, regulatory guidance documents clarify that the use of a comparator product from a different source may be accepted, provided a number of conditions are fulfilled (EMA 2014a; U.S. FDA 2015a). Proof of analytical similarity between the domestic reference product and the proposed foreign comparator is one critical condition, meaning that there is no regulatory relief for multiple sources of reference product in the analytical similarity study program. Other countries either have a similar requirement, or accept the use of a non-domestic reference product, whether or not subject to specific conditions (e.g. the product must be sourced from an approved reference country). A detailed understanding of the regulatory requirements, and possible exceptions, in all markets of interest is therefore essential to enable informed decisions on the sourcing of the reference product samples.

Analysis of the reference product samples must provide an accurate representation of the range of variability that typifies the reference product manufacturing process. The range of results measured for the reference product forms the primary basis for analytical similarity evaluation, therefore each lot of reference product must ideally originate from a different batch of drug substance. If different reference product samples originate from the same lot of drug substance, their results will not reflect the variability between drug substance production runs, and thus cause underestimation of the true range of variability for the reference product. Such correlation between reference product lots is however unknown to sponsors of biosimilar products, and can only be suspected. One obvious factor of risk for sourcing of correlated lots, is to procure lots which have the same or a very similar expiry date. Depending on the batch size of the drug substance and drug product, one batch of bulk drug substance may be used to manufacture several batches of drug product, usually in a consecutive fashion – although not necessarily in a single manufacturing campaign. Spacing reference product procurement over time helps to reduce the risk of lot correlation, but does not eliminate it. Use of isotope ratio monitoring mass spectrometry has been reported to enable identification of the origin of protein expression (Apostol et al. 2001) and may be an analytical option to determine lot correlation. It must however be noted that the drug substance source, although arguably the most important correlation factor, is not the only factor of correlation between reference product lots. Other variables, such as excipient sources, or campaigning of reference drug product manufacture, also contribute to

lot correlation and therefore have the potential to cause underappreciation of the true variability of the reference product. Spaced purchase of reference product introduces other risks to the program: several examples of sudden shifts in the quality profile of biological medicines are reported in the public domain (Schiestl et al. 2011; Kim et al. 2017). Product samples sourced before and after such a shift then originate from different populations, complicating statistical processing of the pool of pre- and post-shift results by standard inferential methods.

The problems with the statistical analysis are both technical and inferential. Shifts in the process create a population of data that is no longer normally distributed, but instead may be multimodal, or at least flat-topped (platykurtic distribution). The shift needs to be around 3 standard deviations before it has a noticeable effect on the shape, where you start to see a small side mode, but in case of shifts greater than 3 standard deviations the composite has a distinctly bimodal shape (see Fig. 11.1).

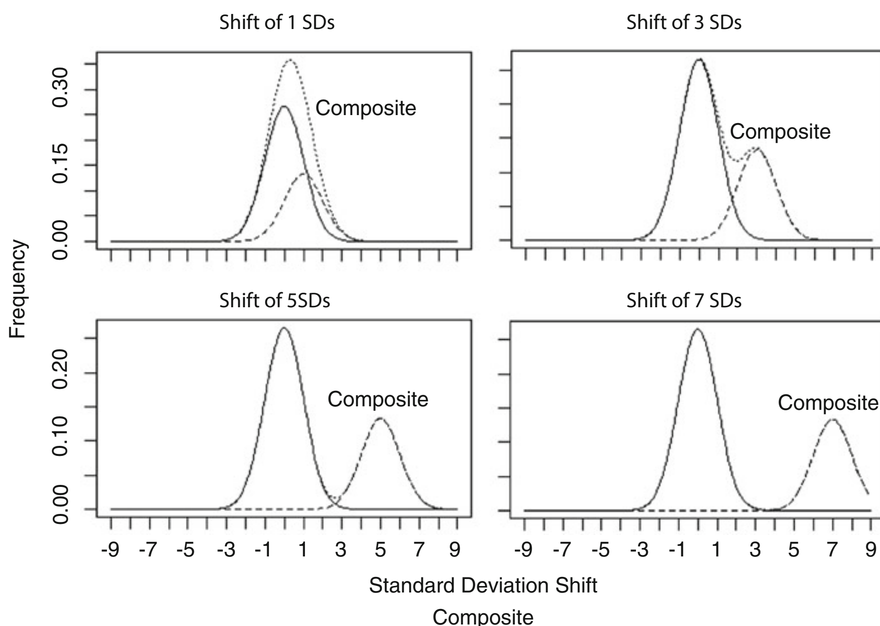


Fig. 11.1 Composite distributions caused by a process shift (pre-change process: solid line; post-change process: dashed line; composite distribution: dotted line)

In case of a shift occurring in the reference process, the underlying variability will be inflated. The inflated variability will manifest itself in two ways in the calculation of the statistical power of the analytical similarity study: the standard deviation of the measurements will increase, and the variability in the mean will increase. The net effect of both phenomena on the power of the study can be effectively investigated by Monte Carlo simulation. The simulation model employed in this exercise assumes that the biosimilar product has been engineered to match the original (pre-shift) manufacturing process of the reference product, to within $\pm\sigma/8$;

the deviation of the biosimilar process from the reference process will be referred to as the biosimilar bias, and is the same as the difference that is assumed by U.S. FDA in their simulation model (see section "U.S. FDA"). During the development of the biosimilar product, a shift occurs in the reference product manufacturing process, of a magnitude $r\sigma$ for some given r. It is important to note that this shift may be in the direction of the biosimilar bias, or away from it, and the directional consonance or dissonance will have an effect on the operating characteristics of the testing procedure. The simulation algorithm calculates reference lot data for each simulated experiment: five pre-change lots with measured attribute values of $0 + e$, where $e \sim N(0, \sigma^2)$, and five post-change lots with measured attribute values of $r\sigma + e$, where $e \sim N(0, \sigma^2)$. For each shift constant r 100,000 *in silico* trials were conducted. Individual values of biosimilar lots ($n = 10$ per simulated experiment) are calculated as $U + e$, where $e \sim N(0, \sigma^2)$. The values of U used in the simulations are $-\sigma/8$, 0 and $+\sigma/8$.

The effect of the shift in the reference product manufacturing process was determined for both Tier 1 and Tier 2 assessments,[1] by changing the decision rule:

- For Tier 1, the 95% confidence interval of the difference between simulated biosimilar and reference product lots is computed. If the confidence interval wholly lies between $\pm 1.5\ s_{reference}$, then the test passes. Power is calculated as the proportion of simulated experiments passing the equivalence test.
- For Tier 2, the test passes if 90% of the simulated biosimilar lots fall within the interval $[mean_{reference} \pm 2.5\ s_{reference}]$. Power is equally calculated as the proportion of simulated experiments passing the Quality Range test. A multiplier of 2.5 was selected as a compromise value between the conventional multiplier values of 2 to 3.

All reported results are precise to within 0.3% with 95% confidence.

$Mean_{reference}$ and $s_{reference}$ are the sample mean and sample standard deviation calculated from the simulated reference product lots for each individual experiment. The rules for assignment of analytical test measures to Tier 1 and Tier 2 and the associated statistical methodology is discussed in further detail in section "U.S. FDA".

A shift in a highly critical quality attribute (such as target biological activity) is unlikely to occur, as this would not pass regulatory approval. Therefore, the effect of a shift in the manufacturing process of the reference product on the power of Tier 1 assessments has no practical relevance, but was computed for completeness.

In case of Tier 1, as the size of the shift increases, the power drops precipitously when the shift varies from 0 (no shift) to 6 standard deviations (Fig. 11.2). The opposite trend in statistical power is seen for Tier 2: as the magnitude of the shift increases, the power of the study nears and ultimately reaches 100% (Fig. 11.3).

[1] The simulations adopted U.S. FDA's recommended approach to evaluation of analytical similarity data in effect at the time of writing of this chapter. New draft guidance on this subject was however announced by the Agency on the 21st of June 2018. Details of possible revisions to the methodology were not disclosed in the announcement.

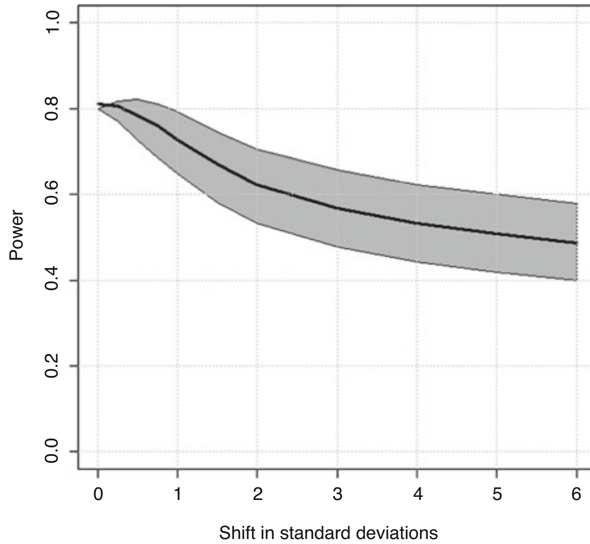


Fig. 11.2 Plot of power of Tier 1 equivalence test vs. size of shift in the reference product process (expressed as number of standard deviations) (thick solid line: biosimilar mean = pre-change reference product mean; shaded region = power for biosimilar mean ranging between $\pm\sigma/8$ difference from pre-change reference product mean)

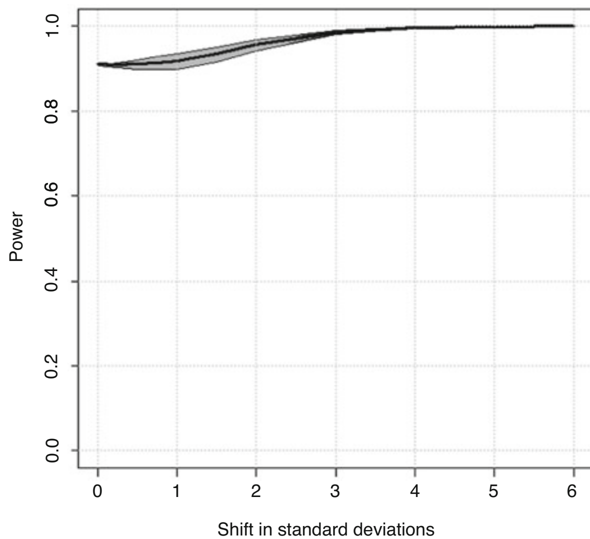


Fig. 11.3 Plot of power of Tier 2 quality range test vs. size of shift in the reference product process (expressed as number of standard deviations) (thick solid line: biosimilar mean = pre-change reference product mean; shaded region = power for biosimilar mean ranging between $\pm\sigma/8$ difference from pre-change reference product mean)

These results show that in case of Tier 1 assessment, the widening of the equivalence acceptance margin (caused by an increase in the calculated standard deviation for the reference product with an increasing size of the process shift) is counteracted by a simultaneous increase in the variability of the mean. For small shifts, the increase in the variability of the mean tends to dominate the power calculations, but for larger shifts, both opposing influences approximately cancel each other out, yielding the plateau. In case of evaluation per Tier 2 criteria, the shift in the process also translates into a widening of the similarity acceptance margin (here defined as a Quality Range), but is not counteracted by the increased variability of the mean: unlike Tier 1, the difference between the biosimilar and reference product (with the associated uncertainty, which increases with an increasing variability of the mean for the reference product) is not computed for Tier 2 evaluation.

The effect of a process shift on the Type I error rate is also different between Tier 1 and Tier 2. While an analysis based on a t-test is fairly robust against departures from normality, especially in the case of thin tails as in this case, there may be deviations of the true Type I error rate from the nominal error rate (Rhiel and Chaffin 1996), though this deviation is likely to be small, yielding Type I error rates of less than 6% when the desired rate is 5%; but this may cause some concerns about the validity of the test results among the statisticians reviewing the application. For Tier 2, Type I error can be estimated by simulations, by defining a 'true difference' that must be detected by the test, and computing the proportion of simulated experiments that passes the test. The true difference was defined as a difference of 2.5 standard deviations between the biosimilar process mean, and the mean of the pre-change reference product process. The change in the probability of passing Tier 2 criteria with an increasing process shift depends on the direction of the shift, as expected (Fig. 11.4): when the reference process shifts in the direction of the biosimilar process, a fast increase in the proportion of experiments passing Tier 2 criteria is observed. This is explained by both a shift of the overall mean for the simulated reference product lots, bringing it closer to the biosimilar mean, and a widening of the Quality Range due to an increase in the calculated standard deviation. In the opposite case, when the reference process shifts further away from the biosimilar process, the probability of (falsely) reaching a conclusion of similarity slightly decreases in case of small shifts, as the shift in the mean slightly overcompensates the widening of the Quality Range. For larger shifts, the widened Quality Range dominates and the Type I error increases.

From a practical standpoint, the impact of shifts in the manufacturing process of the reference product is only relevant for Tier 2 assessments. Two situations can be distinguished:

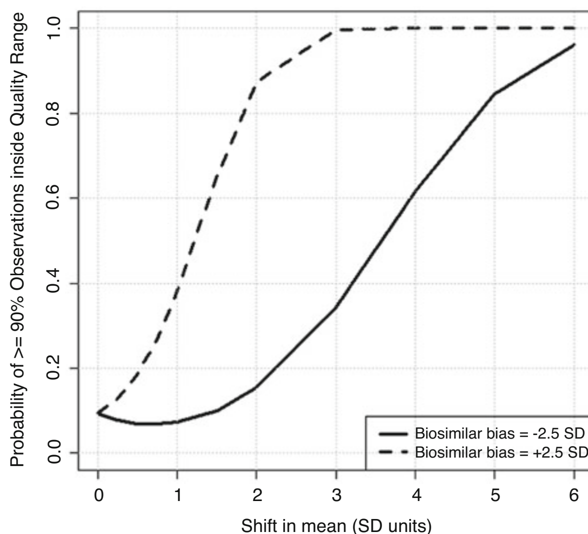


Fig. 11.4 Effect of a shift in the reference process (0 to +6 standard deviations) on probability of passing Tier 2 criteria, for biosimilar bias of −2.5 and +2.5 standard deviations

1. Smaller shifts (below three standard deviations) are difficult to detect for the biosimilar product sponsor, particularly if only a limited number of reference product batches were analyzed. Undetected shifts do not increase the patient risk for shifts less than 1.5 standard deviations, but do have an adverse effect on Type I error rate when exceeding a shift of 1.5 standard deviations.
2. Larger shifts (three standard deviations or greater) increase the Type I error rate, but the probability of detecting the process shift also increases, allowing the sponsor to take appropriate measures.

Regardless of the size of the process shift, the power of the study either remains unchanged or increases, hence decreasing the sponsor risk. The detection of true process shifts is therefore particularly important for protection of the patient, and may form an argument for complementary visual assessment of data patterns instead of sole reliance on mathematical decision rules.

Selection of representative samples of reference product also entails that these samples are within their shelf life at the time of analysis, and have been handled in accordance with label instructions at all times. Samples that have suffered excursions in storage temperature, or other incidents that may compromise the quality of the product, must not be used for analytical testing unless duly justified.

Biosimilar Product

The selection of test samples of the proposed biosimilar must be appropriately representative of the product to ensure a meaningful similarity assessment. Adherence to the general criteria discussed in section "General Considerations" will help to choose the appropriate presentations, production stage (generally drug product) and lot ages—whilst ensuring to include all lots that were used in pivotal development studies (e.g. comparative in vivo studies). Specific attention should be dedicated to biosimilar batch origin and pedigree: batches must be representative of the product that is proposed for commercial manufacture, with minimal inter-lot correlation to avoid underestimation of true product variability.

The most obvious manner to ensure that the biosimilar test lots are representative of the proposed commercial product is to choose lots that originate from the manufacturing process that is proposed in the registration file. For the avoidance of ambiguity, this should be understood as lots that were manufactured using the final manufacturing procedure, at the intended scale, in the production facility that is proposed in the registration dossier, using the same equipment and facility controls, with all process input and auxiliary materials from the sources described in the dossier, and meeting the described quality standards—both for the drug product, and the drug substance. Since process and other changes are common during biological (including biosimilar) product development, it is very difficult to restrict the selection of lots to those originating from the final process. One of the most common changes during late-stage development of biological products is an increase of the manufacturing scale, for several reasons: firstly, commercial-scale production is unnecessary from a demand perspective during product development, leading to more waste, which is environmentally and financially unwise. Secondly, production of a larger number of smaller-scale lots (compared to a smaller number of larger-scale lots) is more attractive from a scientific perspective: an increased number of production runs and lots offers better insight into the variability of the process and a larger sample of materials available for comparative testing. Production scale-up may coincide with a manufacturing site transfer, which in turn triggers various related changes, such as differences in equipment, small revisions to the process to accommodate the change in scale, possibly different suppliers of certain materials already qualified by the receiving site, and so forth. Regulatory Agencies acknowledge this reality of product development and can tolerate inclusion of batches originating from a different version of the biosimilar manufacturing process, provided that the differences versus the proposed commercial process are not clinically meaningful, as demonstrated through appropriate comparability assessments in line with ICH guideline Q5E (ICH 2004).

Correlation between biosimilar test lots should be understood as the degree of dependence between those lots, usually due to (predictor) variables that are identical between those lots. Correlation between product lots is unavoidable

and expected: certain variables are deliberately fixed to ensure consistency in production. Examples are the master cell bank or the manufacturing site—these may change during a product's lifecycle, but will be identical across several lots. Other production variables exhibit some degree of variation between lots, whether within a controlled range (for suspected or known critical process parameters), or without explicit process or other control. Variation of all (critical) variables across the lots that are selected for analytical similarity testing is very difficult to achieve in practice; for example: the same lot of raw material (or of other materials: chromatographic media, disposables . . .) can be used in several production runs. Similar correlation can also be expected between lots of the reference product, and will in any case be unknown to the sponsor. The degree of correlation between biosimilar lots is therefore best modeled to what is reasonably expected for the reference product: each lot of finished product should originate from a different fermentation batch of the protein (drug substance), and lots of drug substance and drug product should be manufactured in different campaigns, that are appropriately spaced across their respective shelf lives. These measures offer balance between the degree of process variability that is reflected in the sampled product lots, and (assumed) correspondence between the degree of inter-lot correlation for the biosimilar and reference product.

Selection and Ranking of Product Quality Attributes

Introduction

ICH guideline Q8(R2) defines the term 'quality attribute' as a 'physical, chemical, biological or microbiological property or characteristic'. Identification of quality attributes (and particularly: potentially critical quality attributes i.e. attributes that must be within an appropriate limit, range, or distribution to ensure the desired product quality) is a normal part of pharmaceutical development of any medicinal product (ICH 2009). In the case of biosimilar product development, early identification of quality attributes with their associated ranges and relative criticality is highly advisable as it enables a target-directed development approach, building similarity into the product from the onset. Consequently, a mature list of quality attributes and a robust understanding of the criticality of each attribute should be available by the time the program reaches the stage of pivotal analytical similarity testing.

The content of this section focuses specifically on the selection and ranking of product quality attributes in the context of analytical similarity testing. The principles, concepts and approaches discussed can however equally be applied when creating the quality target at the start of biosimilar development.

Selection of Quality Attributes

A discussion of the mechanics of attribute identification and selection best starts with appropriate clarification of the term "quality attribute". The regulatory definition of the term leaves some room for interpretation, which can be obstructive in later stages of the process. The two most common errors in this regard are confusion between quality attributes and analytical test measures, and inadequate definition of the attribute.

A quality attribute is a description of a specific physical, chemical, biological or microbial property of the product, which in many cases can be measured by different analytical techniques. It is best defined in a manner that supports evaluation of its relationship with the clinical performance of the product. This is different from an analytical test measure, which is usually linked to a specific technique (or group of techniques sharing the same separation principle), and is merely a measurement of one or several quality attributes. For example: the charge pattern of a product is an analytical test measure, providing analytical information on molecular variants of the product carrying a different charge. Charge isoforms can be the expression of different quality attributes, alone or in combination, such as chemical modifications of the protein, charged glycan species, truncated species, N- or C-terminal protein variants etc. As another example: protein mass may seem a genuine quality attribute, but it is not possible to assess the relationship between the mass of the protein (not to be confused with size variants) and the clinical performance of the product. Such an assessment is however possible for the true quality attributes that are reflected by protein mass, such as primary sequence, or the relative abundance of mass isoforms associated with specific protein variants.

The description of the quality attribute must be appropriately specific so that its clinical relevance can be uniquely assessed. To explain with an example: consider the popular quality attribute 'degree of deamidation'. Protein deamidation can occur at different loci of the protein, but not all deamidation events have the same effect on protein function, *in vivo* fate, host response, or stability of the product. It therefore deserves recommendation to define different deamidation-related quality attributes, e.g. grouped per their suspected or known clinical impact.

Quality attributes for the biosimilar candidate can be identified from a wide variety of sources. In an early stage of product development, attribute selection is mostly informed by external intelligence sources, such as regulatory sources and scientific literature. With growing product knowledge gained from the company's internal experimental work, the list of attributes is further refined and improved. Figure 11.5 provides a summary of the most important sources for selection of product quality attributes.

A few important rules must be followed when creating and refining the list of quality attributes. First of all, this list does not serve the single purpose of designing the analytical similarity study program. It is better to be comprehensive and select all relevant quality attributes (including quality attributes that are not relevant for the analytical similarity study e.g. microbial safety parameters, certain

External Sources	Internal Sources
• Regulatory • Guidance documents • Pharmacopoeia • Regulatory assessment reports • Labels • Post-marketing change history • Literature • On reference product • On related products • Other • Reference product COAs	• Analysis of reference product lots • Pharmaceutical studies on biosimlar e.g.: • Stability • Forced degradation • Formulation characterization • Analytical testing • Process design (input materials, unit operations, ...)

Fig. 11.5 Sources of product quality attribute selection

physical properties of the drug product), as this can prove invaluable for other uses; examples are process development, or definition of the control strategy for the product. Secondly, care must be taken not to omit attributes that may be less obvious, for example:

– Control of non-target biological functions: all possible effector mechanisms of the product must be considered, including those that are not intended or expected to contribute to the mechanism of action (for example: Fc-related effector functions for therapeutic antibodies only intended to neutralize soluble signal molecules by inhibitory binding, e.g. bevacizumab, adalimumab). Equally, possible affinity for endogenous proteins that are homologous to the product's biological target, or non-target receptors for which the product may have affinity and would thus alter its disposition, must be taken into account.
– Target heterogeneity: in case of polymorphism of a product ligand, the affinity and kinetics of binding with the different phenotypes may be clinically relevant and must therefore be evaluated.

Finally, quality attribute listing is not a one-shot exercise. The body of product knowledge increases with time, a process that continues until the end of the product's lifecycle. Periodic re-evaluation of the list will therefore ensure that it reflects the current state of knowledge at all times, which in turn prompts necessary updates to all related product monitoring and control systems.

Criticality Assessment and Ranking of Quality Attributes

The objective of criticality assessment is to determine the relative importance of each quality attribute for the clinical performance of the product, and hence the priority of controlling that attribute within limits known to deliver product of the expected quality. In general, each criticality assessment procedure aims to respond at least the following two questions for each attribute:

1. What is the effect on the patient receiving the treatment if the attribute is not controlled within acceptable limits? What is the expected impact on the biological function/pharmacodynamics/efficacy, pharmacokinetics, safety and immunogenicity of the product?
2. What type of information/evidence is used to support the response to the former question, and what degree of uncertainty remains (i.e. what is the probability of reaching a false conclusion)? This question measures the extent of scientific inference that is necessary for evaluation of the clinical impact. Indirect evidence of the relationship between an attribute and the clinical properties of the product requires use of assumptions, extrapolation and logical inferences, and thus results in greater uncertainty of the trueness of the conclusion.

The information sources that can be used to support the assessment are to a large extent the same as those used for identification of quality attributes, presented in Fig. 11.5. To ensure a meaningful criticality assessment with a reasonable level of certainty of the impact assessment decisions, appropriate use must be made of internal evidence. Various experiments can offer invaluable insights into the relationship between individual quality attributes and the biological functions or fate of the product. Appropriately designed forced degradation studies help to elucidate the effect of specific physicochemical changes of the product on biological effector functions through *in vitro* assays, including the possibility to quantify the relationship (see section "Selection of Stress Conditions and Test Methods" for further discussion). Product-related substances or impurities can also be obtained from in-process samples (unpurified product) or targeted chemical or enzymatic treatment of the product. A more specific assessment of the structure-activity relationship of a given isoform is possible when isolating it from an enriched sample using suitable purification techniques. Other data also contribute to the understanding of attribute criticality. For example, a significant degree of variability in the results obtained from analysis of different batches of the chosen reference product suggests that the associated quality attribute is not tightly controlled and may therefore not be critical. For the sake of completeness: valuable insights may also be gained from the results of *in vivo* studies, if available, although such data must be interpreted with caution given the substantial number of variables that contribute to the response being measured.

Different assessment models are available for determination of the criticality of quality attributes. Several examples are described in the A-Mab case study (CMC Biotech Working Group 2009), although some companies choose to create their internal version of these tools (Alt et al. 2016; Stangler 2011). Broadly, criticality assessment models can be categorized into qualitative or quantitative approaches, each with their merits and drawbacks, as summarized in Table 11.2.

Product developers have a free choice of their criticality assessment approach, can engineer their own unique corporate procedure bottom-up, or improve an existing model. Whatever option is selected, successful criticality assessment procedures share a few common traits:

Table 11.2 Qualitative vs. quantitative criticality assessment approaches

	Quantitative	Qualitative
Description	For each dimension in the assessment model, a scale with discrete numerical scores is defined. Each score is associated with a description of the meaning of that score. The evaluator selects the appropriate score for each assessment dimension, and then computes an aggregate score, typically (but not always) using a predefined simple arithmetic such as multiplication of the sub-scores for each dimension The aggregate scores are used to rank the attributes in order of criticality. Some models also categorize the attributes, for example high-moderate-low criticality	Similarly to quantitative models, a scale of different levels with associated definitions is used for evaluation within each criticality dimension, although no numerical scoring system is used The criticality of the attribute can then be determined in various ways, e.g. – Use of a grid with one axis per assessment dimension; criticality categories are predefined for each intersection (combination) of dimension levels (=visual approach) – Use of a set of rules defining the criticality category for each combination of dimension levels (=descriptive approach)
Merits[a]	– The numerical score for each attribute enables sub-sampling within each criticality category for specific purposes, and offers finer resolution of the relative importance of quality attributes – Easy determination of overall attribute criticality (computation)	– Intuitive approach: can be easily understood and hence implemented – Usually simpler than quantitative models and thus less laborious
Drawbacks[a]	– Small changes in sub-scoring due to subtle differences in interpretation can result in a different criticality category (effect of multiplication)	– May promote more subjective assessment ('expert opinion' vs. fully justified and documented scoring decision)
Case example	Sandoz (U.S. FDA CDER 2016)	Amgen (Karow 2016)

[a]The described merits and drawbacks of each approach are indicative only. Careful design and thoughtful implementation of the assessment model can overcome the mentioned differences

– The fundamental questions listed at the start of this section must be addressed.
– Determination of the impact a quality attribute may have on the product's clinical performance often requires a judgment of the worst-case situation (e.g. attribute levels, or deviation from the desired target) that is conceivable for the attribute. In such instances, the evaluator must consider a situation of complete lack of control, be it by process design or control, quality control, or other measures taken to assure the intended quality of the product. The criticality of a quality attribute is an intrinsic property; the possible impact that the attribute may have on efficacy or safety does not change when it is better controlled (although the *risk* of occurrence of the impact decreases). The criticality rank of an attribute dictates the importance of maintaining it within an acceptable range, and thus the level of control required. To illustrate this, consider the case of host cell proteins, which are widely recognized as risk factors for immunogenic responses to therapeutic protein products (U.S. FDA 2014). Host cell proteins are efficiently cleared by conventional downstream purification platforms and are therefore rarely a cause of concern in clinical practice. The low levels of host cell proteins typically present in biological medicines is however the consequence of their known critical impact; controlling them to trace levels does not change their criticality.
– In case different quality attributes relate to the same mechanism of clinical impact, an assessment hierarchy must be defined to ensure that the criticality is not assigned to each individual quality attribute, but only to the most relevant quality attribute. The criticality of the other quality attributes must then be appraised independently from their link with the already counted mechanism of clinical impact.
– The list of quality attributes covers a wide diversity of different categories, such as molecular variants of the product, undesired substances and organisms, biological functions, structural characteristics, physical properties and so on. Finding a single approach that is suitable for criticality assessment of all of those items is difficult. A combination of different approaches, or different variations of the same basic model but adapted to the type of attribute being assessed, may be a more practical solution than finding a unique one-size-fits-all procedure.
– Regardless of the mechanics of the model being selected, all assessment conclusions must be fully justified, based on sound scientific principles and concepts, and must be substantiated with references to the evidence sources used as a basis for decision-making.

There is currently no mandatory regulatory standard for criticality assessment of quality attributes. Notwithstanding, U.S. FDA has provided clear recommendations in their draft guidance document on statistical approaches to evaluate analytical similarity (U.S. FDA 2017). In this document, the Agency proposes to evaluate the potential impact of an attribute on clinical performance, and the degree of uncertainty around a quality attribute, thus following the basic rules of Tool 1 described in the previously referenced A-Mab case study (CMC Biotech Working Group 2009). The Agency deviates from the scoring principles of Tool 1 by advising

to prioritize attributes *known* to be of high impact (=high impact × low uncertainty) over attributes *suspected* to be of high impact (=high impact × high uncertainty). The draft guidance document was withdrawn by U.S. FDA on the 21st of June 2018. In their press release, the Agency announced future publication of a new guidance document that will better address the scientific and regulatory issues raised in public comments.

Unlike U.S. FDA, EMA has not taken a formal position on the desired approach to attribute criticality ranking, although allusion is made to the expectation of criticality-ranked quality attributes in their 2017 draft Reflection Paper on statistical methodology for the comparative assessment of quality attributes (EMA 2017a).

Analytical Testing Plan

Criteria for Selection of Analytical Techniques

Expectations for the choice of analytical techniques used to evaluate analytical similarity are clearly outlined in regulatory guidance documents. EMA and U.S. FDA recommend an extensive characterization program using state-of-the-art analytical technologies, capable of discerning minor differences between the biosimilar candidate and the reference product. Sponsors must understand the limitations of individual methods and apply orthogonal testing to ensure complete coverage of all relevant product quality attributes (see section "Selection of Quality Attributes"), thus minimizing the extent of residual uncertainties to be addressed in downstream *in vivo* studies (EMA 2014b; U.S. FDA 2015b).

One implication that may not be fully appreciated, particularly by new entrants in the biosimilar industry, is that pharmacopoeial monographs for biotherapeutic agents do not constitute a sufficient standard for similarity testing. Monographs only describe a minimum testing panel, often involving widely available yet not state-of-the-art analytical technology. Companies must therefore critically review the suitability of each analytical procedure described in the monograph (if one exists for the product of interest) for similarity testing, and choose alternative technologies or procedures if appropriate.

The ability of an analytical method to measure small differences between two products depends on its resolving power (in case of a separation procedure), the sensitivity of detection (i.e. signals that can be identified as true analytical responses distinguishable from background noise), but also the precision of the analytical procedure. Intrinsic variability of the analytical measurement method may mask small yet relevant differences between the products. It is furthermore important to understand the limitations of each analytical procedure, and to compensate it by use of orthogonal test procedures. For example: differences between products may become undetectable by the method due to sample preparation procedures (chemical or physical alterations necessary for analysis), due to the test procedure

itself, or due to limitations in analyte detection or measurement of responses. Such limitations must be carefully considered, by a thorough understanding of the method's operating principles and how information can be lost or altered by the measurement itself. In most cases, comprehensive assessment of quality attributes is only possible by a combination of analytical methods based on orthogonal test principles.

Selection of analytical methods can also be driven by the nature of the resulting data. In particular, methods capable of producing (semi-)quantitative results may, *ceteris paribus*, be preferred over methods only yielding qualitative information, which is less amenable to objective comparisons against predefined decision rules (e.g. by use of statistical methodology). Thorough knowledge of analytical technology is again needed to fully appreciate all possible data processing options for a given technique.

Although scientific considerations predominate during analytical method selection, final decisions are also guided by practical constraints. Examples are the availability of or access to specific techniques, or financial constraints.

Once a first broad selection is made, usually based on theoretical grounds in combination with regulatory standards (i.e. inclusion of methods mandated by applicable regulation or otherwise expected by regulatory agencies = "gold standards"), the testing panel can be further refined through experiments. This can be achieved by deliberate alterations to the product, e.g. by chemical, physical or biological treatment, yielding samples used to challenge all relevant methods. Methods that perform poorly despite optimization attempts, and do not provide unique information (unavailable from other analyses), are unsuitable for use in the analytical similarity assessment program. Refer to section "Selection of Stress Conditions and Test Methods" for further discussion of method selection during forced degradation screening experiments.

Verification of Analytical Method Suitability

Analytical methods used for assessment of similarity must be demonstrated to be suitable for their intended purpose. Clear guidance on the specific parameters that must be evaluated to satisfy regulatory agencies is however unavailable. Note that the terms "method qualification" and "method validation" are deliberately avoided in this context, to avoid confusion caused by different interpretations of these terms. Instead, the generic term "method suitability testing" will be used in this section and should be understood as the experimental confirmation of the suitable performance of an analytical test method for its intended purpose.

Many different approaches are used by analytical laboratories to produce evidence that the method is well-controlled and consistently delivers reliable results which can be used for regulatory decision-making. These approaches share some

common principles. For one, the intrinsic nature of the method informs the type of suitability study required. Compendial test procedures generally require limited studies, with the expected testing defined in the Pharmacopoeia. Methods that measure fundamental biophysical properties of the product, with little or no sample manipulation, do not normally require extensive studies either (provided suitably calibrated equipment is used). Demonstration of method specificity (particularly, lack of interference by components other than the target analyte, such as excipients) and method precision (repeatability and intermediate precision) is sufficient in most cases, unless another parameter deserves verification (e.g. linearity of response). In this regard, it must be noted that the ability of the method to measure subtle changes in the product is ideally demonstrated during the method screening phase (see section "Criteria for Selection of Analytical Techniques"). Methods involving sample preparation procedures and/or other interactions with the test material (e.g. all separation procedures) necessitate more elaborate suitability studies.

The intended use of the method also determines the extent of method suitability testing. Methods that are only proposed for characterization testing at a single laboratory, under repeatability conditions (head-to-head analyses), do not require the same evaluation as quality control methods used for routine analysis of the product. As another example: the performance of methods controlling the most critical quality attributes (e.g. measurement of target mechanism of action) is assessed in far greater detail than non-critical test methods.

A simple generic decision guide for determining the appropriate level of method suitability testing is provided in Fig. 11.6.

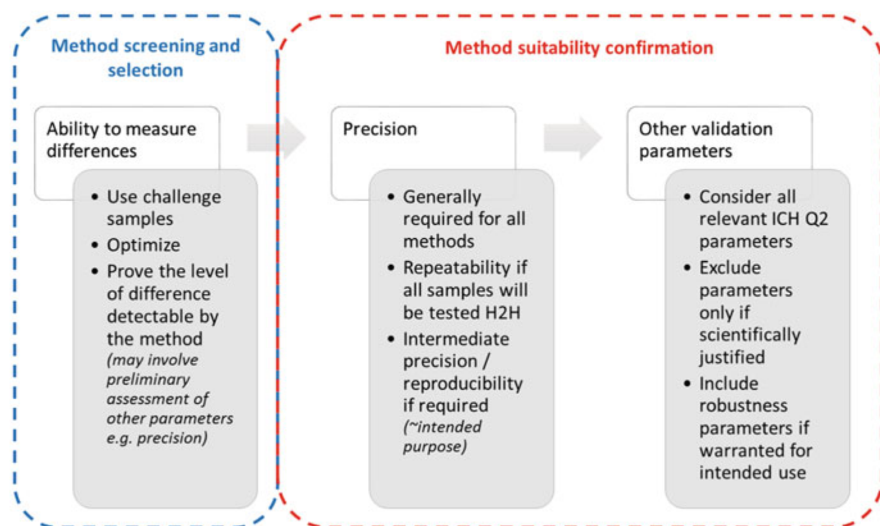


Fig. 11.6 Method suitability parameter selection

Testing and Data Interpretation

The ability to measure small differences between products is also dependent upon the operational design of the similarity study. These principles are generally well understood, as they are the same as those followed for analytical comparability testing, a common experiment within the biopharmaceutical industry. Measures can be taken at the level of the individual test procedure, for example: analysis of product characteristics associated with significant heterogeneity can be improved by selectively removing part of the heterogeneity using appropriate chemical or enzymatic treatments (alone or in combination). As such, instead of comparing complex data patterns which may hide subtle yet relevant differences, samples are processed to facilitate interpretation and increase the probability of finding differences, should any exist. Such selective processing of samples is however not possible in all cases, due to e.g. destabilization of the product (preventing further analysis), or only incomplete cleavage of the targeted bond (biasing the evaluation).

The most common measure to avoid extraneous variability complicating similarity assessment, is direct comparative analysis, whereby all samples are analyzed in a single analytical sequence under repeatability conditions. Head-to-head testing is not possible in all cases: the number of samples that can be analyzed in a single run is limited for certain methods. Or, the sourcing of test batches is spread across a significant interval of time, preventing analysis of all samples within their shelf life in a single study. When comparative testing must be staged for such reasons, it is important to control sources of analytical variability to the extent possible. The same analytical procedure must be used, with careful control of all critical test variables (identified during method development, optimization and suitability studies). As an additional precaution, it is advisable to analyze the same, stable control sample in all occurrences alongside the test samples, serving as a point of reference and thus a link between the different studies.

A focus on differences is also essential when processing the test data. The emphasis of analytical similarity testing does not lie on confirming the correspondence between the products, but on searching evidence of differences between the products. This may seem as semantics, but represents a fundamentally different approach to the data. For example: when comparing the disulfide bonds between products, the review of the data should not be limited to confirmation that both products possess the expected disulfide bonds. Instead, low-level signals of mismatched disulfide bonds must also be compared between the products, as well as the levels of structural variants (such as thioether links or trisulfide bonds) (Liu and May 2012). Another manifestation of the difference in approach is in the choice of the reportable data for comparison. Conventional reporting often restricts to (semi-)quantitative information on major isoforms, bearing a substantial risk of missing small yet consistent differences in the analytical patterns between the products.

Assessment of Analytical Similarity Data

Introduction

Prospective planning of the assessment of the analytical data is indispensable to avoid confirmation bias during review of the results. Pre-definition of the assessment criteria and of the assessment methodology is therefore necessary, and informs other study design decisions, such as the type and number of test samples of the proposed biosimilar and reference product.

In general terms, definition of similarity acceptance criteria implies a judgment of the extent of difference that can be tolerated between the proposed biosimilar and the reference product, without impact on the product's clinical performance. The effect of variations in product attributes on biological activity and/or clinical characteristics has been extensively studied for many quality attributes. Well-known examples are the relationship between core fucosylation of N-glycans and Fcγ3A receptor binding affinity (which, in turn, correlates with antibody-dependent cellular cytotoxicity (ADCC)) (Abès and Teillaud 2010; Zhong and Somers 2012); or the relationship between aggregation and immunogenicity (whereby the immune response and clinical consequence depend on aggregate characteristics such as loss or preservation of native epitopes, aggregate size and solubility; as well as on other factors such as frequency and route of administration, host immune status, presence of product impurities, epitope shielding by attached PEG or oligosaccharides, to name a few) (Rosenberg 2006; U.S. FDA 2014; Freitag 2012). An exact quantitative relationship between a product quality attribute and the effect in humans (PK/PD, efficacy, safety, immunogenicity) is however rarely known. Without such knowledge, definition of an 'acceptable difference' is very difficult to impossible. The most conservative approach to similarity acceptance is then to require absence of difference between test and reference product. Such stringency may be appropriate for critical product attributes with a likely or known effect on the product's clinical performance; for other quality attributes, some degree of variation will have no adverse clinical impact and should hence be permitted. EMA and U.S. FDA acknowledge this rank order of attribute criticality, and propose or accept the concept of applying different levels of rigor in similarity evaluation commensurate with attribute criticality (EMA 2017a; Tsong et al. 2015) (refer to section "Selection and Ranking of Product Quality Attributes" for further discussion of criticality assessment). The Agencies however diverge in their expectations and instructions to sponsors of biosimilar medicines for assessment and interpretation of analytical data. A brief overview of the methodology advocated by EMA and U.S. FDA is provided in sections "U.S. FDA" and "EMA" below. The reader is referred to other chapters of this book, or to the list of citations, for further detail.

U.S. FDA

In their draft guidance document (U.S. FDA 2017),[2] U.S. FDA offered precise instructions to sponsors of biosimilar products on the methodology for interpretation of analytical similarity data: each analytical test measure in the proposed study must be assigned to one of three tiers, each of which is associated with a specific method of similarity assessment. Note the use of the term 'analytical test measure', not quality attribute: an individual quality attribute can often be measured by various orthogonal analytical techniques. Some techniques yield various types of information, e.g. (semi-)quantitative information from signal processing in addition to an analytical fingerprint of the sample (such as a spectrum or chromatogram). 'Analytical test measure' can then be defined as a specific result of interest that is proposed for similarity evaluation, whereby a single technique may provide more than one measure.

Factors of consideration for tier assignment of individual analytical test measures are the criticality of the corresponding quality attribute (high-moderate-low); the nature of the analytical output (e.g. continuous numerical vs. analytical fingerprint vs. binary); the measured levels of the target analyte; and analytical considerations such as method performance (sensitivity, precision ...) or other tests in the study with relevance for the same clinical characteristic. As an example of the latter consideration: for a product with important Fc-related bioactivity, such as trastuzumab (EMA 2017b), the analytical study will include binding assays versus targets such as HER2, Fcγ receptors and complement c1q; and cell-based bioassays such as ADCC and CDC (complement-dependent cytotoxicity). Cell-based bioassays provide a more direct measure of the clinically relevant mechanism of action, and would therefore be assigned to a higher tier (most likely Tier 1), which is associated with more stringent similarity evaluation criteria. The binding assays measure each component of these antibody effector functions separately (Clynes et al. 2000), and enable evaluation of binding affinity for specific (medically relevant) (Varchetta et al. 2007; Tamura et al. 2011) receptor polymorphs, but have less direct clinical relevance (compared to the cell-based assays) and can therefore be assigned to a lower tier (Tier 2). Overall, U.S. FDA expects the most clinically relevant potency assays to be included in Tier 1 (U.S. FDA 2017). Other analytical test measures will be assigned to Tier 2 or Tier 3 based on criteria such as those listed above. Figure 11.7 provides a summary of the most important decision rules for tier assignment of analytical test measures.

U.S. FDA proposed inferential statistical methodology for assessment of data in Tier 1 and Tier 2, restricting these tiers to evaluation of continuous data only. Analytical measures that are assigned to Tier 3 require qualitative similarity

[2]The draft guidance document was withdrawn on the 21st of June 2018, in response to public comments raising regulatory and scientific issues with the proposed guidance. U.S. FDA announced publication of a new guidance document at an undisclosed date. The methodology described in this section reflects the Agency's recommendations prior to withdrawal of the draft guidance document.

Rule 1: Criticality of the Attribute

- Highest criticality → Tier 1
- Moderate criticality → Tier 2
- Low criticality → Tier 3

Rule 2: Nature of Analytical Information

- Continuous numerical: tiering depends on other rules
- Other data format (discrete, visual, categorical...): Tier 3

Rule 3: Orthogonal Measurement of Attribute

- Best performing test (for detection of differences): highest tier (per other rules)
- Other tests: may be assigned to lower tier

Rule 4: Attribute Levels

- *Example*: reference / biosimilar < LOQ (of suitably sensitive test) ⇒ not amenable to statistical analysis ⇒ Tier 3
- *Example*: low attribute levels in biosimilar and reference suggest that a significant impact is unlikely ⇒ lower tier possible

Rule 5: Common Clinical Performance Link

- Different critical attributes relating to the same clinical characteristic (e.g. efficacy) ⇒ assign the most direct link to the highest tier (others: tier down)

Fig. 11.7 Decision rules for tier assignment of analytical test measures

assessments, such as visual comparison of (overlaid) analytical patterns and data plots, conclusions drawn from descriptive statistics of numerical data, or other methodology that is appropriate for the type of analytical information to be processed.

U.S. FDA proposed statistical equivalence testing for Tier 1, whereby the *X*% confidence interval of the difference between the proposed biosimilar and reference product must be entirely contained within the Equivalence Acceptance Criteria that are calculated as *f* times the sample standard deviation of available reference product results. By default, a 90% confidence interval is proposed (equating to two one-sided t-tests each at a significance level α of 5%), and a multiplier *f* of 1.5. FDA has calculated that under their standard model, and assuming a true mean difference between the proposed biosimilar and reference product of $\sigma_R/8$ (with σ_R being the population standard deviation for the reference product), a statistical power of 84% is achieved if ten lots of each product are tested (Tsong et al. 2015). Alternative proposals, such as an increased Type I error (e.g. α of 10% instead of 5%) can also be accepted, if compensated by other control measures. In the example provided, stringent control of the concerned analytical measure will be expected as part of routine product control (e.g. by setting appropriate specification criteria).

Tier 2 data evaluation starts with computation of a Quality Range, which is defined as $\overline{X}_R \pm k \times s_R$, whereby $\overline{X}_R$ and s_R are the sample mean and sample standard deviation of available reference product results, respectively; and k is a multiplier (Tsong et al. 2015). The value of the multiplier must be proposed and justified by the product sponsor, and varies depending on considerations such as the criticality

of the concerned quality attribute and the performance of the analytical test method. Usually, a multiplier value ranging between 2 and 3 is expected by FDA (Dong 2015). Similarity is then defined as a minimum proportion of biosimilar results that must fall within the calculated quality range. FDA proposes a threshold of 90% of results falling within the quality range.

Several concerns have been raised by stakeholders with regards to this methodology, such as the risk of setting artificially tight acceptance criteria due to (unknown) correlation between tested reference product lots; large sample sizes required to achieve the desired statistical power; the effect of including other comparators (licensed in a different jurisdiction) on the statistical power of the analytical similarity study; and so forth. The statistical power of a proposed program can be estimated by use of Monte Carlo simulations, whereby the impact of different assumptions can be easily tested. For a detailed discussion of FDA's proposed statistical assessment methodology and its possible limitations, refer to other chapters of this book.

EMA

Although EMA has first established a legal basis for marketing authorization of similar biological medicinal products in 2004 (European Parliament and Council 2004), and has reviewed tens of applications since that time, the Agency waited until March 2017 to publish its view on the assessment of similarity in quality attributes in a draft Reflection Paper (EMA 2017a). In this document, EMA discusses relevant concepts for comparative evaluation of quality characteristics between drug products, including but not limited to biosimilars, but does not provide clear instructions for the design of the (statistical) data analysis plan. This is a distinctly different approach from the one taken by U.S. FDA, in that the latter had opted to offer clear, prescriptive guidance to industry.

The draft Reflection Paper acknowledges several concepts that are also found in U.S. FDA's analytical similarity data assessment model. For instance: the notion of using statistical techniques of different rigor for different categories of quality attributes is discussed, with statistical equivalence testing as the most suitable technique for the purpose of similarity assessment (although it is specified that certain quality attributes deserve a different approach, e.g. non-inferiority testing for comparison of impurity levels). Another example of a shared vision between the Agencies is the concept of 'residual uncertainty', whereby the degree of analytical similarity achieved determines the required extent of downstream in vivo comparisons. Overall, however, the document dedicates much attention to the factors complicating or preventing the use of inferential statistical methodology in the context of quality data comparisons, whereby certain conclusions may be interpreted as representing a different position compared to U.S. FDA's approach.

EMA's experts put much emphasis on the conditions that must be met in order to use inferential statistical methodology. For example: in the case of continuous

measurements, sampled data can be used to estimate population parameters for each drug product, which can subsequently be compared (through a suitable metric, e.g. the difference or ratio of means). Parameter estimates assume an underlying distribution of the data, usually a normal distribution, verification of which is complicated by the often limited datasets in quality comparisons. EMA nevertheless expects sponsors to discuss whether the assumptions underlying the choice of a specific methodology can be considered fulfilled. The use of inferential techniques may also be limited by the ability to obtain representative samples, usually achieved through random sampling. This, in addition to the level of understanding of the sources of variability (also impacting the sampling plan), must be duly considered when selecting the statistical methodology. The expectation to justify the validity of a proposed statistical methodology exemplifies a fundamental difference compared to U.S. FDA, where departure from the assessment principles selected by the Agency would instead require strong justification. Put otherwise: unlike U.S. FDA, EMA does not advocate a default data assessment model for analytical similarity, but expects a substantiated proposal from the product sponsor. U.S. FDA's proposed methodology is however under revision, and more flexibility for sponsors developing biosimilar products has been announced in future draft guidance.

In their draft Reflection Paper, EMA discusses common approaches for comparison of quality data, and explains why some conventional methods are unsuitable for regulatory decision-making. Comparative quality data assessments, such as those used to evaluate the similarity between a proposed biosimilar and its reference product, must aim to compare the entirety of the material that originates from each manufacturing process. Test results from samples of each manufacturing process are manifestations of these processes and can be used to estimate data distribution parameters such as the mean and variance (provided the quality attribute of interest is measured on a continuous scale, representative samples were drawn, and underlying assumptions of process consistency and data distribution are fulfilled). Such estimates are associated with uncertainty, which is addressed by computation of appropriate statistical intervals. A statistical interval can also be computed for the metric chosen to describe the difference or distance between the two data populations (biosimilar and reference product), e.g. the difference or ratio of the means. To determine whether the two processes produce material of non-inferior or equivalent quality (depending on the objective of the comparison), the computed interval for the difference between the products must be judged against a predefined acceptance criterion, representing the maximum allowable difference. Acceptance criteria, per EMA, are not based on statistical considerations, but should be scientifically justified. Typically, a difference limit is defined as the maximum difference for the concerned quality attribute that does not result in a meaningful difference in clinical outcomes. In many cases, the relationship between the quality attribute and the clinical performance of the product is insufficiently understood, causing some level of arbitrariness in the definition of acceptance ranges; or other considerations may drive acceptance range setting. Importantly, EMA clearly warns against amalgamating statistical intervals and acceptance ranges, and argues

that many traditional approaches, such as the use of min-max ranges, tolerance intervals and 'x'-sigma approaches, disregard this distinction. Such approaches involve determination of an acceptance range using test results of the material of comparison (i.e. the reference product in case of biosimilar studies), with the expectation that measurements of the test product (i.e. the biosimilar) fall within the calculated range. The reasons why such approaches are unsuitable for similarity assessment are described in the draft Reflection Paper. Of particular interest is the critique on x-sigma approaches, which is essentially the methodology recommended by U.S. FDA for Tier 2 quality measures. EMA argues that an 'x-sigma' acceptance range assumes normality of the underlying data distribution, which often cannot be verified. Furthermore, the choice of 'x' remains arbitrary. The sigma value used for calculation of the range is typically the sample standard deviation, which is an estimate of the population standard deviation and therefore associated with uncertainty. Finally, it can be argued that such approaches tend to reward small sample sizes of test material (increasing the probability of fitting within the calculated range) and may not allow statements on the similarity of the originating process as a whole. Overall, EMA's position seems to differ from U.S. FDA in several respects:

– EMA advises against calculation of acceptance ranges based on results from actual (reference product) samples. This is different from U.S. FDA's advice to compute the equivalence margin for Tier 1 assessment as a function of the sample standard deviation calculated from reference product results.
– U.S. FDA's Tier 2 assessment is a modified 'x-sigma' approach, which is not recommended by EMA.

As a final note: EMA stresses the need for adequate control of false positive conclusions in the context of analytical similarity, particularly when using such data as a basis for approval of an abbreviated in vivo development program. Specific measures to control overall Type I error in the analytical program are not discussed in (now withdrawn) draft U.S. FDA guidance (U.S. FDA 2017) or in available BLA case history.

Forced Degradation

Introduction

Forced degradation (or: stress) studies of therapeutic proteins serve a wide variety of different purposes across a product's development lifecycle, such as drug candidate selection; formulation and process development; identification of degradation pathways and evaluation of the product's stability profile; and determination of the effect of accidental exposure to chemical or physical stress conditions (e.g. during transport or storage), to name a few (EBE 2015; Schmidt 2016; Hawe et al. 2012;

ICH 1995; Chan 2016). Whereas many of these purposes may also be relevant for biosimilar product development, stress studies also serve other objectives which are particularly relevant for analytical similarity assessment. Stressed materials are commonly used for analytical method screening, development and validation, to evaluate the ability of a method to resolve, detect and possibly quantify relevant structures formed during degradation. A carefully designed stress experiment may also contribute to the elucidation of structure-function-relationships, either directly (by studying the correlation between the formation of degradation products and biological activity) or indirectly (by generation of increased quantities of variants of specific interest for isolation and subsequent structural and functional characterization). Such data in turn can aid in criticality ranking of quality attributes. Finally, comparative stress studies are a powerful tool for detection of differences between the biosimilar candidate and reference product, which may not otherwise be detected in head-to-head characterization tests. Appropriate conditions can magnify such small differences through differential reaction to chemical or physical stress, resulting in different degradation kinetics and/or patterns (EBE 2015; Chan 2016). This also explains why forced degradation is a better tool for evaluation of similarity than stability studies: whereas standard stability testing serves to confirm that product quality does not significantly change when stored under recommended conditions, forced degradation studies are meant to expose the product to conditions that cause measurable changes. Intentional degradation in response to a selected chemical or physical 'challenge' provides a better model for detecting (possibly small) differences between products than exposure to conditions intended to minimize changes in quality.

Forced degradation can also complement analytical characterization data that are otherwise difficult to interpret, for example: a difference in the levels of certain modifications found during side-by-side analysis of biosimilar and reference product batches may be caused by a different age of the test materials at the time of analysis (either known or unknown, e.g. a reference product batch originating from a drug substance lot at the end of its shelf life, information which is not publicly accessible). Comparative stress experiments can then be used to measure possible differences in the propensity of each product to form such degradation products, information that may be helpful to interpret the observed difference. Forced degradation may also help to assess structural features that are difficult to comprehensively evaluate by characterization only. One such example is the similarity of chemically modified (e.g. PEGylated) proteins, where the quality profile of the reagent may (critically) impact the stability or *in vivo* performance of the product but is difficult to analyze in the finished product. Differences in the reagent quality (e.g. impurity levels) can cause differential reactions to stress and thus provide an indirect model for assessment of reagent quality (Wang et al. 2014).

The design and interpretation of forced degradation studies described in this section focuses specifically on achieving the objectives related to analytical similarity assessment. Since generic experimental conditions for forced degradation of proteins are unavailable (nor possible, given the widely different responses

of protein formulations to chemical or physical stress), a two-staged approach is advised: first, a screening experiment is performed to identify the most optimal conditions for the product under study. Generally, a level of degradation not exceeding 20% is targeted (EBE 2015; Schmidt 2016; Chan 2016). Too extensive degradation can cause highly complex degradation patterns and kinetics (including interactions between degradants), leading to uninterpretable results. It may however be necessary to target a greater level of degradation to challenge less sensitive test methods, such as higher order structure analyses and functional assays. Another objective of the screening study is analytical method screening and selection, particularly to determine the sensitivity of assays for detection and measurement of degradation products, and to study existing correlations between orthogonal assays. This information can be used to rationalize the testing panel for further analytical studies.

In a second step, a comparative stress study is executed to assess differences in degradation kinetics and/or patterns between the biosimilar candidate and the chosen reference. The design of this study is informed by the results of the screening study.

The design of the forced degradation study program is discussed in further detail in the next sections.

Test Materials

The only evaluation of relevance for analytical similarity assessment is to determine the effect of stress on the finished (drug) product. In the case of lyophilized products, both the lyophilized cake and the reconstituted solution should be subjected to forced degradation (EBE 2015).

Particular attention should go to the choice of the container holding the product during the stress experiment. Certain conditions, including oxidation, pH modification, and conformational stress, require addition of a chemical agent for degradation. The reaction of interest is therefore independent from interactions with the container closure system, and transfer to a suitable sample container (often small-scale plastic recipients) can be justified. Other conditions, involving physical stress (e.g. heat, agitation, thermal cycling, light), could in principle be tested in the drug product container. This is in many cases not practical or desirable: the container fill volume may be far in excess of the required sample volume for testing at a specific measurement point, rendering the study uneconomical. The degradation reaction may not be optimal in the product container, such as in the case of agitation, which is strongly influenced by the interfacial contact surface (air, glass) (Hawe et al. 2012; Kiese et al. 2008). Transfer of the test material to sample containers that are appropriate for the stress condition may therefore be justified. In all cases is it recommendable to transfer the biosimilar and reference product test material to identical sample containers for comparative forced degradation studies.

A possible contribution of container closure system leachables to product degradation however cannot be ruled out. The risk of differences in degradation between products caused by differences in baseline levels of leachables, or by other factors impacting product degradability (e.g. degradation of excipients such as polysorbate (Chan 2016)) can be mitigated by using biosimilar and reference product batches of similar age at the time of the comparative forced degradation study.

The screening study aims to assess the most suitable combination of chemical and physical conditions to yield a reasonable level of degradation. Evaluation of the effect of inter-lot differences on degradation, although relevant, is not the objective of the study—in fact, such differences may complicate the interpretation of the result as they add extraneous variability to the experiment. Therefore, a single lot of product, usually the biosimilar product (for practical and financial reasons), is used for this purpose. The opposite reasoning applies to the comparative study: in this case, the experimental conditions are fixed, to enable measurement of existing differences, if any, between the biosimilar and reference product. Several batches of each product must be tested in order to account for within-product variables that influence the product's sensitivity to stress. A rule of thumb is to use not less than three batches of each product. These batches must not necessarily all be tested at once, provided that testing occurs in pairs (one biosimilar batch vs. one reference product batch in a head-to-head study), and that all batches are stressed and analyzed using the same study protocol. In case statistical assessment of the results is proposed (see section "Design and Interpretation of the Comparative Stress Study"), a greater number of batches may be needed to achieve adequate statistical power.

Selection of Stress Conditions and Test Methods

The stress agents most typically selected for protein forced degradation studies are heat; change in pH (incubation in acidic and alkaline conditions); exposure to oxidizing agents; agitation (mechanical stress); light; and freeze-thawing (EBE 2015; Hawe et al. 2012; Chan 2016). Therapeutic protein formulations may be exposed to such stressors in real-life situations, so information on the impact that may be anticipated in such circumstances is highly relevant. The relevance of the stress condition is however not the most essential criterion for analytical similarity assessment. Other stress conditions, which have less practical relevance but may enhance detection of differences between two products, can also be considered; examples are conformational stress (use of chaotropic agents such as guanidine hydrochloride), high ionic strength, or selective chemical degradation e.g. to evaluate surface-exposure of amino acids sensitive to the chemical agent.

The level of degradation caused by each of these stress factors is highly dependent on the protein and its formulation, and is influenced by the specific experimental set-up (Hawe et al. 2012). The initial screening study allows to find

those stressors that are most informative for similarity assessment, and the specific experimental conditions that result in a relevant extent of degradation. The first focus of the design of the screening study is therefore to identify the appropriate input variables (or: factors) for optimization. The range of conditions studied should be kept as wide as possible, to enable an evidence-based selection of study conditions for the (later) comparative assessment. Figure 11.8 summarizes a possible process for selection of screening factors for heat stress: protein degradation at elevated temperatures is known to be influenced by factors such as the temperature and duration of exposure; the composition of the product (including its pH, protein concentration, type and concentration of excipients, presence and concentration of protein and excipient degradants causing or catalyzing degradation), and the contact surface with the container and headspace interface (Hawe et al. 2012). Optimization of all of these factors is not possible or desirable; for practical reasons, not more than two or (maximum) three factors should be optimized. In the example provided, the composition of the sample and the interfacial contact surface are fixed, by using the unchanged solution and preparing identical fills in identical sample containers, respectively. The most significant factors, temperature and duration, are considered for optimization. *Of note*: a change in the composition of the product may be of interest in certain situations. During development of ABP 501, Amgen evaluated thermal degradation in both the biosimilar product's own formulation buffer, and in a buffer with the same composition as the reference product (Humira®), and found a significant difference in ABP 501's aggregation rate (Liu et al. 2016).

For each factor retained for screening, two or three levels are selected for the experiment, depending on the experimental design chosen (discussed further below). Different information sources together with other considerations can play a role in the ultimate decision of the factor settings, including literature reports,

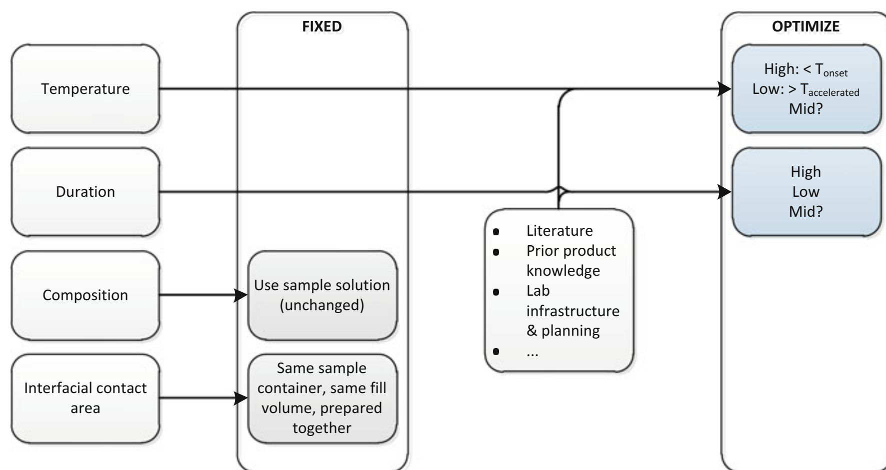


Fig. 11.8 Selection of heat stress screening factors

in-house experience with the molecule, available lab equipment, resource planning and other. A similar selection process is followed for all other stress conditions of interest, whereby care must be taken to map and control or evaluate all relevant experimental variables. Failing that, the pivotal comparative stress study may yield unexpected results caused by uncontrolled variation in critical variables.

The analytical testing panel used for screening must be comprehensive for the same reason as indicated above: to support evidence-driven selection of the test battery used for the comparative study. All techniques that are or may be stability-indicating should be considered for testing. The richness of data that results from a comprehensive testing panel during screening will allow not only selection of the most sensitive techniques for further stress studies, but may also help to demonstrate the suitability of the methods for their intended purpose, and to detect correlations between different structural and/or functional modifications. For example:

- LC-MS analysis of protein digests of degraded samples provides semi-quantitative information on protein modifications (formed in response to specific stress conditions). Analysis of the same samples by charge-based analysis techniques (cIEF, IEX) and other relevant techniques (e.g. RP-UPLC) can help to reveal which techniques are most suited to measure specific degradants.
- Simultaneous analysis of degraded samples by a suite of physicochemical analysis techniques and different functional assays, including measurement of specific effector functions and cell-based assays, can help to establish not only the relationship between specific degradation reactions and the protein's functional activity, but also between discrete biological functions and the effect on target tissues. For example: a specific degradation reaction may measurably affect the dissociation constant for a target ligand, but may not measurably impact a cell-based assay involving that receptor (e.g. if receptor binding is not the rate-limiting step).

Use of a comprehensive testing panel for screening does not mean that each technique must be used for each condition tested. Techniques measuring attributes that are not expected to relevantly change in response to a specific stress condition, or for which more pronounced changes are expected in other experiments, may be excluded from the test panel for that condition. The second focus of the design of the screening study is therefore to select the most relevant response measures for each stress condition, whereby 'response measure' is defined as the specific analytical information that is used for data processing. Quantitative response measures are preferably used for screening, to enable statistical analysis and response surface modeling of the results. The process of analytical response selection for the screening study is illustrated in Fig. 11.9.

The selected screening factors and analytical responses are then used to design the experiment. Use of statistical designs (full or fractional factorial experiments) is highly recommended, as these offer significant benefits compared to a traditional study where each factor is varied independently across the range of interest. One benefit is the ability to predict the response between the actual conditions tested by mathematical modeling (provided the analytical response measure is

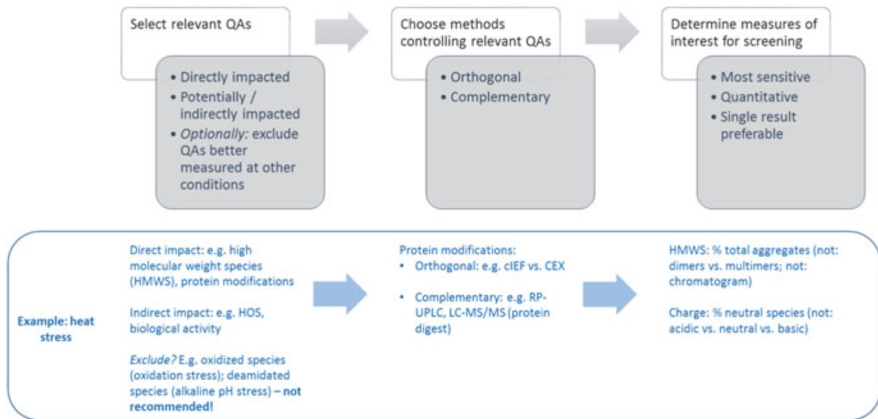


Fig. 11.9 Selection of analytical response measures

amenable to such analysis), which can then be visualized in a response surface plot. The reliability of the model depends in part on the number of levels actually tested for each factor, e.g. only linear regression is possible in case of two-level experiments, whereas a second-order regression model can be fitted in case of three-level experiments. Another benefit of statistical modeling is the ability to detect interactions between the factors tested, if those are present, in addition to main effects. Taken together, a statistically controlled experimental design enables determination of the most suitable combination of factor settings using a minimal number of experimental runs. The suitability of the thus selected settings may subsequently be confirmed in a follow-up experiment.

For example, let us consider a study intended to optimize two selected variables. A common design for this is a central composite design, which is illustrated in Fig. 11.10. The design consists of two components that work together: screen or linear points (shown in black dots, and typically positioned at the corners of a box), and the optimization or quadratic points (shown in grey dots, sitting on the axial

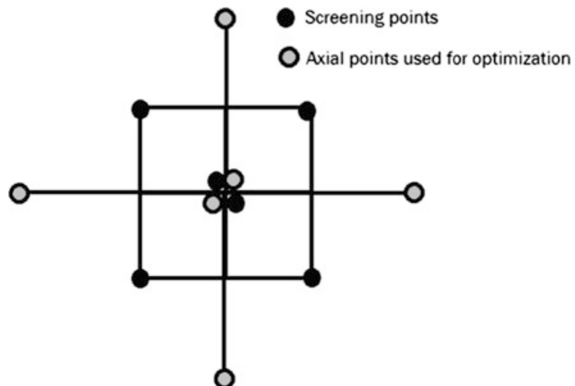


Fig. 11.10 Central composite design built on top of a screening design. The points in the original screening design are shown as black dots, and the augmented points for the optimization design are shown as grey dots. The central points are grouped together, but should be replicates of the same point

positions, either in the box or outside it). The center points are a key component of the optimization design, and should be replicated to achieve good statistical estimation of the response surface. The central composite design allows for a fitting of a quadratic surface, along with the interactions between the variables. Generally, the statistical properties of the design are best when the axial points are out of the box, generally set to the point $k^{1/2}$, where k is the number of variables (2 in this case, so set to 1.4). But if this is not feasible, these points can be brought in towards the box or even on the center of the box lines, and the model will retain its validity. Some advantages of using a central composite design (CCD) are: (1) if a screening design has already been performed, then the CCD can be added to the screening design to reduce the total work burden; (2) most response surface design software can create an appropriate design; (3) the design has some room for error (i.e. is fairly robust again a few missing observations should something go awry in the experiment); and (4) the design estimates the optimal point relatively efficiently. Regression analysis of this design can be performed with a linear regression package, and typically response surface or contour plots will be plotted to estimate the optimum of the process. An example of a response surface plot and a contour plot for a fictitious experiment determining the effect of agitation speed and temperature on protein aggregation is provided in Figs. 11.11 and 11.12, respectively.

Forced degradation studies must also include control runs for appropriate interpretation of the results. Control samples are samples that are treated in the same manner as the test samples, with the exception of the stressing agent(s).

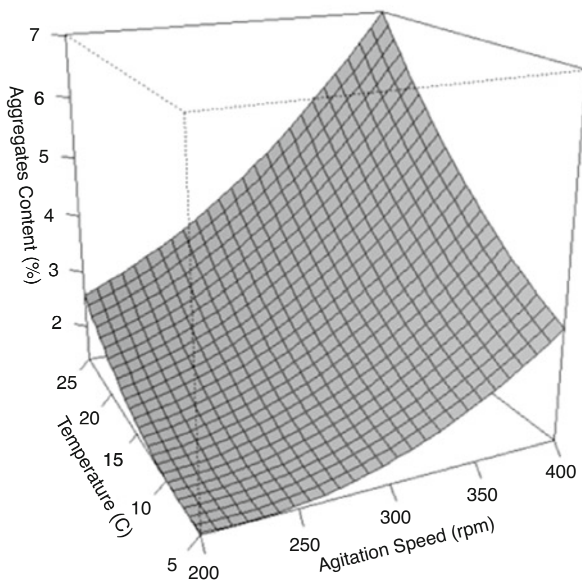


Fig. 11.11 Response surface plot for fictitious agitation optimization study using central composite design

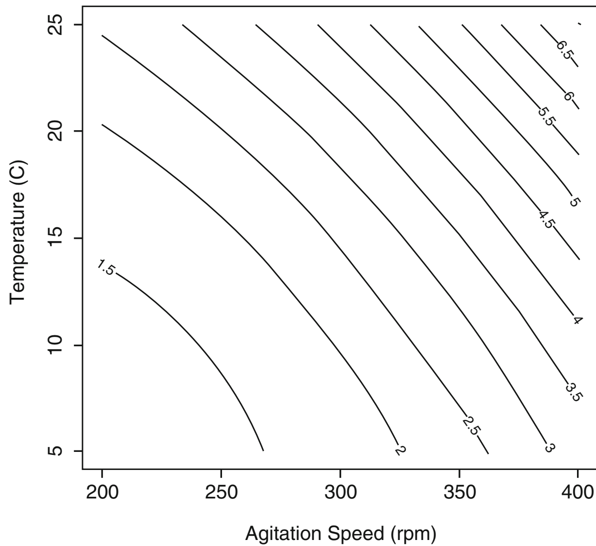


Fig. 11.12 Contour plot for fictitious agitation optimization study using central composite design

Examples are dark controls used in photostability tests, which are incubated next to the light-exposed samples in an identical sample container but protected from light; or samples identical to heat-stressed samples that are stored at the product's recommended storage temperature for the duration of the heat-stress study. Replication of these control samples supports estimation of the experimental standard error, which can be used for statistical analysis of the data.

Another measure to minimize 'experimental noise' complicating the interpretation of the data is the simultaneous analysis of all samples from the same experiment in the same test sequence (to the extent possible). This can be achieved by careful experimental planning, ensuring that all samples reach the end of the study at the same time (by exposing them to the relevant test condition at different times). This principle is illustrated in Fig. 11.13 for clarity. Alternatively, samples can all be incubated at the same time and the stress factor removed or neutralized for samples that reach their endpoint before the overall study end.

The results of such screening experiments can be used for different purposes, as previously discussed, including for optimization of the design of the study comparing the degradation kinetics and patterns between the biosimilar and reference product. Particularly, the experimental conditions that provide the targeted level of product degradation for each stress factor of interest can be derived from the obtained results. This may also include evidence for exclusion of certain stress conditions from the comparative study, if no meaningful degradation is found. The target for degradation may be different for different analytical tests and may be used in the selection of appropriate testing points (e.g. selection of early testing points offering sufficient degradation for more sensitive tests, and late testing points

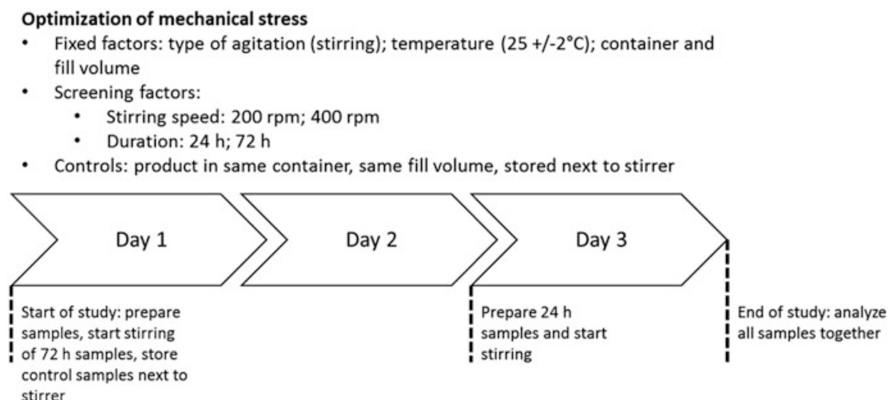


Fig. 11.13 Screening study planning example

for less sensitive tests). Furthermore, the original testing panel can be refined, by excluding certain methods demonstrated to be less sensitive for measurement of degradation (at specific stress conditions). The final testing panel must however still cover all relevant quality attributes, even if only limited or no degradation was observed during screening studies, as the reference product may respond differently to the same stress conditions than the (biosimilar) sample used for screening.

Design and Interpretation of the Comparative Stress Study

Table 11.3 compares the design of the comparative stress study vs. that of the screening experiments. This shows that the comparative study is either informed by or similar in design to the screening study in many respects. Most of the interlinkages and differences between the screening and comparative study were already discussed in the previous sections.

The results of the comparative stress study are used for evaluation of the similarity in degradation patterns and degradation kinetics between the products. Degradation patterns can be compared by careful review of overlaid analytical patterns, such as chromatograms, spectra, electropherograms … or other types of qualitative comparisons. In a first instance, end-of-study samples can be compared, but results from less exposed samples must also be considered, particularly when the extent of degradation at the end of study prevents an appropriate comparison, or to evaluate time or rate of formation of species of interest. Objective criteria for qualitative assessments are difficult to define. One often used decision rule for determining similarity in degradation patterns is to identify analytical signals that are observed in one product, but never found in any batch of the other product when exposed to the same stress conditions.

Table 11.3 Forced degradation screening vs. comparative study

Study parameter	Screening study	Comparative study
Test materials (Refer to section "Test Materials")	One lot, biosimilar only; include control samples	Not less than three lots of biosimilar and of reference product; include control samples
Study conditions	– Study comprehensive set of stress factors – Find optimal combination of experimental settings for each stress factor – Statistically controlled factorial experiment	– Relevant stress factors only (exclude factors that yield minimal or no degradation at reasonable exposure levels) – Testing points and study endpoint informed by screening study – Include at least 4–5 testing points (for evaluation of degradation kinetics)
Analytical test methods	– Evaluate comprehensive list of test methods – Focus on single measure for each method, ideally continuous numerical (where possible), to facilitate mathematical modeling of responses	– Rationalize testing panel (based on screening study results)—but retain capability to measure possible degradation of reference product not detected in the biosimilar screening sample – Collect and compare all relevant analytical information from each method
Study planning	– Ensure that all samples reach their study endpoint at the same time, by careful experimental planning – Test samples together; if not possible, minimize between-assay variation to the extent feasible	
Interpretation	Evaluate most suitable conditions for reaching the targeted level of degradation within an appropriately evaluable timeframe (allowing interim testing points, which may include study endpoints for more sensitive analytical tests)	Compare degradation patterns and degradation kinetics between the biosimilar and reference product

Similarity in degradation kinetics is assessed by comparing the evolution of a relevant numerical analytical measure over the duration of exposure; examples are the formation of a specific degradation product, or the decrease in potency. This is done by plotting the data of interest from each lot of each product, followed by a first visual assessment of the observed data trends. Degradation kinetics may be highly complex and not amenable to regression analysis, in which case only visual assessment is possible. In the event a relevant regression model fits the data (or: part of the data, if departure from the model is only observed for the final testing points), then similarity in degradation kinetics can be statistically tested through comparison of appropriate regression model parameters (representing the rate of change—e.g. slope) between the products.

Differences in degradation patterns or kinetics between the biosimilar and reference product necessitate further investigation, to understand their cause, nature and (biological / clinical) relevance.

Conclusion

Analytical similarity assessment is the cornerstone of biosimilar product development: it is the first and most important step in the sequence of comparative studies, as it provides the most sensitive measurement of differences versus the reference product, and serves as the basis for regulatory decisions on the design of further studies. The importance of a carefully considered and scientifically rigorous program of experiments can therefore not be overestimated. The present chapter offered a set of guiding principles for informed decisions on various design attributes of the analytical similarity assessment program. In the next chapters, readers will learn more about different analytical subjects that are essential for state-of-the-art similarity assessment, as well as gain further insights into the statistical evaluation and interpretation of the analytical data. Ultimately, a well-designed and well-executed testing program will increase the probability of successful biosimilar development and registration.

References

Abès R, Teillaud J-L. Impact of glycosylation on effector functions of therapeutic IgG. Pharmaceuticals. 2010;3:146–57.

Alt N, Zhang TY, Motchnik P, Taticek R, Quarmby V, Schlothauer T, Beck H, Emrich T, Harris RJ. Determination of critical quality attributes for monoclonal antibodies using quality by design principles. Biologicals. 2016;44:291–305.

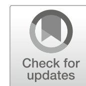
Apostol I, Brooks PD, Mathews AJ. Application of high-precision isotope ratio monitoring mass spectrometry to identify the biosynthetic origins of proteins. Protein Sci. 2001;10:1466–9.

Chan CP. Forced degradation studies: current trends and future perspectives for protein-based therapeutics. Expert Rev Proteomics. 2016;13(7):651–8.

Clynes RA, Towers TL, Presta LG, Ravetch JV. Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. Nat Med. 2000;6:443–6.
CMC Biotech Working Group. A-Mab: a case study in bioprocess development (version 2.1). October 30, 2009.
Dong X (Cassie). 2nd Statistical and Data Management Approaches for Biotechnology Drug Development; 2015 Sept 29–30; Rockville, MD, USA.
EMA. CHMP/437/04 Rev 1: guideline on similar biological medicinal products. October 23, 2014a.
EMA. EMA/CHMP/BWP/247713/2012: guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1). May 22, 2014b.
EMA. EMA/CHMP/138502/2017: reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (draft). March 23, 2017a.
EMA. Herceptin European public assessment report - scientific discussion. [Online] [Cited: May 10, 2017b.] http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000278/WC500049816.pdf.
EMA. Aranesp product information - all authorised presentations. [Online] [Cited: February 23, 2018.] http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_All_Authorised_presentations/human/000332/WC500026150.pdf.
European Biopharmaceutical Enterprises (EBE). Concept paper: forced degradation studies for therapeutic proteins. March 24, 2015.
European Parliament and Council. EU directive 2001/83/EC (As Amended). s.l. : EudraLex Vol. 1; Consolidated Version.
European Parliament and Council. Directive 2004/27/EC. March 31, 2004.
Eurostat. Eurostat: statistics explained. http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Observation_unit.
Freitag AJ. The immunogenicity of protein aggregates: studies on a murine monoclonal antibody in wild-type mice. Ph.D. thesis, Faculty of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München, Munich, Germany (2012)
Hawe A, et al. Forced degradation of therapeutic proteins. J Pharm Sci. 2012;101:895–913.
ICH. Q5C: stability testing of biotechnological/biological products. November 30, 1995.
ICH. Q1D: bracketing and matrixing designs for stability testing of new drug substances and products. February 7, 2002.
ICH. Q5E: comparability of biotechnological/biological products subject to changes in their manufacturing process. November 18, 2004.
ICH. Q8(R2): pharmaceutical development. August 2009.
Karow M. Applying risk ranking for similarity into the QTPP for antibody biosimilars. 2016 PDA biosimilars conference, Baltimore, MD, USA. 20–21 June 2016.
Kiese S, et al. Shaken, not stirred: mechanical stress testing of an IgG1 antibody. J Pharm Sci. 2008;97:4347–66.
Kim S, Song J, Park S, Ham S, Paek K, Kang M, Chae Y, Seo H, Kim H-C, Flores M. Drifts in ADCC-related quality attributes of Herceptin: impact on development of a trastuzumab biosimilar. MAbs. 2017;9:704–14.
Liu H, May K. Disulfide bond structures of IgG molecules. MAbs. 2012;4(1):17–23.
Liu J, Eris T, Li C, Cao S, Kuhns S. Assessing analytical similarity of proposed Amgen biosimilar ABP 501 to adalimumab. BioDrugs. 2016;30:321–38.
Rhiel GS, Chaffin WW. An investigation of the large-sample/small-sample approach to the one-sample test for a mean (sigma unknown). J Stat Educ. 1996;4(3). https://doi.org/10.1080/10691898.1996.11910515
Rosenberg AS. Effects of protein aggregates: an immunologic perspective. AAPS J. 2006;8(3):E501–7.
Schiestl M, Stangler T, Torella C, Cepeljnik T, Toll H, Grau R. Acceptable changes in quality attributes of glycosylated biopharmaceuticals. Nat Biotechnol. 2011;29:310–2.
Schmidt AS. Forced degradation studies for biopharmaceuticals. BioPharm Int. 2016;29(7):54–7.

Stangler T. What to control? CQAs and CPPs. 2011 BWP workshop on setting specifications, London, UK. September 9, 2011.
Tamura K, et al. FcγR2A and 3A polymorphisms predict clinical outcome of trastuzumab in both neoadjuvant and metastatic settings in patients with HER2-positive breast cancer. Ann Oncol. 2011;22:1302–7.
Tsong Y, Dong X, Shen M. Development of statistical methods for analytical similarity assessment. J Biopharm Stat. 2015 Sep 15. https://doi.org/10.1080/10543406.2015.1092038. [Epub ahead of print].
U.S. FDA. Guidance for industry: immunogenicity assessment for therapeutic protein products. August 2014.
U.S. FDA. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. April 2015a.
U.S. FDA. Guidance for industry: quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. April 2015b.
U.S. FDA. Guidance for industry: clinical pharmacology data to support a demonstration of biosimilarity to a reference product. December 2016.
U.S. FDA. Guidance for industry (draft): statistical approaches to evaluate analytical similarity (withdrawn June-2018). September 2017.
U.S. FDA. Aranesp (BLA 103951), SUPPL-5374: label. Drugs@FDA: FDA approved drug products. [Online] [Cited: February 23, 2018.] https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/103951s5374lbl.pdf.
U.S. FDA CDER. Quality review BLA761042 - GP2015 (etanercept), Sandoz. 2016.
Varchetta S, Gibelli N, Oliviero B, Nardini E, Gennari R, Gatti G, Silva LS, Villani L, Tagliabue E, Ménard S, Costa A, Fagnoni F. Elements related to heterogeneity of antibody-dependent cell cytotoxicity in patients under trastuzumab therapy for primary operable breast cancer overexpressing HER2. Cancer Res. 2007;67:11991–9.
Wang W, Ignatius AA, Thakkar SV. Impact of residual impurities and contaminants on protein stability. J Pharm Sci. 2014;103:1315.
Zhong X, Somers W. Recent advances in glycosylation modifications in the context of therapeutic glycoproteins. [book auth.] Dr. Hon-Chiu Leung. Integrative proteomics. London: InTech Open Access; 2012.

Chapter 12
Application of an Adaptive Analytical Characterization Strategy to Support Development and Approval of Biosimilars



John P. Gabrielson, Jared A. Young, and Brent S. Kendrick

Abstract Availability of biosimilar drugs in the United States and other regions of the world plays an important role in decreasing drug prices and increasing access to life-enhancing therapies. To ensure safety, efficacy, and bioequivalence of these biologically-derived products, their commercialization pathway is governed by stringent, and sometimes region-specific, clinical and quality requirements. Biosimilars represent a unique class of biopharmaceuticals in which the clinical commercialization pathway is streamlined, but the stringency of analytical characterization remains high. Comprehensive analytical characterization is required to demonstrate analytical similarity of numerous quality attributes between the biosimilar and reference products. In this chapter, the authors present an overarching, yet adaptive, analytical characterization strategy intended to satisfy world-wide regulatory expectations for approval of protein-based biosimilars in key regions of the world with known requirements for marketing authorization. The authors then demonstrate how analytical characterization methods may be managed within a broad and general lifecycle characterization framework. It is the authors' belief that a carefully designed characterization strategy, such as the one recommended in this chapter, will enable more rapid development of high-quality biosimilar drugs and aid in successfully bringing them to market.

Keywords Analytical characterization methods · Protein characterization · Reference product characterization · Biosimilar product characterization · Adaptive characterization strategy · Lifecycle management

J. P. Gabrielson (✉) · J. A. Young · B. S. Kendrick
Elion Labs, KBI Biopharma, Inc., Louisville, CO, USA
e-mail: jgabriel@elionlabs.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_12

Introduction

Protein-based pharmaceutical products are naturally complex and heterogeneous. After being produced from living cells, the protein of interest is extensively purified through multiple downstream processing steps. Even so, the finished product contains a mixture of related forms (isoforms) of the protein molecule, which may or may not be equally safe and efficacious (Harris et al. 2010). Due to the complexity of the isoforms, which can differ in many ways including hydrophobicity, charge state and aggregation state, a variety of analytical characterization test methods are needed to assure product quality and, in the case of biosimilars, to maximize the probability of matching the reference product's isoform distribution (Tsuruta et al. 2015). This extensive and detailed analytical characterization of protein-based pharmaceutical products is required during early development of the drug, during clinical trials, as manufacturing changes are made, to support licensure of the product, and when post-marketing changes are made (42 U.S.C 2017; U.S. Department of Health and Human Services, Food and Drug Administration 2015; International Conference on Harmonization 2009a; European Medicines Agency 2014; International Conference on Harmonization 2009b).

Regulatory guidelines govern product characterization requirements for biologics. FDA Guidance for Industry M4Q (U.S. Department of Health and Human Services, Food and Drug Administration 2018a) states "for desired product and product related substances, details should be provided on primary, secondary and higher-order structure; posttranslational forms; biological activity; purity; and immunological properties." This is foundational to the quality by design (QbD) principles described in ICH Q8(R2) (International Conference on Harmonization 2009c), Q9 (International Conference on Harmonization 2009d) and Q10 (International Conference on Harmonization 2009e). Specifically, product characterization enables identification of critical quality attributes (CQAs), creation of a quality target product profile (QTPP), and a risk-based approach leveraging new and prior knowledge in pharmaceutical development as suggested in ICH Q8(R2). This, in turn, helps drug developers (1) identify the appropriate analytical procedures for the product's control strategy, (2) ensure that the manufacturing process operates within a state of control, and (3) assess raw material and process sources of variation affecting process performance and product quality.

As described elsewhere in this book, development and licensure of biosimilar products, i.e., biologics intended to mimic an approved biologic, are regulated differently by health authorities than their non-biosimilar counter-parts. The FDA, for example, now allows biosimilars to be approved via a 351(k) licensure pathway under the Public Health Service Act as specified in the Biologics Price Competition and Innovation Act (Biologics Price Competition and Innovation Act 2018). This abbreviated approval pathway, which lessens clinical trial requirements without omitting them, is meant to encourage development of biosimilars by focusing approval requirements on demonstration of nonclinical, clinical, and analytical equivalence of the biosimilar to a reference product. Although clinical requirements

are streamlined, the analytical characterization requirements to license a biosimilar are just as stringent, and potentially more so, than those applied to license a novel protein product submitted through a traditional approval pathway. Nevertheless, whether biosimilar or not, the protein attributes which must be studied are the same—primary structure, posttranslational modifications, secondary structure, tertiary structure, thermal stability, quaternary structure and aggregation, impurity profile, degradation pathways, and biological mechanisms of action.

In this chapter, we provide a roadmap to comprehensive characterization of biosimilars. First and foremost, a biosimilar developer's strategy should be founded on the requirements for in-depth characterization of protein products. Second, the strategy should entail distinct but related characterization plans for the biosimilar and its reference product, and third, it should be adaptive. No single characterization strategy is appropriate for every molecule; moreover, once established, the characterization plan should be refined during all stages of biosimilar development as the sponsor learns more about the biosimilar and reference products, interacts with regulators, and develops the analytical tools needed to fully characterize the biosimilar and its reference.

Analytical Characterization of Protein Attributes

The characterization plan applied to a biosimilar product should align with the regulatory guidance in place for protein product characterization generally, and it should be based primarily on well-developed scientific considerations. As discussed in this section, analytical characterization of biosimilars diverges only to the extent that the strategy should be based heavily on the biosimilar developer's understanding of the reference product. Moreover, as with any protein-based pharmaceutical product, the analytical characterization approach applied to biosimilars should follow an integrated control strategy and be adapted to the features of the molecule, its manufacturing process, mode of delivery, mode of storage, stage of development, and the analytical procedures available to study it (Vandekerckhove et al. 2018; Flynn and Nyberg 2014; Schenerman et al. 2004).

Protein Product Characterization

Unlike chemically-synthesized small molecule drugs, protein-based pharmaceutical products are produced in biological systems. These cell-based protein production systems are used to synthesize highly complex protein structures. The manufacturing of the protein also yields undesired byproducts such as host cell proteins, process reagents, residual DNA, and natural variations of the intended protein product (Liu et al. 2010). Process-related impurities, such as residual DNA and host cell proteins, are removed from the final product during downstream purification

Table 12.1 Product attributes for characterization

Product attribute	Purpose
Primary structure, disulfide structure and glycan structure	Ensure fidelity of DNA transcription to expected amino acid sequence
	Elucidate post-translational modifications, glycan structures, and disulfide structures
Structural product variants, including misincorporations and other amino acid modifications, glycan variants, disulfide variants, hydrophobic, and charge variants	Detect, identify, and quantify product-related variants present in drug substance and drug product
	Assess criticality of product attributes by studying relative potency of product variants
Size variants, including low molecular weight fragments and high molecular weight aggregates	Determine nature of size variants, including whether they are covalent or non-covalent, reducible or non-reducible
Subvisible and visible particles	Guide the overall analytical control strategy for the product
Secondary structure, tertiary structure, and thermal stability	Ensure the intended protein and related protein variants are properly folded and have sufficient conformational stability
Degradation pathways	Elucidate product susceptibility to degradation under relevant stress conditions
	Inform analytical and stability control strategies
Biological mechanism(s) of action	Elucidate biological mechanism(s) of action and justify bioassay methodologies
	Assess attribute criticality based on relative potency and/or receptor binding characteristics of purified product-related variants

steps; for product approval, adequate removal of process-related impurities must be demonstrated through multiple logs of process clearance (International Conference on Harmonization 2009a; Shukla et al. 2017; Chon and Zarbis-Papastoitsis 2011). Product-related impurities and product-related substances are often more difficult to remove during drug substance purification because they may be in equilibrium with other forms of the product, and they may exhibit similar physicochemical characteristics resulting in co-purification. As such, protein-based pharmaceutical products inevitably contain a heterogeneous mixture of naturally occurring variations of structures and closely related forms, which should be understood by the drug developer, and which together are considered constituents of the final product. Product attributes and structural variants that typically need to be characterized include those listed in Table 12.1 (Schenerman et al. 2004).

The product attributes listed in Table 12.1 should be assessed for elucidation of protein structure and determination of the product's CQAs, whether or not the product is being developed as a biosimilar. Criticality of each product attribute can be assessed through a variety of approaches to characterize its biological relevance and/or structure-function properties (Goetze et al. 2010; Liu et al. 2009; Arthur

et al. 2015). Understanding the attribute criticality serves as an important starting point from which to develop a comprehensive, product-specific, and attribute-focused characterization plan. Adopting an attribute-focused strategy, rather than a procedure-focused strategy, correctly places the focus of analytical characterization on the product. As discussed later in this chapter, the task of identifying, developing, and appropriately validating analytical procedures and an integrated process control strategy becomes easier and much more efficient when the product's critical attributes and non-critical attributes are understood and assessed for impact to the product's quality, safety, and efficacy.

In the case of biosimilar products—those products designed to closely match a reference product and therefore allowed to follow an abbreviated licensure pathway—the analytical characterization strategy should be developed broadly to include the reference product in addition to the biosimilar itself. In fact, much of the characterization work should be duplicated, though not concurrently, to ensure that the biosimilar is in fact analytically similar to its reference.

Reference Product Characterization

Early in the development of a biosimilar product, sponsors should characterize the reference product, which is sourced in its commercial drug product configuration. The advantages of reference product characterization early in the biosimilar development lifecycle are many: early definition of the CQAs of the reference material, and by extension, the likely CQAs of the biosimilar; establishment of a QTPP from which the analytical control strategy can be derived; and development of analytical procedures to measure the CQAs. Ultimately, process and product development and characterization will be more directed when the sponsor better understands the target. Reference product characterization is therefore an integral piece of the biosimilar developer's overall characterization plan.

Fundamentally, the goal of biosimilar development is to ensure that the biosimilar product is highly similar to its reference product for every measured attribute, although the allowable difference between the biosimilar and reference product depends on the criticality of that attribute (i.e., biosimilarity tier) (U.S. Department of Health and Human Services, Food and Drug Administration 2018b). Although some attributes are more critical than others, it is generally in the biosimilar sponsor's interest to measure and characterize numerous product attributes of both the reference and biosimilar products. As more attributes are characterized, less residual uncertainty exists between the products. It is not possible to achieve the goal of 'highly similar' without thoroughly characterizing the reference product alongside the biosimilar.

Most in-depth characterization work (e.g., amino acid sequence confirmation, impurity characterization, etc.) can be performed on one representative reference product lot. However, for product attributes that may vary between lots, even when the manufacturing process is well controlled, the characterization plan needs to

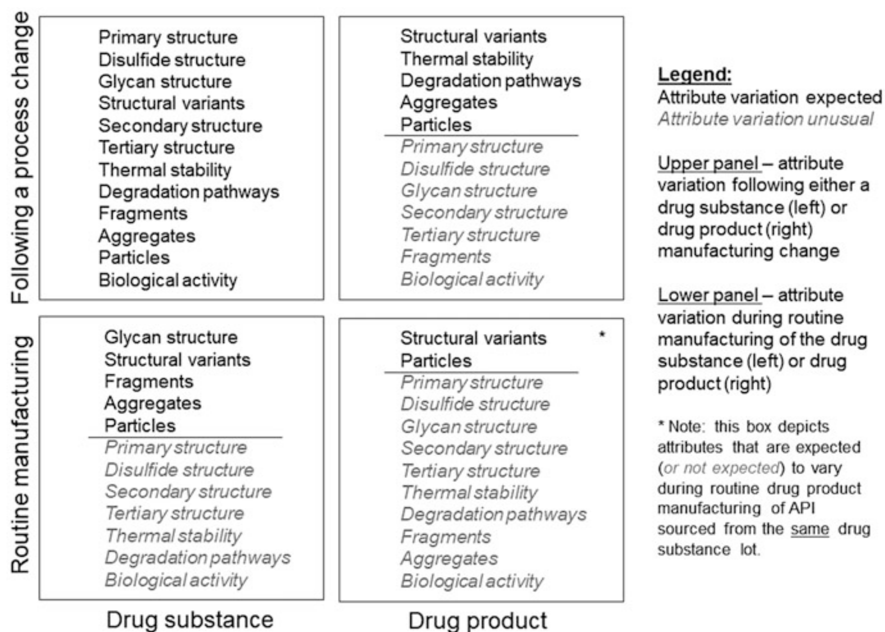


Fig. 12.1 Protein attribute-process risk classification

include at least three, and preferably six, unique lots. For purposes of formal biosimilarity testing, the number of reference product lots is much higher, often 30 or more, and should be determined on an attribute-specific basis depending on how critical the attribute is (its biosimilarity tier), and to what degree the attribute varies across lots. A risk assessment will help inform the biosimilar sponsor's analytical characterization strategy. A protein attribute and process risk classification scheme based on product attributes expected to change, versus those not expected to change, between manufacturing lots and/or processes provides a good foundation for conducting such a risk assessment. An example of this type of attribute-process risk classification is provided in Fig. 12.1.

The attribute classification diagram in Fig. 12.1 conveys only one example of how attributes may vary for a hypothetical monoclonal antibody product. Each protein product is different and should be assessed individually to identify which attributes are controlled at which manufacturing steps. Nonetheless, the figure is instructive in that most protein product attributes display significantly more variation during drug substance manufacturing (particularly upstream), than they do during drug product manufacturing. One implication of this is that drug developers should give some consideration to the alignment of the product attribute with the appropriate sample type for characterization. For example, it may be most informative to characterize glycan structures of the product at drug substance stage,

whereas characterization of subvisible particles should generally be performed using the drug product in its final container.

When reference product lots are sourced from the market, biosimilar sponsors should be aware that not all lots are statistically independent. Using Fig. 12.1 as an example, many protein attributes vary little, or not at all, between drug substance lots, whereas some attributes vary considerably between drug substance lots. Few attributes change between drug product lots if they are filled from the same drug substance. When developing a biosimilar, therefore, the sponsor should recognize that the reference product lots being characterized may not be fully independent, particularly lots which are obtained at similar times with similar expiration dates. Rather than assuming independence when the lots are likely correlated, the sponsor is better served using the data it collects while characterizing the reference product to reverse engineer the lot parentage. Once determined, the correlation among lots may be statistically modeled in the formal biosimilarity evaluation.

In summary, reference product characterization serves the following purposes as a component of the biosimilar sponsor's overall characterization plan:

1. The reference product serves as a target early in the development of the biosimilar for purposes of determining CQAs, QTPP, and clinical specifications;
2. Initial characterization results for the reference product may inform the biosimilar characterization plan; and
3. The reference product provides early clues about the anticipated level, criticality, and variation of product attributes and how those product attributes need to be controlled.

Biosimilar Product Characterization

As with the reference product, characterization of a biosimilar product should be conducted in a series of stages. However, unlike the reference product, the extent and timing of biosimilar characterization is dictated by the sponsor's development program, e.g., availability of representative material, initiation of clinical trials and related regulatory submissions, and commercial process development. Thus, while the biosimilar sponsor's characterization of a reference product typically spans many years, governed by availability of statistically independent lots, comprehensive reference product characterization that focuses on identifying the criticality of product attributes can be performed at any time with any reference product lot. Conversely, for the biosimilar product, a staged characterization approach may be warranted, one which grows in scope and complexity as the biosimilar program matures. The complex interplay between reference product characterization and biosimilar characterization is depicted in Fig. 12.2.

During the biosimilar development lifecycle, it is convenient and efficient to utilize the same characterization methods, apply the same characterization strategy, and characterize the same attributes of the biosimilar and reference products. Thus,

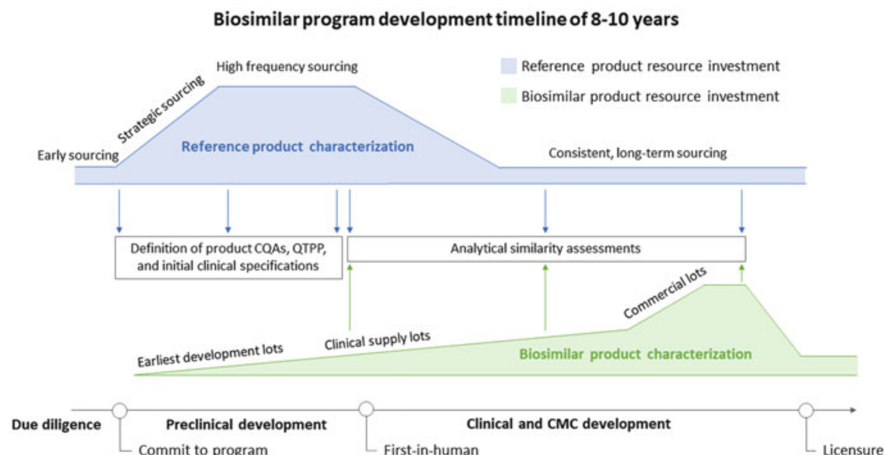


Fig. 12.2 Progressive and adaptive reference product and biosimilar product characterization strategies throughout biosimilar program development

although characterization of each product separately is important, comparisons between the two must be at the heart of the characterization plan, as shown in Fig. 12.2. Analytical characterization results obtained for the biosimilar should be similar to results of the reference product in every way possible. Differences must be justified, and ultimately, the data for both the biosimilar and reference product should be included in the regulatory application for licensure.

Adequate demonstration of analytical similarity is a foundational element in the overall assessment of biosimilarity. Although unlikely in the near future in our view, it is possible that comprehensive and detailed analytical characterization of the biosimilar and reference products, without clinical trials, may eventually be sufficient for licensure of biosimilars, but only if both products are assessed by numerous highly resolving analytical methods sufficient to minimize residual uncertainty, and even then only if the analytical similarity evaluation results in nearly indistinguishable products, achieving so called 'fingerprint-like similarity' (U.S. Department of Health and Human Services, Food and Drug Administration 2015; U.S. Department of Health and Human Services, Food and Drug Administration 2016). Given the structural complexity and natural heterogeneity of protein pharmaceutical products, fingerprint-like similarity is a very high bar, which is not likely achievable with current process and analytical technology.

Often, early development lots of the biosimilar product are characterized to support the initial regulatory submission which enables clinical trials. These initial data are useful as they begin a progression of information that culminates with the data package submitted for licensure. Between those two regulatory applications, the sponsor should methodically increase its understanding of its biosimilar product, during process improvements, scale-up, formulation studies, and long-term stability

studies. As part of its development program, the sponsor must perform both comparability studies of its own product, particularly from clinical material to commercial material, along with similarity evaluation of the biosimilar to the reference product at multiple points in the development lifecycle (International Conference on Harmonization 2009b).

Adaptive Characterization Strategy

It is not possible to recommend one characterization strategy that may be applied successfully across the many and varied biosimilars currently being developed. Instead, we provide a framework with which to create a strategy that is comprehensive and yet focused on the needs of the specific biosimilar program. Such an *adaptive* strategy will take different forms for different products, and we have touched on some of these considerations already: a product quality attribute assessment (PQAA) is used to define which attributes to study, and with what degree of rigor; the stage of development should inform the degree of characterization needed to ensure that the sponsor's understanding of its biosimilar product increases over time as the program matures; and the characterization plan for the biosimilar may change as the sponsor learns more about the reference product and incorporates that knowledge into the biosimilar program, as shown in Fig. 12.2. Thus, while the characterization strategy itself necessarily differs for different products, the pillars of an adaptive characterization strategy may be defined by answering the following questions:

1. Is the characterization plan aligned with the product's unique risk profile, including its specific product attributes, intended patient population, and anticipated administration and dosing?
2. Does the characterization plan allow the sponsor to move quickly to respond to program changes, availability of new analytical procedures, and new learnings from the reference product?

By defining the product's QTPP early on, performing a thorough PQAA, understanding critical patient factors, and utilizing a progression-of-information approach to incorporate reference product understanding into the biosimilar characterization plan, the sponsor will achieve a biosimilar characterization strategy acceptable to health authorities. Ultimately, the approval readiness of a biosimilar depends less on the strategy undertaken to characterize it and more on the data package produced by executing that strategy.

Lifecycle Management of Analytical Characterization Methods

Our focus in this chapter is on protein attributes, specifically how biosimilar sponsors must develop a product characterization plan that aims, primarily, to identify and characterize the protein attributes that are most relevant to the specific biosimilar product, using the reference product as a guide. Selecting appropriate analytical procedures to characterize the reference and biosimilar products is also important, but this is of secondary importance to identifying and characterizing the attributes themselves. An attribute that does not appreciably impact the product's quality, safety, or efficacy does not become important simply because it can be measured easily and with high sensitivity, just as our inability to measure a critical attribute does not make that attribute less critical.

Analytical characterization procedures should be selected for a purpose and then shown to achieve that purpose. Although this may seem overly simplified, the fundamental principles of analytical method lifecycle management described in ICH Q2(R1) (ICH Harmonized Tripartite Guideline 2005; United States Pharmacopeia 2017) and USP <1220> (United States Pharmacopeia 2017), among other sources, apply not only to quality control analytical procedures but also to the analytical procedures used for protein product characterization. In fact, by applying a lifecycle management approach to analytical characterization procedures, biosimilar sponsors can ensure that they are, in fact, selecting and using the appropriate tools for the job. Only then will biosimilar sponsors have the assurance that they have correctly placed focus on the most critical protein attributes.

Martin et al. (2018) present an approach for analytical method lifecycle management using concepts consistent with Quality by Design (QbD), ICH Q8 (2009c), and the FDA Process Validation Guidance (2011). The performance requirements of any analytical procedure are defined by the analytical target profile (ATP). Defining an ATP is an important (and often overlooked) first step in analytical method lifecycle management, and it defines the analyte or attribute to be measured, the concentration range, procedure performance criteria, and product specifications, if applicable (Barnett et al. 2016). The ATP is established primarily to define the purpose of the analytical procedure.

Glycosylation provides a useful example of why establishing an ATP is important. In years past, a drug developer may have spent very little time identifying the attribute to be measured and defining the ATP for the procedure in advance of procedure development. Instead, a scientist may have immediately launched into traditional method development work, such as establishing sample preparation conditions, selecting a resolving column, and optimizing the chromatography with no significant regard to which chromatographic peaks are important to optimize. However, for some products, the purpose of glycan characterization is to identify or quantify specific, potentially immunogenic, glycans and verify that they are not present in the product or exist at sufficiently low levels (i.e., a limit test for specific glycan species), whereas in other cases, the purpose of glycan characterization is to

assess the similarity of the overall glycan profile between a reference product and a biosimilar. If the method purpose and ATP are not defined at the outset, the scientist risks significant rework, or worse, developing a method that is not suitable for its purpose.

To be fit for its intended use, the analytical procedure should produce reportable results that meet the requirements of whatever decision is intended to be made by the procedure. In the case of analytical characterization of biosimilars, the decision criterion is usually quite simple: to definitively demonstrate, statistically or otherwise, that the attribute as it exists in the biosimilar product is sufficiently similar, or not inferior, to the same attribute in the reference product. It follows, then, that the analytical characterization procedures used to measure attributes of biosimilar products should be sufficiently precise and able to detect changes in the attribute with sufficient sensitivity to enable a determination of similarity, or lack thereof. Examples of approaches to assessing method precision and similarity-quantitation limits for biophysical methods are covered in Teska et al. (2013) and Dinh et al. (2014).

An analytical procedure can be considered to have a lifecycle with three stages: (1) procedure design, (2) procedure performance qualification, and (3) continued procedure performance verification. The first stage includes selecting an appropriate technology, deciding upon the method parameters, and identifying potential sources of variation. The second stage demonstrates that the analytical procedure is fit for its intended use. This process is described in USP <1225> as method validation, but in the new lifecycle paradigm, experiments are used to demonstrate that the procedure output, and its uncertainty, meet the requirements specified in the ATP. Finally, stage 3 of the lifecycle provides ongoing assurance that the analytical procedure continues to operate as required. This includes activities such as result trending during routine use of the procedure, transfer of the procedure between laboratories, and application of a procedure control strategy. One important element of analytical lifecycle management is that it is cyclical, not linear. An effective procedure monitoring process will naturally lead to re-design of the procedure.

In the context of analytical characterization of biosimilars, lifecycle management of analytical procedures has several important implications. First, applying the principles of method lifecycle management to characterization methods will inevitably force the biosimilar sponsor to consider the specific attributes of its product, their criticality, and why those attributes must be monitored and controlled. Establishing an ATP and then designing the procedure with the ATP in mind must be accompanied by an understanding of the product attributes themselves. Second, the cyclical nature of method lifecycle management is well suited to the maturation process of a biosimilar program. Early in development, a first iteration of analytical characterization procedures may be deployed; as the sponsor learns more about its product and its methods (through stage 2 and stage 3), the procedures can be re-designed and implemented for improved characterization at later stages of development, e.g., for licensure-enabling characterization. Finally, without demonstration of a procedure's fitness for use, the biosimilar sponsor is ill-equipped to evaluate analytical similarity. It is not possible to conclude that

the biosimilar and reference products are analytically similar without knowing the performance capabilities of the procedures used to characterize the products—namely, their precision and how sensitive they are to changes in the quantity and state of the measured attribute(s).

As a final example, we present in Table 12.2 a list of analytical characterization procedures for detailed characterization of a biosimilar monoclonal antibody. The table of attributes is based on the authors' experience characterizing both original and biosimilar antibody-based products, with reflection on the procedures used to characterize FDA-approved biosimilar antibody products (infliximab-dyyb, adalimumab-atto, bevacizumab-awwb, and trastuzumab-dkst). As discussed throughout this chapter, it is imperative that an analytical characterization plan include more than a list of procedures, like the one listed in Table 12.2. At its core, the plan should be attribute-centric, risk-based, and supported by analytical procedures managed through a continuous improvement lifecycle.

Future Outlook

In our view, the need for and extent of characterization of biosimilars will continue to increase in coming years. Many health authorities throughout the world, notably the FDA, have created approval pathways for biosimilar products to increase drug availability and reduce drug prices. In this environment, we believe health authorities will continue to look for opportunities to streamline the approval requirements for biosimilar products, while still ensuring that they approve safe and efficacious therapies which sufficiently match the intended reference product. Analytical characterization will remain a foundational component, and focal point, of the overall biosimilarity demonstration needed for approval of biosimilars, and while clinical requirements may be relaxed over time, it is our expectation that analytical requirements will continue to increase.

To successfully bring biosimilars to market, sponsors should invest significantly in reference product characterization as a guidepost for developing an appropriate characterization strategy for their biosimilar product. With a well-developed reference product characterization plan in place, biosimilar sponsors can more effectively develop an adaptive characterization strategy for their own biosimilar product, one that evolves over time as new information is revealed about both products, and one that depends on the stage of development, target patient population, and most importantly, the attributes of the product. The selection of analytical characterization procedures plays an important role in this strategy: by applying lifecycle management principles to their analytical characterization procedures, biosimilar sponsors will generate the supporting data they need to justify analytical similarity.

Table 12.2 Analytical procedures to characterize monoclonal antibody biosimilar products

Product attribute	Analytical characterization procedures[a,b]
Primary structure	• Intact molecular mass (LC-MS) • Reduced and deglycosylated molecular mass (LC-MS) • Peptide mapping with UV and MS detection (reduced and non-reduced) • Peptide mapping with MS/MS sequence confirmation • Amino acid analysis or Edelhoch method for extinction coefficient
Disulfide structure	• Non-reduced and reduced peptide mapping LC-MS • Intact RP-LC
Glycan structure	• HILIC-LC with FLD and MS detection of cleaved glycans • *Sequential endo-/exo-glycosidase treatment to confirm specific glycan structures, including possible co-eluting isobaric glycans*
Structural variants (charge, hydrophobic, etc.)	• Analytical scale or preparative purification by IEX-LC and RP-LC • Analysis of purified peaks by relevant purity methods and peptide mapping
Secondary structure	• FT-IR spectroscopy • Far UV CD spectroscopy • *Raman spectroscopy*
Tertiary structure	• Near UV CD spectroscopy • NMR spectroscopy (1D and 2D methods) • *Intrinsic fluorescence spectroscopy*
Thermal/conformational stability	• Differential scanning calorimetry
Degradation pathways	• Thermal forced degradation and accelerated stability • Photo-stability • pH degradation (high pH, low pH, physiological pH) • Forced oxidation (hydrogen peroxide, *tert-butyl peroxide*, AAPH) • *Agitation and mechanical stress*

(continued)

Table 12.2 (continued)

Product attribute	Analytical characterization procedures[a,b]
Size variants (fragments, aggregates, etc.)	• SE-HPLC-MALS • Sedimentation velocity AUC • CE-SDS (reduced and non-reduced) • Analytical scale or preparative purification by SEC and denaturing SEC • Analysis of purified peaks by relevant purity methods, size characterization methods, and peptide mapping • *Asymmetrical flow FFF* • *Dynamic light scattering*
Particles (subvisible, visible)	• Visual inspection • Dynamic flow imaging analysis • Light obscuration • Nanoparticle tracking analysis and/or resonant mass measurement • G3-ID with imaging and Raman spectroscopy
Biological activity(ies)	• Cell-based in vitro bioassay (orthogonal to release method) • Binding-based relative potency analysis • ADCC and CDC cell based assays (MoA-dependent) • FcγRIII binding assay (MoA-dependent) • FcRn binding assay (MoA-dependent)

[a]The analytical characterization procedures included in this table are a composite list derived partially from the authors' experience and largely from the analytical procedures used for Inflectra® (infliximab-dyyb), Amgevita® (adalimumab-atto), Mvasi® (bevacizumab-awwb and Ogivri® (trastuzumab-dkst) as also shown in Table 22.4 of Chap. 22

[b]Italicized gray text is used to signify analytical procedures which may not be necessary, depending on product-specific characteristics, or those that may provide limited incremental value when performed in addition to other procedures that evaluate the same attribute

Declaration of Interest The authors are employees of KBI Biopharma, Inc., which serves as a contract development and manufacturing organization in the biopharmaceutical industry, including contract analytical testing and manufacturing of biosimilars.

References

42 U.S.C. 262 - Regulation of biological products. Available at: https://www.gpo.gov/fdsys/granule/USCODE-2010-title42/USCODE-2010-title42-chap6A-subchapII-partF-subpart1-sec262. Accessed 5 Sept 2017.

Arthur KK, Dinh N, Gabrielson JP. Technical decision making with higher order structure data: utilization of differential scanning calorimetry to elucidate critical protein structural changes resulting from oxidation. J Pharm Sci. 2015;104:1548–54.

Barnett K, et al. Analytical target profile: structure and application throughout the analytical lifecycle. Pharmacopeial Forum. 2016;42:1–15.

Biologics Price Competition and Innovation Act of 2009. Available at: https://www.fda.gov/downloads/drugs/ucm216146.pdf. Accessed 22 May 2018.

Chon JH, Zarbis-Papastoitsis G. Advances in the production and downstream processing of antibodies. New Biotechnol. 2011;28:458–63.

Dinh NN, Winn BC, Arthur KK, Gabrielson JP. Quantitative spectral comparison by weighted spectral difference for protein higher order structure confirmation. Anal Biochem. 2014;464:60–2.

European Medicines Agency. Guideline on similar biological medicinal products: CHMP 437/04 Rev 1. 2014. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176768.pdf. Accessed 22 May 2018.

FDA. Guidance for industry: process validation: general principles and practices. 2011. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070336.pdf.

Flynn GC, Nyberg GB. Using quality by design principles in setting a control strategy for product quality attributes. In: State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization, Monoclonal antibody therapeutics, vol. 1. Washington, D.C.: ACS Publications; 2014.

Goetze AM, Schenauer MR, Flynn GC. Assessing monoclonal antibody product quality attribute criticality through clinical studies. MAbs. 2010;2:500–7.

Harris RJ, et al. Analytical characterization of monoclonal antibodies: linking structure to function. In: Current trends in monoclonal antibody development and manufacturing. New York: Springer; 2010. p. 193–205. https://doi.org/10.1007/978-0-387-76643-0_12.

ICH Harmonized Tripartite Guideline. Validation of analytical procedures: text and methodology ICH Q2(R1). 2005.

International Conference on Harmonization. ICH Q6B: specifications: test procedures and acceptance criteria for biotechnological/biological products. 2009a. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q6B/Step4/Q6B_Guideline.pdf. Accessed 22 May 2018.

International Conference on Harmonization. ICH Q5E: comparability of biotechnological/biological products subject to changes in their manufacturing process. 2009b. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q5E/Step4/Q5E_Guideline.pdf. Accessed 22 May 2018.

International Conference on Harmonization. ICH Q8(R2): pharmaceutical development. 2009c. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf. Accessed 22 May 2018.

International Conference on Harmonization. ICH Q9: quality risk management. 2009d. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q9/Step4/Q9_Guideline.pdf. Accessed 22 May 2018.
International Conference on Harmonization. ICH Q10: pharmaceutical quality system. 2009e. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q10/Step4/Q10_Guideline.pdf. Accessed 22 May 2018.
Liu HF, Ma J, Winter C, Bayer R. Recovery and purification process development for monoclonal antibody production. MAbs. 2010;2:480–99.
Liu YD, van Enk JZ, Flynn GC. Human antibody Fc deamidation in vivo. Biologicals. 2009;37:313–22.
Martin GP, et al. Lifecycle management of analytical procedures: method development, procedure performance qualification, and procedure performance verification. Available at http://www.uspnf.com/sites/default/files/usp_pdf/EN/USPNF/revisions/lifecycle_pdf.pdf. Accessed 22 May 2018.
Schenerman MA, et al. CMC strategy forum report: analysis and structure characterization of monoclonal antibodies. BioProcess Int. 2004;2(2):42–51.
Shukla AA, Wolfe LS, Mostafa SS, Norman C. Evolving trends in mAb production processes. Bioeng Transl Med. 2017;2:58–69.
Teska BM, et al. Comparison of quantitative spectral similarity analysis methods for protein higher-order structure confirmation. Anal Biochem. 2013;434:153–65.
Tsuruta LR, Lopes dos Santos M, Moro AM. Biosimilars advancements: moving on to the future. Biotechnol Prog. 2015;31:1139–49.
U.S. Department of Health and Human Services, Food and Drug Administration. Quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. 2015. Available at: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf. Accessed 22 May 2018.
U.S. Department of Health and Human Services, Food and Drug Administration. Clinical pharmacology data to support a demonstration of biosimilarity to a reference product: guidance for industry. 2016. Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm397017.pdf. Accessed 22 May 2018.
U.S. Department of Health and Human Services, Food and Drug Administration. M4Q: The CTD - quality. Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073280.pdf. Accessed 22 May 2018a.
U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for industry: statistical approaches to evaluate analytical similarity. Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm576786.pdf. Accessed 22 May 2018b.
United States Pharmacopeia. USP <1220> The analytical procedure lifecycle. 2017. 9.
Vandekerckhove K, et al. Rational selection, criticality assessment, and tiering of quality attributes and test methods for analytical similarity evaluation of biosimilars. AAPS J. 2018;20:68.

Chapter 13
Higher Order Structure Methods for Similarity Assessment

Jared A. Young and John P. Gabrielson

Abstract Guidance from the EMA and FDA suggest a stepwise approach for assessing biosimilarity, an approach that leverages both structural and functional characterization of the biosimilar product to define appropriately-sized non-clinical and clinical studies. Because higher order structure (HOS) dictates protein function and stability, HOS is a key product quality attribute for which demonstration of analytical similarity is essential; by extension, characterization of the HOS of a protein biosimilar can aid in reducing residual uncertainty and informing appropriately sized non-clinical and clinical studies. A review of seven biosimilar advisory committee briefing documents showed a wide range of diversity in HOS methods utilized for similarity assessment to date. No correlation was observed between the types of methods selected, the number of methods, the number of reference product lots characterized, or the subsequent non-clinical or clinical study designs. The diversity in method selection appears to arise from two factors: the range of opinions across the industry on the ability of HOS methods to inform technical decisions, and the regulatory risk tolerance of different organizations. These two factors inform an organization's overall HOS similarity strategy, and each organization must balance speed, sensitivity, specificity, and cost to select the HOS characterization methods it applies to the similarity exercise. We recommend a quantitative approach for HOS method selection and analytical similarity study design. Qualifying HOS methods provides a quantitative measure of method sensitivity and specificity to better inform a method ranking process from which appropriate methods may be selected. These methods should then be applied with appropriate lot selection and with a sufficient number of lots, emphasizing trends over time in the reference product material. Quantitative assessment of method sensitivity and specificity combined with appropriate lot selection provides objective measures to reduce residual uncertainty and better inform the analytical similarity conclusion.

J. A. Young (✉) · J. P. Gabrielson
Elion Labs, KBI Biopharma, Inc., Louisville, CO, USA
e-mail: jyoung@elionlabs.com; jgabriel@elionlabs.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_13

Keywords Higher order structure · Characterization · Method selection · Similarity · Method strategy · Circular Dichroism · Fourier Transform Infrared Spectroscopy · Differential Scanning Calorimetry · Nuclear Magnetic Resonance · X-ray crystallography · Intrinsic Fluorescence

Introduction

The function and stability of biopharmaceutical protein products are dictated by the complex three-dimensional protein structure; thus, similarity assessment of higher order structure (secondary, tertiary and quaternary structure) is an essential component of analytical similarity and is an important precursor to defining appropriately sized non-clinical and clinical studies. Guidance from the EMA and FDA both suggest a stepwise biosimilar approach which leverages structural and functional characterization of the product to define the required scope of non-clinical studies and clinical studies (FDA guidance for industry 2015; Guideline on Similar Biological Medicinal Products 2005; Guidance for Industry 2015). "The more comprehensive and robust the comparative structural and functional characterization—the extent to which these studies are able to identify (qualitatively or quantitatively) differences in relevant product attributes between the proposed product and the reference product (including the drug substance, excipients, and impurities)—the more useful such characterization will be in determining what additional studies may be needed" (FDA guidance for industry 2015). Per EMA and FDA guidance, a link between the extent of analytical characterization and non-clinical and clinical studies is expected. Robust and comprehensive demonstration of attribute similarity, including HOS, reduces residual uncertainty, and if the products are judged to be highly similar, such characterization should justify reduction in the scope of subsequent studies. To date, however, based on publicly available information, no measurable correlation has been observed between the HOS strategies used and the resulting non-clinical and clinical trial designs.

Current Industry Landscape

A diverse range of strategies have been deployed across the biopharmaceutical industry to demonstrate similarity in the higher order structure (HOS) of biosimilar products compared to their intended reference products. We reviewed publicly available advisory committee briefing documents for seven separate products filed between January 2015 and July 2017 and compiled the list of methods used for analytical similarity assessment of higher order structure, as summarized in Table 13.1 (Celltrion 2016; Hospira 2017; Sandoz 2016; Amgen 2016; Amgen 2017; Sandoz 2015; Mylan 2017). All companies used between 3 and 6 analytical methods for higher order structure assessment (excluding disulfide structure and

Table 13.1 Summary of methods utilized for seven different biosimilar products from five different companies

Method	Number of biosimilars
Near UV CD	7
FTIR	6
DSC	6
Far UV CD	5
1-D NMR	3
2-D NMR	2
Intrinsic fluorescence	2
X-ray crystallography	1
HDX	1
Antibody array	1

Table 13.2 Comparison of HOS method utilization across two organizations

Method	Organization 1	Organization 2
FTIR	2/2	1/2
Far UV CD		1/2
Near UV CD	2/2	2/2
DSC	2/2	1/2
HDX		1/2
Antibody array		1/2
1-D NMR		2/2
2-D NMR		1/2
X-ray crystallography		1/2

The four products evaluated are all large multi-domain proteins, with similar critical quality attributes. The differences in similarity assessment strategies did not appear to be influenced by protein size or complexity

aggregate characterization). The exact number of lots characterized by each method was typically redacted from public documents.

FTIR, CD and DSC are established core methods, with every company using at least two of these three methods in their analytical similarity data packages. Otherwise, no other clear trends are evident from the publicly available information across the seven products. The diversity in approaches does not appear to arise from modality-tailored strategies, as no clear patterns arose within the same modality, e.g., monoclonal antibodies. Instead, we believe the differences arise at the first step in creating an analytical similarity assessment plan, namely defining the criticality of HOS. HOS criticality was ranked from “high” to “low” by different organizations, and no two organizations employed the same set of HOS methods. Clearly, the relative importance placed on HOS data varies significantly between companies, as evidenced in the briefing documents.

To demonstrate the wide differences between organizations, two organizational strategies are compared based on publicly available ODAC briefing documents. An overview of the methods used by each organization is provided in Table 13.2. Each

organization was developing two biosimilar products, and we reviewed the publicly available information on each product.

Organization 1 employed a streamlined strategy, utilizing the same three methods for both biosimilar programs. FTIR spectroscopy was used to assess secondary structure, and near UV CD spectroscopy was used to assess tertiary structure. DSC provided an orthogonal method to evaluate thermal and conformational stability, which is indicative of changes in both secondary and tertiary structure. These three methods are commonplace in the biopharmaceutical industry, and are often used to demonstrate secondary and tertiary structure similarity with minimal resource requirements. In addition to applying a minimalistic approach for the number and type of HOS methods, Organization 1 analyzed fewer lots than Organization 2. Organization 1 sourced more than 50 lots over 6 years for their evaluation of analytical similarity, but a reduced number of these lots were tested by higher order structure methods. The reduction was justified as higher order structure was considered insensitive to variations in process conditions. The exact number of lots tested by higher order structure methods was not disclosed.

Organization 2 appeared to focus its strategy on attributes rather than specific HOS characterization methods. FTIR spectroscopy was utilized to assess secondary structure for one molecule, and far UV CD spectroscopy was utilized for the other molecule. Near UV CD spectroscopy was utilized to assess tertiary structure for both. High sensitivity fingerprint-like comparisons were conducted using 1-D NMR for both products. At least one highly specific method, 2-D NMR or X-ray crystallography, was used for one lot of each product. Another relatively specific method, HDX or antibody array, was also used to assess reference product and biosimilar product lots. The number of lots tested by each method varied depending on the resource requirements of the method. Relatively high throughput FTIR and CD spectroscopy methods were used to evaluate a high number of lots (>10 US, and >30 EU). Low throughput methods like 2-D NMR and X-ray crystallography were used to study only one or two lots.

The common components of each application by Organization 2 were:

- Secondary structure characterization on a high number of lots [Far UV CD or FTIR],
- Tertiary structure characterization on a high number of lots [Near UV CD],
- High sensitivity fingerprinting on a several lots [1-D NMR],
- High specificity characterization on several lots [HDX or antibody array], and
- Very high specificity characterization on one or two lots [2-D NMR or X ray crystallography].

Industry Perception of HOS Characterization

In addition to differing organizational views about attribute criticality, HOS similarity strategies are influenced by the perceived value of the HOS data resulting

from the similarity assessments. If an organization does not believe HOS methods provide actionable information, it will likely attempt to reduce the amount of HOS characterization it performs. Survey data published in 2015 demonstrates the wide range of viewpoints across the industry regarding the utility of HOS data. There are stark differences in the opinions of self-described subject matter experts (SMEs) and non-experts[1] on the ability to make decisions based solely on HOS data, and the adequacy of current HOS methods. The difference in these opinions can manifest in an organization's HOS similarity assessment strategy in different ways. For example, protein function is dictated by protein conformation, and preserving protein structure is therefore essential to product efficacy and potentially safety. The final risk classification of secondary and tertiary structure quality attributes depends upon the risk assessment model employed, but the potential impact on product safety and efficacy likely precludes HOS from being classified as a low risk attribute. However, some organizations may focus on the perceived lack of sensitivity of HOS characterization methods more than the criticality of the underlying attributes. Similarly, organizations in which HOS SMEs believe sufficient HOS methods exist to inform technical decisions may assign more importance to the analytical similarity results obtained from those HOS methods, while organizations lacking HOS SMEs may view HOS characterization predominately as a regulatory risk mitigation requirement and focus on developing a strategy that mitigates regulatory risk while also minimizing cost.

Impact of HOS Data

From publicly available information, it is not possible to determine if HOS data influenced the clinical study designs for any of the biosimilar products submitted to date. Multiple factors influence clinical and non-clinical studies: characterization data, product indication(s), size of the target patient population, risk tolerance, and the potential to be the first biosimilar approved for a given reference product, among others. However, from the available information, there was very little correlation between the size of clinical trials and the level of detail and extent of the HOS similarity packages. The most streamlined approach we reviewed used three methods for HOS assessment, and a clinical study design of 620 patients to demonstrate equivalence in the primary endpoint (Amgen 2017). The most comprehensive approach used six methods for HOS similarity evaluation, with a focus on high sensitivity and specificity, and a clinical efficacy study design with 774 patients (Sandoz 2016). If any change in clinical trial design resulted from the demonstration of analytical similarity, the effect was not strong enough to measure

---

[1]SME is defined as a person who self-reported spending more than 75% of their time working with HOS data.

across the number of filings we studied. Both the streamlined and the more thorough and detailed HOS similarity strategies were approved by both the EMA and FDA.

The impact of a delay in market authorization, however, can be clearly measured. Biosimilar developer Coherus BioSciences received a Complete Response Letter from the FDA for their pegfilgrastim program in June of 2017, leading to a 27.2% decline in stock price overnight. This represented a $348 million dollar decrease in Coherus' market capitalization, and resulted in a 30% reduction in their workforce Coherus Biosciences, Inc (2017). The impact of a delay in approval has even larger implications when considering pegfilgrastim generates $4.6 billion per year in sales (Amgen Financial Report. 2017). While this Complete Response Letter was not related to the HOS data utilized for demonstration of analytical similarity (Davio 2017), and although regulatory approval of a biosimilar does not always coincide with its market availability, this case study provides an important measure of the value of speed-to-market for biosimilar sponsors. A year of sales has enormous financial ramifications, and the value of risk mitigation dwarfs the cost of higher order structure testing. Regardless of varying organizational opinions about the ability of higher order structure data to inform similarity decisions, any uncertainty in evolving regulatory expectations will likely lead biosimilar sponsors to a risk averse strategy to avoid significant delays.

Method Selection Considerations

While important, we believe risk mitigation is only one factor in how organizations should design their HOS characterization strategies. Several other factors need to be considered when selecting higher order structure methods. Method capabilities, including sensitivity and specificity, must be considered, along with the speed and cost of analysis. These factors are depicted in Fig. 13.1 for well-established HOS methods. Additionally, attribute criticality must also be considered during both method selection and subsequent similarity assessment design.

The severity ranking of quality attributes and subsequent risk assessments should heavily influence both the HOS methods selected, and the number of lots used for analytical similarity assessment. Several risk assessment models have been published and successfully implemented across the industry, resulting in a range of criticality assignments for HOS. Organization 2, for example, has consistently defined both secondary and tertiary structure criticality as "high", which is reflected in both the number of orthogonal methods utilized and the number of reference product lots tested. Other organizations, alternatively, have not placed the same emphasis on HOS characterization, using fewer methods, testing fewer reference product lots, or both. Regardless of the defined criticality for HOS, few organizations have treated any spectroscopic method quantitatively (i.e., higher than a tier 3 attribute in FDA's recently retracted guidance). Although spectroscopic methods are information-rich, they are difficult to translate into meaningful quantitative acceptance criteria. FTIR spectroscopy, for example, is a highly precise method

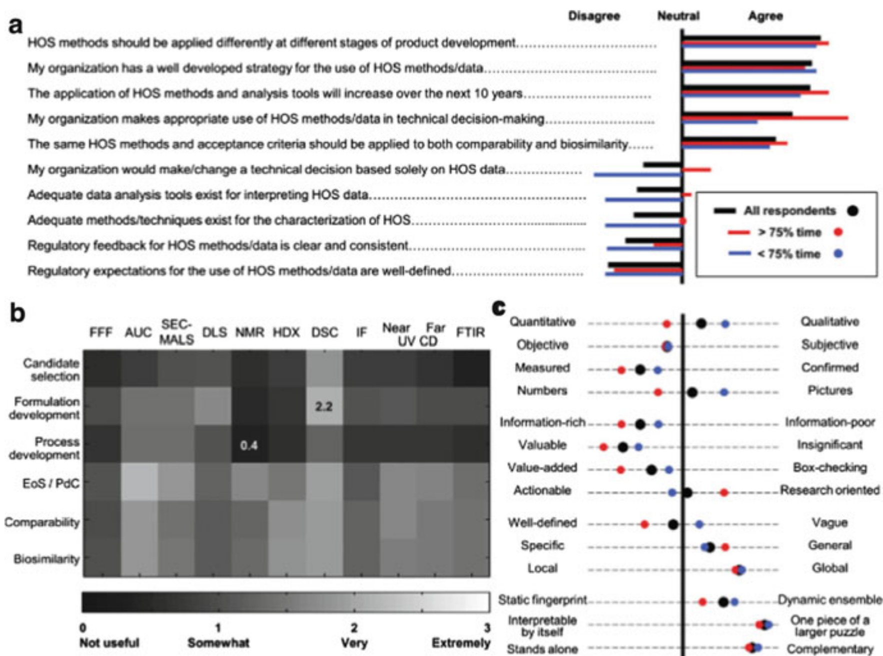


Fig. 13.1 The survey was intended to capture the difference in opinion between non-experts and subject matter experts (SMEs). An SME is defined as a person who self-reported spending more than 75% of their time working with HOS data. 11 SME and 16 nonexperts were surveyed. (**a**) Extent of agreement with 10 statements related to HOS characterization. (**b**) Average usefulness of a selection of HOS characterization techniques for various biopharmaceutical development activities quantitatively ranked. (**c**) Impressions of HOS. Survey respondents were asked to select between two contrasting terms that may or may not be mutually exclusive. Figure was reproduced with consent from Technical Decision-Making with Higher Order Structure Data: Starting a New Dialogue (Weiss and Gabrielson 2015)

generating spectra with thousands of data points that reflect the secondary structure of a given protein. Roughly 10 different overlapping amide I resonance bands have been identified for various secondary structures. Methods for spectral deconvolution have been published, but are typically avoided as the deconvolution algorithms often induce more variability than direct spectral comparison.

As methods increase in sensitivity and specificity, and as more HOS data are collected, the costs required to generate the HOS data package inevitably increase. This correlation between value and cost is illustrated in Fig. 13.2 (FDA guidance for industry 2015).

The relationship between value and cost shown in Fig. 13.3 highlights the diminishing returns associated with generating HOS data for various purposes during development of biologics. This hypothetical relationship correlates well with the industry trends observed for HOS data acquired for analytical similarity assessments, shown in Table 13.1. The most common methods utilized for HOS

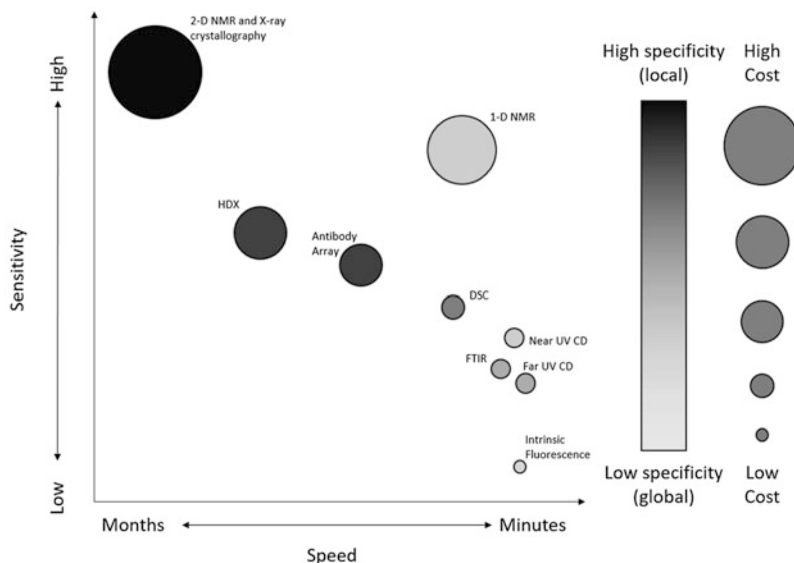


Fig. 13.2 Visualization of key factors involved in HOS method selection decisions. Method sensitivity is the ability of a method to detect differences in structure, particularly subtle changes. Method specificity refers to the spatial resolution of those structural changes (e.g., full protein, domain-specific, peptide, amino-acid, or atomic resolution). It is important to note that method sensitivity and method specificity are entirely independent; a highly sensitive method may provide global information (no specificity) or local information (high specificity). Speed, specificity, and sensitivity were qualitatively defined. Cost was based on method costs by contract research laboratories. Market prices were assumed to reflect capital costs, as well as operational costs. The fact that the symbols are not randomly scattered, and vary diagonally with both color and size, demonstrates a correlation between speed, cost, sensitivity, and specificity. Higher specificity methods take longer to run, and typically are most costly. Figure was adapted from Technical Decision Making with Higher Order Structure Data: Perspectives on Higher Order Structure Characterization from the Biopharmaceutical Industry (Weiss et al. 2015)

assessment are near UV CD, far UV CD, FTIR and DSC, which provide useful data at relatively low cost. The more resource intensive methods (e.g., NMR, X-ray crystallography) are less often utilized, suggesting these methods are considered unnecessary costs to meet the acceptable data threshold of most companies for their HOS data package.

Method Specificity and Fingerprint-Like Similarity

The goal of an HOS similarity package is to demonstrate a sufficiently high degree of HOS similarity between the reference product and the biosimilar product being developed. Reliable and efficient detection of differences in structure, should

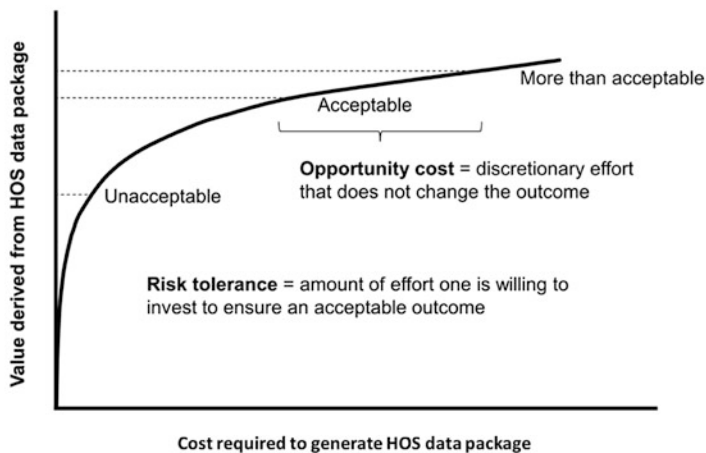


Fig. 13.3 Illustration a hypothetical relationship between effort required to generate an HOS data package and the resulting value. Figure was reproduced with consent from Technical Decision-Making with Higher Order Structure Data: Starting a New Dialogue (Weiss and Gabrielson 2015)

such differences exist, are the key features for HOS characterization methods in the context of analytical similarity. Therefore a technique's sensitivity should be emphasized over its specificity. High specificity methods, like 2D-NMR or X-ray crystallography, provide information on protein structure at specific regions and can be used to detect structural changes in specific domains or even at specific amino acids. These methods are important for elucidating the structure of the biosimilar product, and they also provide value as investigative tools when other methods detect differences between the biosimilar and reference products. However, for the purpose of analytical similarity, being able to pinpoint the location of a change in HOS is only important after other methods reliably detect changes in the protein structure. The application of highly sensitive, global methods aligns with the FDA's philosophy on fingerprint-like analytical similarity: "The results of integrated, multi-parameter approaches that are extremely sensitive in identifying analytical differences (i.e., fingerprint-like analyses) permit a very high level of confidence in the analytical similarity of the proposed biosimilar product and the reference product, and it would be appropriate for the sponsor to use a more targeted and selective approach to conducting animal and/or clinical studies to resolve residual uncertainty and to support a demonstration of biosimilarity." (Clinical Pharmacology Data 2016) Spectroscopic methods generate highly reproducible, information rich spectra reflective of the global structure of the bulk protein in solution. Near UV CD spectra, for example, contain hundreds of data points and different regions of the spectra correlate to the protein conformation around specific amino acids throughout the entire protein. These signals combine to form a highly repeatable spectral signature of the protein structure. It is unlikely that a single HOS

method is currently able to demonstrate fingerprint-like similarity on its own, but when taken together, these "fingerprint-like analyses" fit into the ensemble of data to achieve this end (Weiss et al. 2015).

Method Strategy

Most HOS methods have historically been treated as qualitative methods. Six of the seven most utilized methods are spectroscopic, and many organizations set criteria for similarity assessments based on visual inspection of the spectra. Rather than relying on inherently subjective conclusions of similarity (or lack thereof), we recommend qualifying HOS methods to obtain an objective measure of method capability. The results of method qualification can then be used to quantitatively define the sensitivity of the method to determine if it is suitable for use in analytical similarity, and if it is, the same qualification data can also be used to set criteria for analytical similarity evaluations. It can be useful, although it is not always necessary, to evaluate methods across several sources of variability such as day, lot, instrument and analyst. Obtaining objective measures of method performance is extremely useful in evaluating a method's fitness-for-purpose and setting an appropriately-sized similarity assessment study. Notably, Organization 1 from the example provided earlier in the chapter was the only organization to utilize quantitative similarity assessment criteria for spectroscopic methods, and was able to justify the most streamlined similarity assessment strategy of the briefing documents reviewed. By evaluating the effect of the manufacturing process on method variability, this organization was able to demonstrate that HOS was insensitive to variations in process condition, leading to a reduction in the number of lots tested by HOS methods.

When selecting HOS characterization methods for a particular study, an organization should first define and rank the importance of method performance parameters, like precision and specificity, and resource investment considerations like cost and throughput. When defining the performance parameters of the method, product specific characteristics must be considered. For example, far UV CD can achieve higher sensitivity for alpha helical proteins, whereas FTIR typically achieves higher sensitivity for beta-sheet proteins. After cost and performance parameters have been defined, the company can prioritize these factors in the context of a specific study; for example, for purposes of analytical similarity evaluation, the organization may prioritize method sensitivity highest, followed by speed, cost, and lastly, specificity. Finally, the organization can then rank each potential method for each factor defined and prioritized in advance. By assigning numbers to both the priority rankings and method performance rankings, an overall score can be computed to quantitatively compare methods to each other for purposes of achieving the desired study goals. Of course, this exercise can be completed in a less formal way using an organization's prior knowledge, and two or more orthogonal methods that measure the same attribute in different ways may be included to ensure similarity of more critical product attributes. Although the exact process and

organizational priorities may differ, a documented method ranking process provides strong justification for method selection decisions, which can be beneficial during interactions with health authorities. Some common HOS methods are discussed below based on speed, cost, and sensitivity as a starting point for a method ranking process.

Circular Dichroism

Circular dichroism (CD) is an absorption based spectroscopic technique which uses circularly polarized light to detect ordered structure around optically active chiral molecules. Circular dichroism is observed when optically active moieties, such as chromophores, absorb left and right handed circularly polarized light differently. The difference between left-handed and right-handed absorption produces a difference spectrum containing both positive and negative signals. The secondary structure of proteins can be determined by CD spectroscopy in the far UV wavelength region (ca. 190–250 nm), where the peptide bond (amide bond) produces CD signals characteristic of well-ordered secondary structures (e.g., helices and sheets). In the near UV region (ca. 240–340 nm), aromatic amino acids produce CD signals characteristic of their local environment, which are reflective of the tertiary structure of the protein. Paragraph Break Near UV CD has become the standard method for characterization of protein tertiary structure due to its low cost, high throughput, and high sensitivity. Data acquisition for a single sample typically requires about 1 mg of material, and takes less than an hour. Method specificity is low, but this is not a significant concern for similarity assessment. Most organizations have utilized CD data as a qualitative method (tier 3 attribute); however, tier 2 assessments using spectral comparison algorithms like the weighted spectral difference (WSD) have been implemented with success (Teska et al. 2012). Far UV CD and FTIR are the two most common methods utilized for assessing secondary structure. There are not significant cost, speed or sensitivity differences between FTIR and far UV CD, and final method selection is typically based on product related considerations.

Fourier Transform Infrared Spectroscopy

Fourier transform infrared (FTIR) spectroscopy has become a standard method for evaluating the secondary structure of proteins and peptides. Hydrogen bonding of various secondary structures cause shifts in the profile of the amide I band between 1600 and 1700 cm^{-1}. Ten peaks in the amide I band have been identified and used to deconvolute FTIR spectra and determine an approximate percentage of each type of secondary structure (Dong et al. 1990). FTIR is a high throughput technique capable of analyzing samples in typical formulation conditions in several minutes.

The low cost and high throughput of this technique has made it a standard choice for analytical similarity assessments. Much like CD, it is most common to treat FTIR data as a tier 3 attribute, assessing visual similarity of the absorbance or transmission spectra. Taking the second derivative of the absorbance spectra amplifies visual differences in the amide I band, and it is common to use the second derivative of the spectra for qualitative comparison. FTIR is a high sensitivity method at typical formulation concentrations (>5 mg/ml), and quantitative similarity assessment can be performed using spectral comparison algorithms like QC compare or the WSD algorithm (Teska et al. 2012).

Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is the most widely applied non-spectroscopic higher order structure method for analytical similarity assessment. Figure 13.1 demonstrates the wide applicability of DSC data across various stages of product development. For analytical similarity assessments, the high throughput and sensitivity of DSC make it an extremely valuable tool. Unlike spectroscopic methods, which directly measure the local environment or resonance of specific protein moieties, DSC measures the impact of structural changes on the thermal stability of the protein. This make DSC a useful orthogonal method to other HOS spectroscopic methods.

Due to the high precision of DSC, very small shifts in heat capacity can be reliably measured, and provided that buffer conditions of the samples are identical, it may be inferred that these slight changes in the protein stability indicate protein conformational stability changes. DSC thermograms are typically assessed qualitatively for similarity, and domain melting temperatures (Tm) are often evaluated as tier 2 attributes. Due the very high method sensitivity, minor differences in formulation can cause detectable shifts in thermal stability. This is often overcome by dialyzing samples into a common formulation, or diluting high concentration stock material into the same formulation. DSC experiments are typically conducted at 0.5–2 mg/mL concentrations which may not be at the formulation concentration of the product (Arthur et al. 2014).

1-D NMR

1-D NMR, or ^{1}H NMR, has the allure of higher precision spectral measurements than CD or FTIR, while maintaining high sample throughput. Despite the high precision of the method, the sheer number of peaks for protein samples, particularly for higher molecular weight proteins such as a 150 kDa protein, makes either quantitative or qualitative analysis challenging. Visual comparisons are common for 1-D NMR, but quantitative spectral comparisons were not widely utilized in

the briefing documents we reviewed. When used as a visual comparison method, 1-D NMR suffers from higher costs, and slightly lower throughput than far UV CD and FTIR. Recent publications have proposed novel spectral processing methods to improve 1-D NMR applicability to analytical similarity, and create highly sensitive quantitative measures of overall structural similarity. These emerging techniques show promise, but we were unable to determine if quantitative 1-D NMR methods were used to support the analytical similarity package in any of the publicly available regulatory submissions we reviewed (Poppe et al. 2013).

2-D NMR

2-D NMR has historically been considered the gold standard for detailed structural elucidation and characterization, providing extremely precise data around each amino acid in the protein. However, large proteins suffer from peak overlap and broadening which has made correlation spectra less informative (Poppe et al. 2013; Ghasriani et al. 2016). In addition to the technical challenges present for large proteins, 2-D NMR suffers from high instrumentation cost, high operational costs, and low throughput. 2-D NMR can take weeks to acquire high quality data for a single sample, which is prohibitive for similarity assessments attempting to compare to 10–30 lots of reference product material. Statistical evaluation of 2-D NMR similarity is not practical due to throughput and cost limitations, forcing 2-D NMR to remain a supporting qualitative method, despite the high precision of the data.

Hydrogen-Deuterium Exchange

Hydrogen-Deuterium exchange (HDX) provides precise information on protein structure by exchanging deuterium with exposed, interchangeable hydrogen. Protein digestion and characterization by mass spectrometry (MS) provides precise way to interrogate surface exposed amino acids. General industry opinions shown in Fig. 13.1 suggest that biosimilarity is considered one of the best applications for HDX; however, moderately high costs and low throughput, compared to many of the spectroscopic methods, have prevented wide application of HDX-MS for analytical similarity assessment. Of the documents we reviewed, only one organization included HDX-MS in its analytical similarity assessment. To reduce costs, this organization only compared two US lots, and leveraged CD, FTIR, DSC and X-ray crystallography to demonstrate comparability to EU lots.

X-ray Crystallography

X-ray crystallography provides extremely precise information on protein structure by measuring the diffraction of an X-ray beam through crystalized protein. Precise atomic position can be determined in addition to information on chemical bonds. X-ray crystallography has not been widely adopted for analytical similarity due to high costs and low throughput. Furthermore, protein crystallization conditions are often very different from formulation conditions. This method was only used by one organization, and the analysis was limited to one development lot, one US reference product lot, and one EU reference product lot. These data were used to show a precise comparison between material types, and other spectroscopic methods were used to compare across additional lots. Statistical evaluation of X-ray crystallography data is not practical due to throughput and cost limitations, causing X-ray crystallography to remain a qualitative method, despite the high precision of the data.

Intrinsic Fluorescence

Intrinsic fluorescence uses the well characterized shift in tryptophan fluorescence as this amino acid becomes more solvent exposed. Using the shift in both the wavelength of maximum fluorescence and the maximum fluorescence intensity, changes in the tertiary structure of proteins can be inferred. CD and FTIR spectroscopy have largely replaced intrinsic fluorescence for most proteins and applications, as these methods provided improved precision with similar cost and throughput.

Summary of HOS Methods

Most higher order structure methods are spectroscopic, with the most common methods used being FTIR and circular dichroism. Despite the criticality of HOS for proper protein function, few organizations have treated any spectroscopic method as tier 2, and we suspect that this is for pragmatic reasons. Spectroscopic methods provide rich information and can be shown to be sensitive to HOS changes, but the spectra can be difficult to translate into quantitative acceptance criteria. For example, spectral differences can arise from very minor differences in formulation excipients that can create spectral shifts unrelated to changes in protein structure. Although challenging, distinguishing between experimental variability, instrument variability buffer-induced changes, and true protein structural differences can accomplished. Further complicating analysis is a wide range in data quality across the industry, shown in Fig. 13.3.

We believe that qualifying HOS characterization methods is the only way to overcome these challenges. Determining objective measures of method performance and estimating the impact of significant sources of variability has two benefits. First, and most importantly, method qualification enables a more robust and defensible similarity assessment, which better aligns with the known impact of HOS on protein function. Second, method qualification enables an objective ranking of each method's fitness-for-purpose, based on the unique considerations of each product, each organization, and each study. Organizations that have implemented quantitative similarity criteria for HOS characterization methods have been able to justify at more streamlined similarity assessment strategies.

Conclusion

Significant diversity is evident within the biopharmaceutical industry for the application of higher order structure methods to analytical similarity assessments of biosimilars. The diversity of approaches does not seem to arise from modality-specifc considerations or from potential impact of HOS on clinical study design. Rather, in our view, this diversity likely emerges from a combination of the diverse viewpoints among decision makers at different organizations related to the value of HOS data, and the varying risk tolerances of those decision makers. Organizations with a higher risk tolerance or those that place lower priority on HOS methods tend to design similarity assessments with an increased emphasis on speed and minimization of resource requirements.

Despite the diverse range of strategies within the biopharmaceutical industry, several key trends have emerged. Organizations with approved biosimilars have all used orthogonal methods to demonstrate similarity of protein higher order structure. Orthogonal HOS characterization is essential to reduce the residual uncertainty that exists between a biosimilar and its reference high product. Methods for analytical similarity assessment should be thoughtfully selected, and criteria should be carefully established to meet product-specific, study-specific, and organization-specific method requirements, including speed, cost, sensitivity, and specificity. FTIR, CD, and DSC have consistently been included in similarity assessments for approved biosimilars. Several emerging methods, like NMR spectroscopy, are also being included in analytical similarity assessments. These methods often have technical limitations or prohibitive resource requirements preventing them from being used to characterize multiple product lots. As these technical limitations are overcome, we believe methods with improved, sensitivity and throughput will become standard components in HOS similarity assessments.

References

Amgen. Background information for the Arthritis Advisory Committee. ABP 501 (adalimumab biosimilar). 2016.

Amgen. FDA briefing document Ocologic Drugs Advisory Committee. ABP215 (bevacizumab biosimilar). 2017.

Amgen Financial Report. 2017. https://www.amgen.com/media/news-releases/2017/02/amgen-reports-fourth-quarter-and-full-year-2016-financial-results.

Arthur K, Dinh N, Gabrielson J. Technical decision making with higher order structure data: utilization of differential scanning calorimetry to elucidate critical protein structural changes resulting from oxidation. J Pharma Sci. 2014;2014:1548–54.

Celltrion. Briefing document for the Arthritis Advisory Committee. CT-P13 (infliximab biosimilar). 2016.

Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product. FDA guide for industry. 2016. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm397017.pdf.

Coherus Biosciences, Inc. Form 8-K. Item. 2017; 2:05.

Davio, K.. Coherus receives complete response letter for pegfilgrastim biosimilar. The Center for Biosimilars. 2017. https://www.centerforbiosimilars.com/news/coherus-receives-complete-response-letter-for-pegfilgrastim-biosimilar.

Dong A, Huang P, Caughey W. Protein secondary structure in water from second-derivative amide I infrared spectra. Biochemistry. 1990;29(13):3303–8.

FDA guidance for industry–scientific considerations in demonstrating biosimilarity to a reference product. 2015. https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf. Accessed 16 Jan 2018.

Ghasriani H, Hodgson DJ, Brinson RG, et al. Precision and robustness of 2DNMR for structure assessment of filgrastim biosimilars. Nat Biotechnol. 2016;34:139–41.

Guidance for Industry. Quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. 2015. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf. Accessed 12 Mar 2018.

Guideline on Similar Biological Medicinal Products. CHMP/437/04. 2005 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003517.pdf.

Hospira. FDA briefing document Oncologic Drugs Advisory Committee Meeting. Epoetin Hospira (epoetin alfa biosimilar). 2017.

Mylan. FDA oncologic drugs Advisory Committee Meeting. MYL-1401O (trastuzumab biosimilar). 2017.

Poppe L, Jordan B, Lawson K, Jerums M, Apostol I, Schnier P. Profiling formulated monoclonal antibodies by H NMR spectroscopy. Anal Chem. 2013;85(20):9623–9. https://doi.org/10.1021/ac401867f.

Sandoz. FDA briefing document Oncological Drugs Advisory Committee Meeting. EP2006 (filgrastim biosimilar). 2015.

Sandoz. Arthritis Advisory Committee briefing document. GP2015 (etanercept biosimilar). 2016.

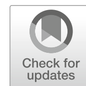
Teska B, Li C, Winn B, Arthur K, Jiang Y, Gabrielson J. Comparison of quantitative spectral similarity analysis methods for protein higher-order structure confirmation. Anal Biochem. 2012;434:154–65.

Vandekerkhhove K, Seidl A, Gutka H, Kumar M, Gratzl G, Keire D, Coffey T, Kuehne H. Rationale selection, criticality assessment, and tiering of quality attributes and test methods for analytical similarity evaluation of biosimilars. AAPS J. 2018;20:68. https://doi.org/10.1208/s12248-018-0230-9.

Weiss W, Gabrielson J. Technical decision-making with higher order structure data: starting a new dialogue. J Pharm Sci. 2015;104:1240. https://doi.org/10.1002/jps.24393

Weiss W, Gabrielson J, Al-Azzam W, Chen G, David D, Das T, Hayes D, Houde D, Singh S. Technical decision making with higher order structure data: perspectives on higher order structure characterization from the biopharmaceutical industry. J Pharm Sci. 2015;105(12):3465–70. https://doi.org/10.1016/j.xphs.2016.09.003.

Chapter 14
Protein Conformational Array Technology for Biosimilar Higher Order Structure Analysis



Xing Wang, Wen-Rong Lie, and Jehangir Mistry

Abstract Biologics Higher Order Structure (HOS) is important to its safety and efficacy but difficult to define. A novel technology called Protein Conformational Array (PCA) was developed using antibody arrays to analyze monoclonal antibody Higher Order Structure. Protein Conformational Array (PCA) technology was developed based on the immunology epitope scanning principle. This technology offers systematic coverage toward the full amino acid sequence of the biologics with high sensitivity and accuracy. For the high volume analytical assays such as cell line selection, bioprocess and formulation development, more accurate and relatively high throughput methods will be more desirable. By using 34 different antibodies covering the whole mAb molecule and measuring the mAb surface epitope exposure, the mAb HOS can be precisely and systematically described. PCA is in two different formats, the ELISA format and magnetic beads-based Luminex format. The first few sections of this chapter will focus on the applications and findings from the ELISA-based PCA and the last two sections will discuss the Luminex system in elucidating detains of mAbs HOS under different stress conditions. Results will be discussed related to mAb HOS stability polymorphism, the multifaceted nature of the PCA technology and its applications in bioprocess and formulation development and comparability studies.

Keywords Higher order structure (HOS) · Protein conformational array (PCA) · Comparability analysis · Biosimilarity · Bioprocess development · Formulation development · Stability · Polymorphism · Immunogenicity

---

X. Wang (✉)
Array Bridge Inc., St. Charles, MO, USA
e-mail: xing.wang@arraybridge.com

W.-R. Lie · J. Mistry
MilliporeSigma Corp., St. Charles, MO, USA
e-mail: wen-rong.lie@emdmillipore.com; Jehangir.mistry@emdmillipore.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_14

Introduction

For biologics, the importance of the molecule's structure to its efficacy and safety is well established (Hermeling et al. 2004, 2005, 2006; Jiskoot et al. 2009; Maas et al. 2007; Mok et al. 2016). However, because of the rotational property of the amino acid α carbon and the large molecular size, a protein could assume enormous number of different conformations (US FDA 2012), some of the conformation may not be desirable for its safety and efficacy (Hermeling et al. 2004; Mok et al. 2016; Arosio et al. 2011; Beck et al. 2014; Berkowitz et al. 2012; Bessa et al. 2015; Buttel et al. 2011; Chang et al. 2005; Laat et al. 2011; Ohkuri et al. 2010; Ratanji et al. 2014; Rosenberg 2006; Porter 2001; Schellekens 2005; Sharma 2007). The high complexity of the antibody Higher Order Structure (HOS) makes the accurate determination of its conformational status a challenge. Currently, several technologies have been used to analyze mAb HOS including circular dichroism (CD), FTIR, fluorescence spectroscopies and differential scanning calorimetry (Arthur et al. 2015; Gabrielson and Weisis 2015; Zurdo 2013; Jiang et al. 2015; Thiagarajan et al. 2016; Wei et al. 2009, 2011, 2016), those technologies play important role in the development of biologics including biosimilars and will be reviewed in a separate chapter. Monoclonal antibodies (mAbs) are the most rapidly growing class of biologics; more than 60 mAbs have been approved so far including many of the best-selling drugs in the world. It is therefore important to develop new technologies with high sensitive and accuracy for the determination of mAb HOS and provides useful information for the selection of the optimal bioprocess and formulation during the mAb development and production. For the high volume analytical assays such as cell line selection, bioprocess and formulation development, and a more accurate and relatively high throughput method will be more desirable. Protein Conformational Array (PCA) technology was developed from the immunology epitope scanning principle (Wang et al. 2013). By using 34 different antibodies covering the whole mAb molecule and measuring the mAb surface epitope exposure, the mAb HOS can be precisely and systematically described (Fig. 14.1).

Currently PCA is in two different formats, the microplate-based ELISA format and magnetic beads-based Luminex format. The first few sections of this chapter will focus on the development and applications from the ELISA-based PCA and the last two sections will discuss the Luminex system in elucidating details of mAb HOS under different stress conditions.

In the recent guidelines for biosimilar development from FDA, a "fingerprint-like" technology is preferred to assess the structural biosimilarity (CBER/CDER 2015). Since the PCA technology is designed to cover the whole mAb molecule and measuring mAb epitope distribution on the surface, it is essentially measuring the conformational changes at molecular level. Furthermore, extensive study has demonstrated that this technology can detect HOS changes in as low as 0.1%

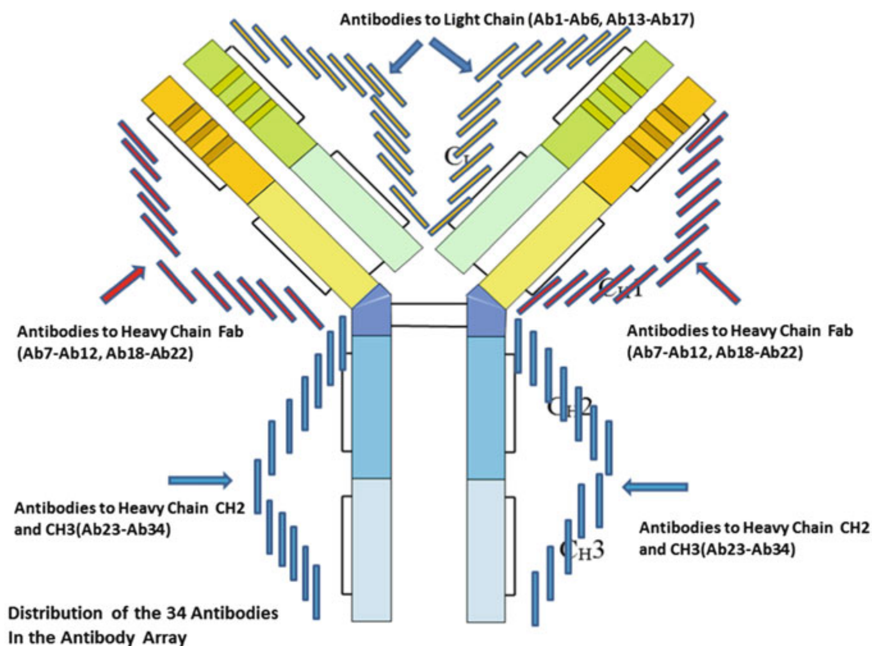


Fig. 14.1 Diagram of the protein conformational array (PCA) design

population of the mAb with high accuracy and precision (discussed in more details in the following sections), PCA is particularly important for biosimilar HOS characterization and similarity assessment.

PCA Development

Antibody Specificity

Since it is known from immunology that linear epitopes recognized by antibodies constitute 5-6 amino acids sometimes 3-4 amino acids, the concept of PCA technology is to develop a series of antibodies that can recognize multiple linear epitopes from the amino acid sequence. One of the criteria required to develop a successful antibody array is the capability to probe regional changes in biologics molecules. To achieve this goal, the antibody specificity is critical. In an earlier publication (Wang et al. 2013), each of the different anti-peptide antibodies were probed with each of the corresponding peptides in a pairwise analysis, i.e. each peptide was analyzed by all the different antibodies and each antibody was probed with different peptides. The results showed that good antibody specificity was achieved. There were a few cases of cross-reactivity from the 30 peptides and 30

antibodies cross-testing (900 data points), but the overall specificity was good, suggesting that these peptide-derived antibodies could be used to probe regional changes in the mAb structure. In this pairwise testing, the main cross-reactivity occurred on the two side of the diagonal line; this was actually expected from the special design of the antibody array. In the antibody array design, peptides covering the entire monoclonal antibody light chain and heavy chain were synthesized with overlapping regions, therefore each peptide will have two overlapping regions from its N-terminal and C-terminal ends respectively, only the peptide corresponding to the very ends of the mAb light chain and heavy chain will have one overlapping region. Because of this special design, the polyclonal antibodies generated from one peptide could potentially recognize the adjacent peptides, and for the same reason, the antibodies generated from the two adjacent peptides could also recognize the peptide in the middle.

Sensitivity of the PCA Assay

Another important parameter for the antibody array technology is the sensitivity of the assay. As a useful analytical technology, it needs to detect either regional changes of the whole mAb population or changes in a sub-population which could reflect the mAb's HOS changes induced from the bioprocess or formulation. To estimate the sensitivity of the technology, an unfolded mAb was generated with the treatment of 8M urea which resulted in the unfolding of proteins and making the linear epitopes available. In the sensitivity testing, unfolded mAb was spiked into native mAb at 0%, 0.1%, 0.2% and 0.5% ratio respectively; it was demonstrated that as low as 0.1% spike will result in a significant increase of ELISA reading, suggesting that at least 0.1% novel epitope exposure could be detected and quantified by PCA ELISA (Wang et al. 2013). Compared with other analytical methods used for mAb HOS analysis, PCA technology is a more sensitive method. More importantly, since it is in an ELISA format, the PCA technology is an accurate analytical method with assay variation typically less than 15%.

For a typical comparability assay, 5 μg/ml mAb is recommended. It is important to point out that, because of the large size of monoclonal antibodies and the complex structure of the molecule, many of the linear epitopes the antibody array raised against are not exposed on the surface of the molecule. Therefore at any time, the antibody array can only detect a sub-population of the mAb for which the corresponding linear epitopes are exposed on the surface, this sub-population can be considered as a "conformational impurity" (Wang et al. 2013). It is believed that this measurement of the "conformational impurity" will provide important information on the biosimilarity of the biosimilar mAbs vs. innovator molecules, it is also believed that this "conformational impurity" may also be related to the mAb immunogenicity potential and/or its efficacy, more studies in this area are under way.

Applications in Biosimilar mAb HOS Comparability Analysis

Since the introduction of the PCA technology, many leading biosimilar developers in the world have used it for mAb biosimilarity analysis and the technology has become one of the standard technology for biosimilar HOS assessment including the first biosimilar mAb approved in the world, Infliximab, first in Europe and then in US (Jung et al. 2014; Chaudhari et al. 2017; DiPaola 2017; Wang et al. 2014). With the specificity and sensitivity of the PCA technology addressed above and to further understand the capability of the technology, the ELISA-based method was used in several case studies to evaluate HOS comparability between the innovator molecules and the biosimilar candidates. The typical outcomes can be categorized into three groups: the first group is that the biosimilar and innovator mAb showed good HOS similarity with minor differences, the majority of the biosimilars tested fell into this group. The second group is that the biosimilar and innovator mAb showed high similarity, in fact the ELISA can't find any HOS differences between these two groups, a few tested biosimilars fell into this group. The third group is that the biosimilar and innovator mAb showed significant HOS differences, a few testing felled into this group as well. In our first case study, multiple batches of innovator Herceptin (Trastuzumab) and biosimilar mAb were analyzed using PCA ELISA (Fig. 14.2).

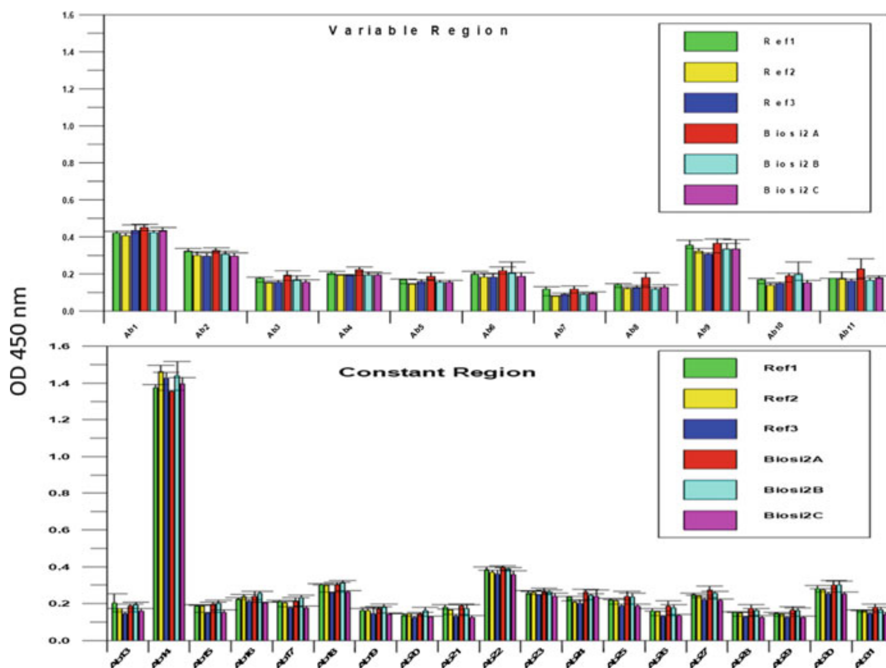


Fig. 14.2 HOS comparability analysis of trastuzumab and biosimilar candidate

For all the coverage areas measured by the 34 antibodies, no differences were detected that were larger than the method variation (15% RSD). Since it has been shown that the PCA ELISA could detect as low as 0.1% new epitope exposure in a native mAb population, this suggested that no more than 0.1% conformational impurity exists in the biosimilar mAb. It should be pointed out that this is a rather rare case in the comparability studies because the majority of biosimilar mAbs tested showed some conformational difference in certain areas of the mAb. On one hand, it was shown that it is possible to develop a cell line that can produce a biosimilar mAb that is highly similar to that of the innovator molecule in its Higher Order Structure. On the other hand, this is only one of a few cases we have observed with no significant HOS differences in PCA ELISA, underscoring the difficulty of achieving complete conformational similarity between originator and biosimilar mAbs. In the second study, Avastin was compared with its biosimilar mAb using PCA ELISA, it was found that there was an increase of signal across the whole antibody panel suggesting a small population of mAb unfolding (Wang et al. 2014). In addition, there were a few regions that showed epitope exposure higher than the average, indicating that additional epitopes were exposure in these regions. Based on the spike testing of unfolded mAb, it represents 0.1–0.2% new epitope exposure, this is considered a minor difference for the mAb HOS compared with that of the innovator molecule. Bioassays on this biosimilar mAb did not detect any differences on the biosimilar molecule. In the third case study, a Humira (Adalimumab) biosimilar candidate was compared with the innovator molecule using 3 batches each and the test showed that one batch of the biosimilar mAb matched very well with the innovator mAb in the HOS profile, however results from the other two batches of biosimilar Adalimumab indicated that there was an increased surface epitope exposure across all the regions covered by the PCA ELISA, suggesting a general unfolding of a small percentage of the mAb representing 0.1–0.2% new epitope exposure. In another case study, the PCA ELISA showed that the innovator mAb and biosimilar mAb showed significant HOS differences. For this biosimilar candidate, bioassay testing showed that the biosimilar mAb was losing potency in an ELISA-based binding assay; however no analytical testing used could detect structural changes. With the PCA ELISA, it was found that there were new epitope exposures in several areas of the mAb; some of the areas are close to the "hot spots" of conformational variability such as the hinge region and the area around the glycosylation site. The PCA technology is currently being used to assess HOS comparability by biosimilar developers, some of the biosimilars evaluated have already being approved in Europe, US and other countries (Jung et al. 2014).

Applications in Bioprocess Development

Biosimilar development is a reiterative process from clone to purification wherein the product is analyzed and tested to check if all attributes are similar or close to reference product. Hence assessing the HOS and other biochemical properties of

the biosimilar during bioprocess development is critical for a successful biosimilar development. In addition to HOS comparability assessment on biosimilar mAbs, the PCA technology can also provide valuable insights in bioprocess development process employed for biosimilar development and manufacturing. In one case study, two biosimilar mAbs were tested for their HOS status during bioprocess development. For biosimilar-1, a total of 30 samples spanning the entire bioprocess were analyzed using the PCA ELISA. Because of the large number of in-process samples, for the first step, three PCA ELISA antibodies known to detect conformational changes in some “hot spots” of the mAb molecule were chosen to detect conformational changes. Among the 3 PCA antibodies selected, antibody 19 covers the mAb heavy chain amino acid 154–179 (based on Trastuzumab amino acid sequence), at the interface between the Fv and CH1 domain. Antibody 25 covers the Heavy chain amino acid 272–293, close to the hinge region and glycosylation site, and finally antibody 30 covers the heavy chain CH3 amino acid 355–379, close to the C-terminal of the heavy chain. In the upstream samples (Fig. 14.3), mAbs from different culture conditions (days of culture) were tested directly (prior to any purification steps) for their conformational status.

This analysis showed a relatively stable level of conformational impurity (new epitope exposure) up to day 9 and significant increase of conformational impurity at day 10 (Davies et al. 2016). For downstream process, samples from three purification steps (Protein A, cation exchange and anion exchange columns) were

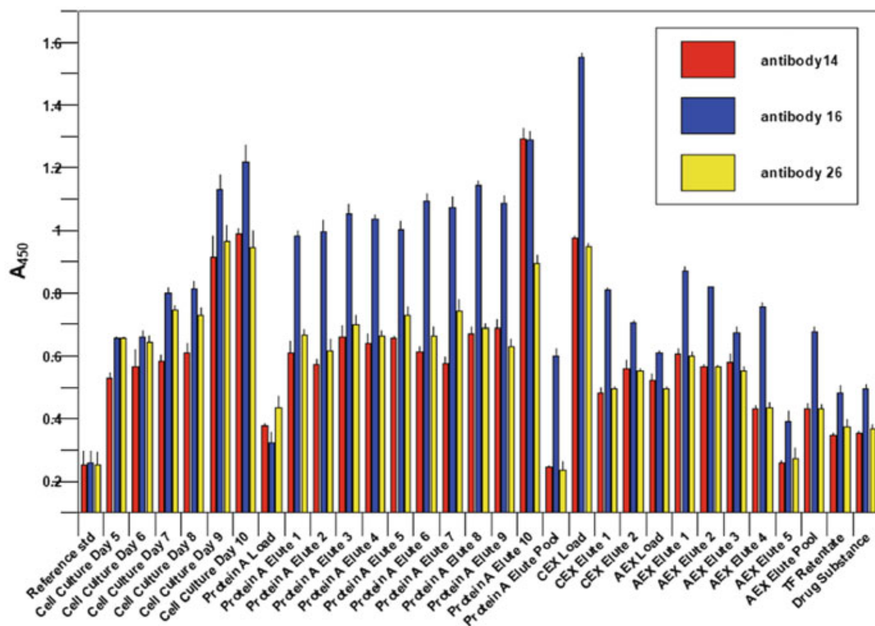


Fig. 14.3 HOS analysis of biosimilar candidate during bioprocess development

tested. The results indicated that Protein A column eluates had decreased epitope exposure in general compared to the mAbs from the medium with the exception of elute 9 which has significant increase of conformational impurity corresponding to the region of Ab19, suggesting that elution conditions had an impact on the conformational status. In the cation exchange (CEX) column purification, there was a relatively small level of increase in mAb epitope exposure. Following elution of the mAb from the CEX column, there was no significant change in epitope exposure from the anion exchange (AEX) column purification except elute 4 which showed significant increase in epitope exposure in the region covered by Ab19. It will be interesting to know the condition differences between eluate 4 and the other four batches. Finally, the deep filtration retentate and drug substance have similar levels of epitope exposure. When compared with reference standard, the conformational impurity profile seems very similar as measured by the three selected antibodies. For biosimilar-2 mAb, 14 samples were tested including upstream samples from cell culture harvested from 5 different time points and downstream samples from three purification columns. The study found that the mAb from Protein A load had epitope exposure similar to those from the upstream cell culture and after the Protein A column, the majority of the conformational impurity was cleared (Fig. 14.4).

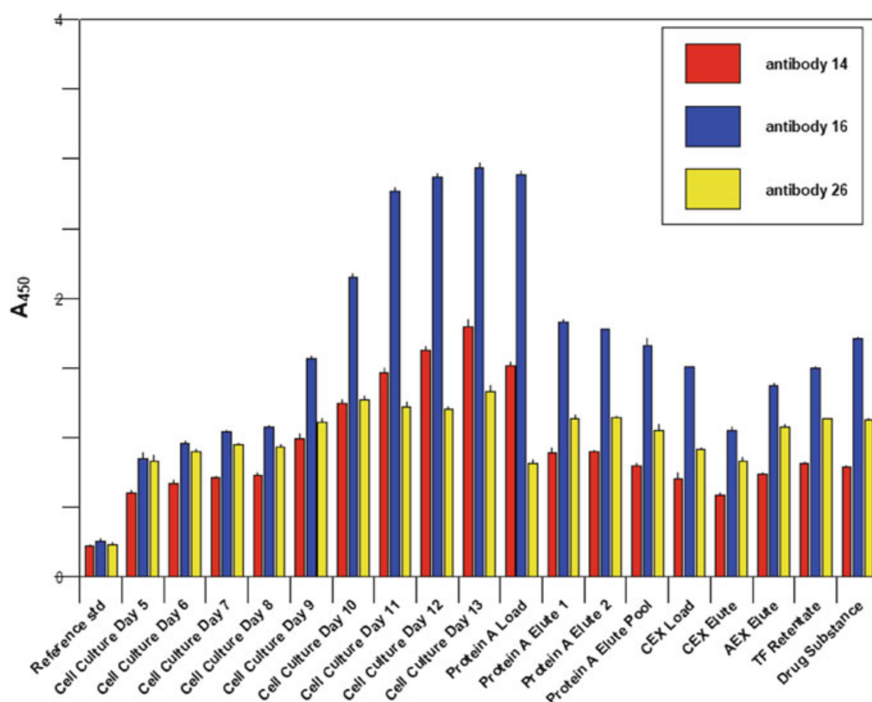


Fig. 14.4 HOS analysis of biosimilar candidate during bioprocess development

On the CEX column, the purification process further decreased conformational impurity as measured by all the three antibodies used. On the anion exchange column (AEX), for an unknown reason, there is a significant epitope exposure in the region covered by Ab19 whereas the other two regions monitored by Ab25 and Ab30 did not show much change. Drug substance has relatively low epitope exposure, however since the reference standard for biosimilar-2 was not available at the time of testing, the HOS status of the drug substance and reference standard can't be compared. After the initial testing of both upstream and downstream process samples by the three selected antibodies above, two samples from Biosimilar-1, cell harvest and drug substance were selected for the analysis with the full antibody panel of 34 antibodies and compared with the reference standard. The results indicated that cell harvest had the highest relative epitope exposure among the three samples tested in both the variable region and constant region. Biosimilar-1 drug substance had decreased epitope exposure as compared to cell harvest but similar epitope exposure compared to the reference standard across the full antibody panel (Fig. 14.5).

It is interesting to note that for the cell harvest sample, some regions had relatively more epitope exposure than others, such as those regions covered by antibody 2, 3, 8, 18, 24, 25, and to a less extent for the rest of the antibody panel (Davies et al. 2016). For Biosimilar-2, there was a significant decrease of

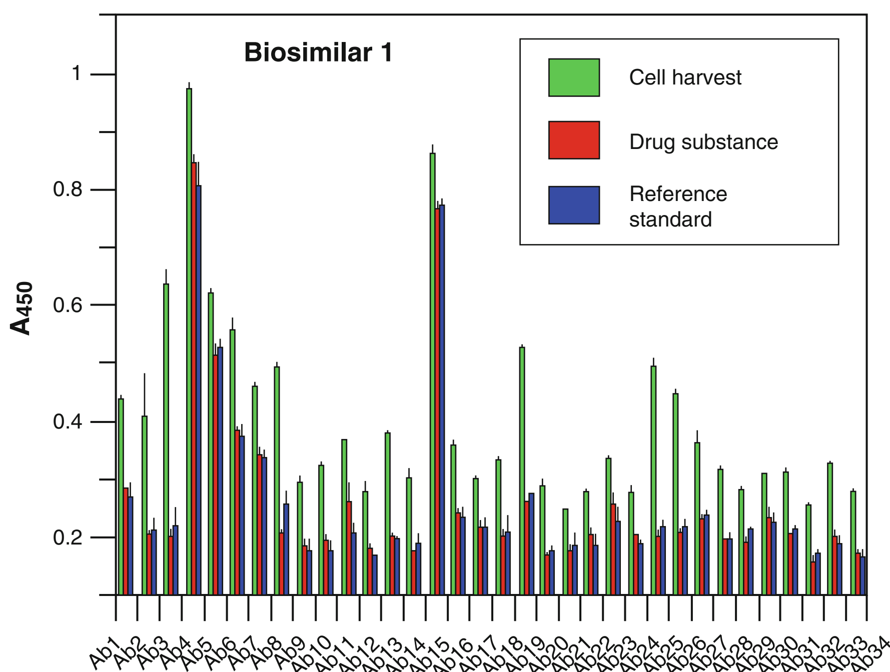


Fig. 14.5 Full panel HOS analysis of biosimilar candidate during bioprocess development

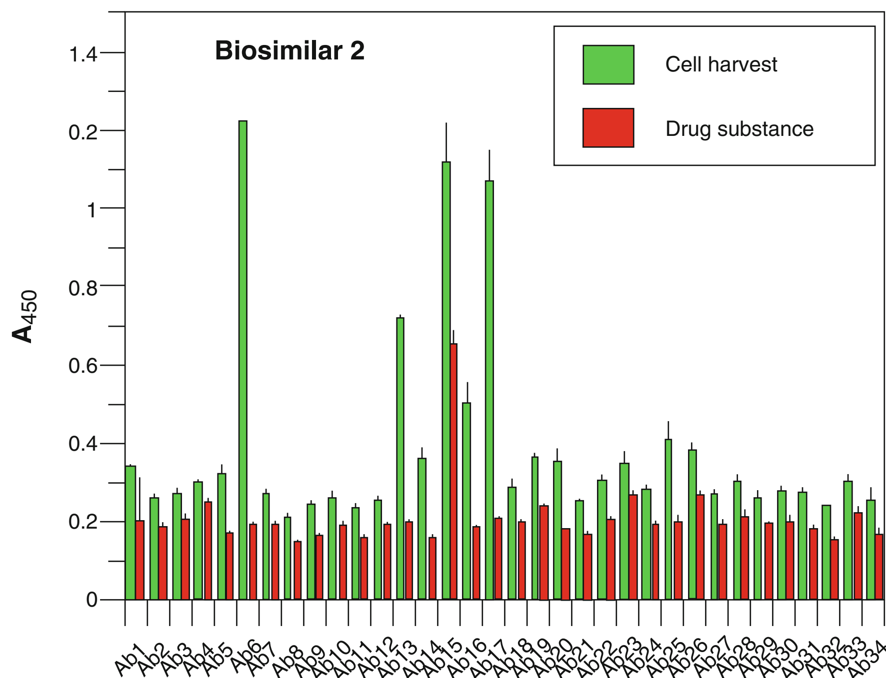


Fig. 14.6 Full panel HOS analysis of biosimilar candidate during bioprocess development

conformational impurity from cell culture samples to drug substance in the mAb (Fig. 14.6). In specific regions covered by Ab6, 13, 15, 16 and 17, there is significant increase of epitope exposure from the cell harvest sample.

Applications in Formulation Development and Accelerated/Stress Condition Stability Testing

The PCA technology was also used in the area of formulation development to test mAb HOS stability. MAb stability is one of the major quality attributes in the development of the molecule (Davies et al. 2015). In a case study, the effect of exposure to increased temperature was examined for both IgG1 and IgG2 mAbs respectively. The IgG1 sample incubated at 55 °C for 10 days was compared to a control sample, while an IgG2 sample (IgG2-a) stored at 40 °C for 14 days was compared to the corresponding control sample. In both cases, the PCA ELISA results suggested significant new epitope exposure in both the mAb variable region (covered by pAb1 through pAb12 in the ELISA panel) and constant regions (covered by pAb13 to pAb31) with significant differences between the IgG1 and IgG2 molecules (Fig. 14.7).

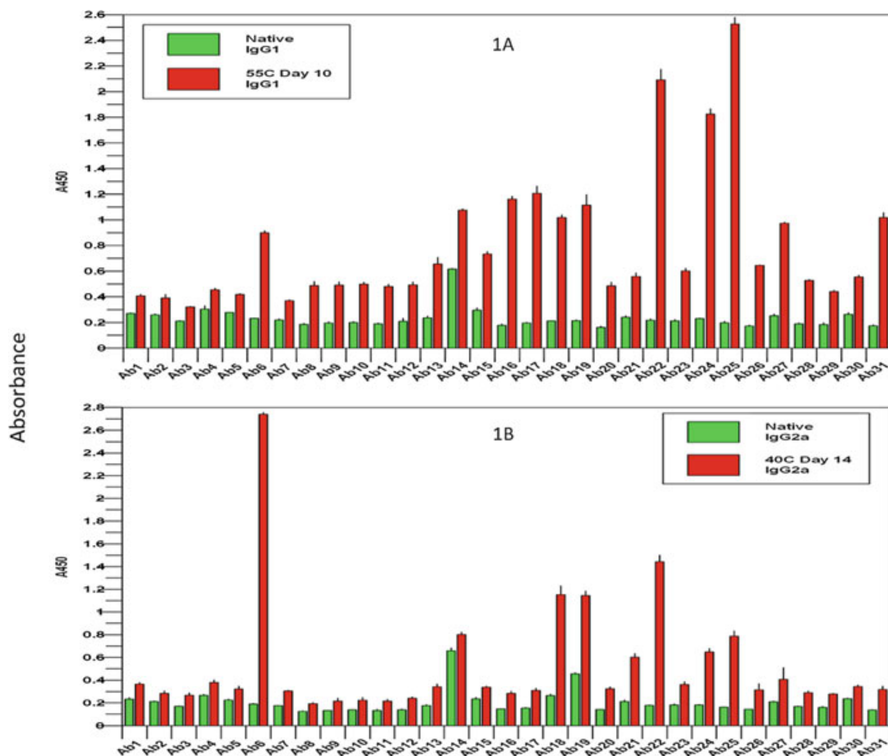


Fig. 14.7 HOS analysis of novel mAb under stress conditions during formulation development

While both datasets suggested a general partial unfolding of the mAb as indicated by the increased signal across the whole antibody panel, the 'hotspots' of greatest epitope exposure differed between the two mAbs. The IgG1 result suggested greatest sensitivity to temperature in the light chain at the boundary between the V_L and C_L domains (pAb6), in both chains near the hinge region (pAbs 14–22) and in the heavy chain C_{H2} region (pAbs 22–25). In contrast, the IgG2 result suggested somewhat better stability in the hinge region (pAbs 14–22) but extreme sensitivity in the light chain at the boundary between the V_L and C_L domains (pAb6). The apparent increased stability of the IgG2 mAb in the hinge region could be explained both by the lower temperature incubation and by the additional disulfide bonds present in this region in an IgG2 vs IgG1 molecule. In the hinge region, the two heavy chains are linked by 2 disulfide bonds in an IgG1 molecule and 4 disulfide bonds in an IgG2 molecule. One of the interesting observation was that the HOS stability of the IgG1 and IgG2 constant regions are very different even though their primary amino acid sequence are highly homologous (Wang et al. 2009), this observation suggested that factors other than the primary sequence contribute to the HOS stability of the mAb.

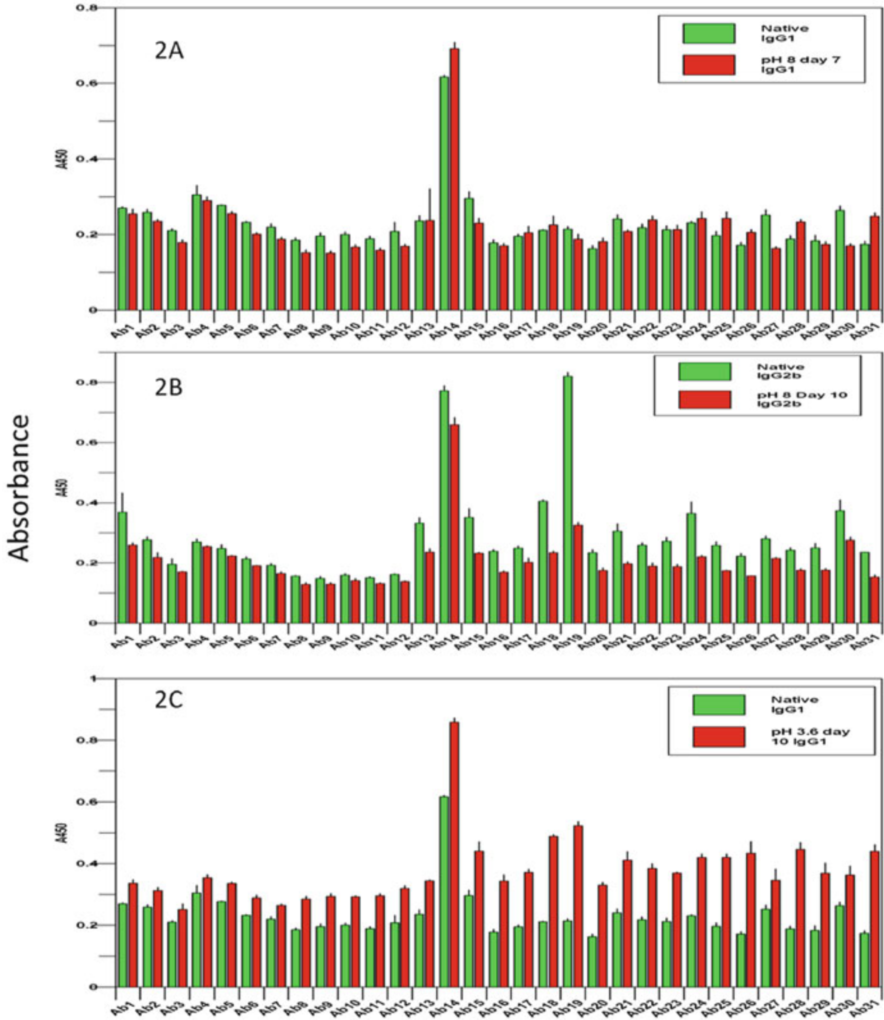


Fig. 14.8 HOS analysis of novel mAb under stress conditions during formulation development

It is known that pH has significant impact on the stability of mAb molecules (Mason et al. 2012). In the case study, the effect of exposure to basic pH condition on both an IgG1 sample and an IgG2 sample were examined (Fig. 14.8).

In addition, data was collected on the effect of acidic pH condition on sample IgG1. For both IgG1 and IgG2, the HOS of the variable region (pAbs 1–12) appeared relatively stable to 7–10 day exposures to pH 8.0, with slight additional exposures occurring at several epitopes. In contrast, the constant region (pAbs 13–31) of the IgG2 sample appeared far more susceptible to pH 8.0 exposures than did the same region of the IgG1 molecule. In contrast to its relative stability to pH

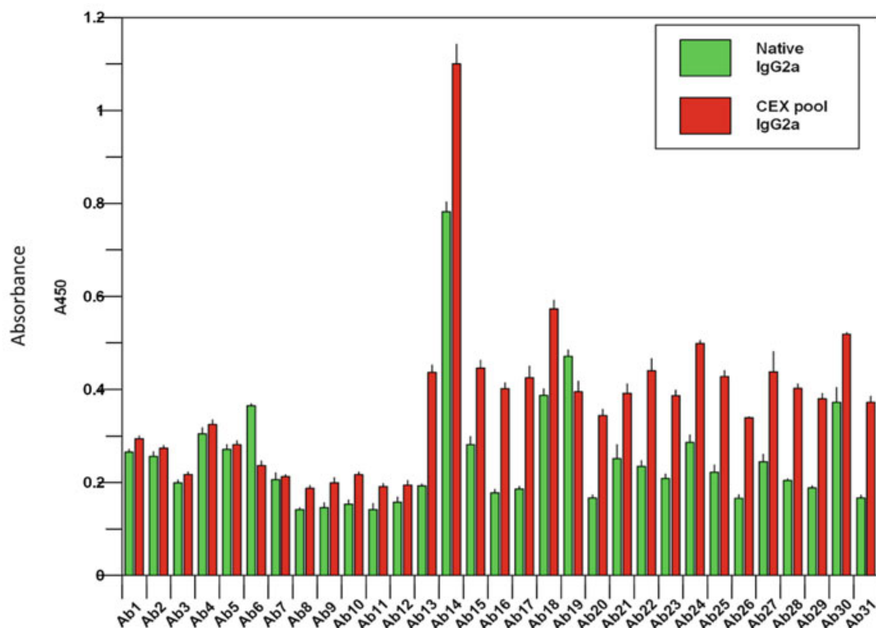


Fig. 14.9 HOS analysis of novel mAb under stress conditions during formulation development

8.0, the IgG1 sample showed significant new epitope exposure across the entire molecule with 10 day incubation to pH 3.6. Low pH (below 4) has been shown by others to cause reversible aggregation in IgG molecules, the result here correlated well with previous findings and also demonstrated that PCA ELISA could be used to characterize mAb aggregates induced by lower pH conditions (Fig. 14.9).

It is known that light exposure could potentially impact the HOS of biologics (Mason et al. 2012). During purification process, mAbs may be exposed to parts of UV-C (200–290 nm), UV-B (290–320 nm) and visible light (400–760 nm) under a variety of buffer and pH conditions. The combination of these conditions was known to promote both chemical and physical degradation which may result in conformational changes. In our case study, a sample of an IgG1 mAb which was irradiated by white light at an intensity of 416,000 lux for 52 h was compared to a control sample, this light exposure resulted in a general increase in epitope exposure with an apparent equal distribution of impact (inferred by new epitope exposure) across the entire molecule (Davies et al. 2015a). The result suggested an unfolding of a small mAb population around 1% as estimated by previous spiking studies. In addition, there is a sub-population of mAbs with specific regional changes around the hinge region (pAb19).

mAb glycosylation plays important role in its biological function, the mAb glycosylation status could impact the PK/PD of the molecule and also induce immunogenicity (Abes and Teillaud 2010; Jefferis 2009; Zheng et al. 2011), there-

fore development of mAbs with glycosylation pattern similar to those with clinical success is an important task in mAb development. Previous studies using X-ray crystallography have demonstrated that the mAb carbohydrate chains do not extend into solvent but form a bridge between the two opposing Cγ2 domains. One of the interesting aspects of glycosylation analysis is to find out the impact of different glycosylation on the Higher Order Structure of the protein. The epitope exposures of deglycosylated and control samples of an IgG2 molecule were compared in a case study. N-glycosylation of mAbs occurs within the C_{H2} region at Asn 297 (Jefferis 2007, 2009) and this epitope is represented in the PCA ELISA by pAb25. In the case study, increased epitope exposure at pAb25 was observed. There was also additional epitope exposure at pAb17 which measure epitopes in the hinge region end of the C_L domain. This result indicated that the removal of the glycosyl group caused some additional epitope exposure but not a dramatic conformational change. This is consistent with an earlier study where mAb with and without glycosylation was shown to have similar secondary as well as tertiary structure as analyzed by Fourier Transform Infrared (FTIP) spectroscopy and Intrinsic Fluorescence respectively (Zheng et al. 2011). To further evaluate the impact of different glycosylation on the HOS status of mAbs, additional studies have been planned with multiple mAbs and glycosylation types. Hopefully these future studies will provide more insight on the relationship between the mAb glycosylation and its HOS impact. With the conditions optimized for biosimilar mAb stress testing, Rituximab was analyzed under different stress conditions and with two different platforms, the ELISA-based assay and Luminex-based multiplex assay, the results were discussed in detail in section “The PCA technology adapted to Luminex xMAP® Multiplex Technology Platform” of this chapter.

HOS Polymorphism Revealed with PCA Technology and its Relevance to Immunogenicity

The classic studies by Anfinsen on ribonuclease A indicated that for small globular proteins, the native structure is determined only by the protein’s amino acid sequence (Anfinsen and Haber 1961; Anfinsen et al. 1961; Anfinsen 1973). Later studies on molecular chaperones demonstrated that, in vivo, for the correct folding of many proteins the assistance of chaperone proteins such as Hsp90, Hsp70 and Hsp60 (Langer et al. 1992; Reymond et al. 1997; Wang and Tabita 1992) are needed. Based on sequence alignment of human IgGs in the public database, it is obvious that for all the human IgGs, their constant regions are highly homologous (Wang et al. 2009). However many studies including our own suggested that factors other than the primary amino acid sequence contributed to the mAb HOS stability (Wang et al. 2013; Davies et al. 2015a). To further define the mAb HOS stability, 4 molecules contain the CH2 and CH3 domains were used and tested their refolding dynamics and HOS stability under various stress conditions, very interesting HOS

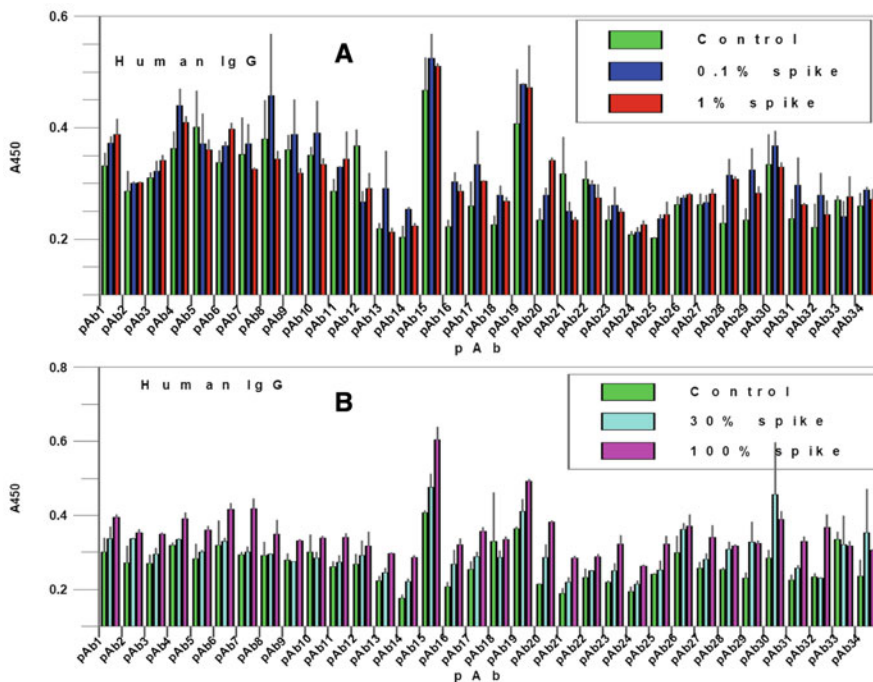


Fig. 14.10 HOS analysis of serum-derived human IgG during the refolding process

polymorphism was revealed. In the first case study, IgG molecules purified from human serum were used to test the refolding dynamics with the PCA ELISA. First, 8M urea was used to unfold the IgGs overnight and the unfolded IgG was spiked into native IgG preparation from 0.1% to 100% (100% means half of the testing IgGs at 5 μg/ml was from 8M urea-treated sample, another half was from native IgGs). The results indicated that IgGs derived from human serum has a very fast refolding process as compared with other IgGs tested (Fig. 14.10).

The results indicated that PCA ELISA will not detect obvious new epitope exposure with up to 30% spike of unfolded IgGs whereas as low as 0.1% spike of unfolded molecules could be detected from CHO-derived mAbs. Three CHO cell-derived proteins were selected to test their folding property with the PCA ELISA, the proteins include Rituximab (IgG1), Etanercept (a fusion protein of TNF-α receptor and human IgG1 CH2-CH3 domains) and an IgG2 mAb under clinical development. The testing showed that Etanercept has the fastest refolding process very similar to the serum-derived IgGs whereas mAb9 showed the slowest refolding, a 0.1% spike of unfolded mAb9 (5 ng/ml of 8M urea treated mAb into 5 μg/ml native mAb) can be quantified based on the new epitope exposure. Rituximab has intermediate rate of refolding process, a 1% spike could be quantified using the PCA ELISA (Fig. 14.11).

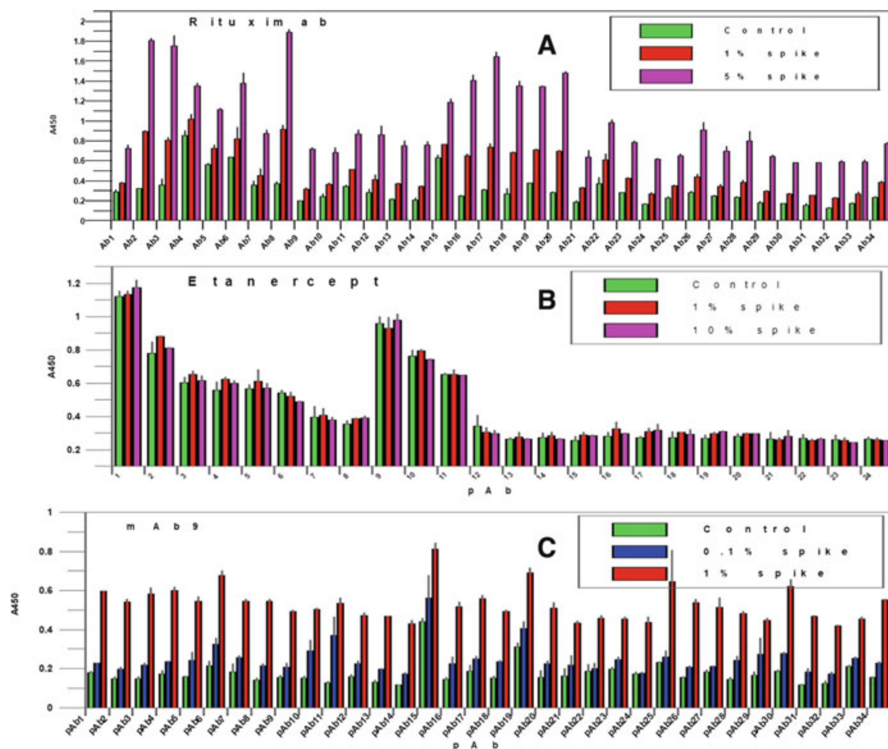


Fig. 14.11 HOS analysis of three different molecules with common CH2-CH3 domains during the refolding process

During biologics formulation development, elevated temperature has been used in accelerated stability testing to predict real world stability property of the molecule (Ohkuri et al. 2010; Ratanji et al. 2014; Zurdo 2013; Jiang et al. 2015; Meager et al. 2011; Vermeer and Norde 2000). In the case study, molecules of different modality or derived from different sources were compared for their HOS stability. It was found that Rituximab was most sensitive to elevated temperature; increased epitope exposure was detected from all the 34 capturing antibodies that were produced from 34 overlapping peptides covering the whole mAb molecule. mAb9 (an IgG2 molecule under clinical development) on the other hand demonstrated the most HOS stability at elevated temperature (Fig. 14.12).

Only minor increases (new epitope exposure) or decreases (inward movement of epitopes) of epitope change were detected from different regions of the molecule. Etanercept showed some instability in the TNF-α receptor domain whereas IgGs derived from human serum showed some instability in the hinge region and to a less extent in the light chain constant region and heavy chain CH1 domain.

The HOS stability of the four selected protein was also tested after oxidation, low pH (pH 3.0) and high pH (pH 9.5) since it is known that all these conditions could

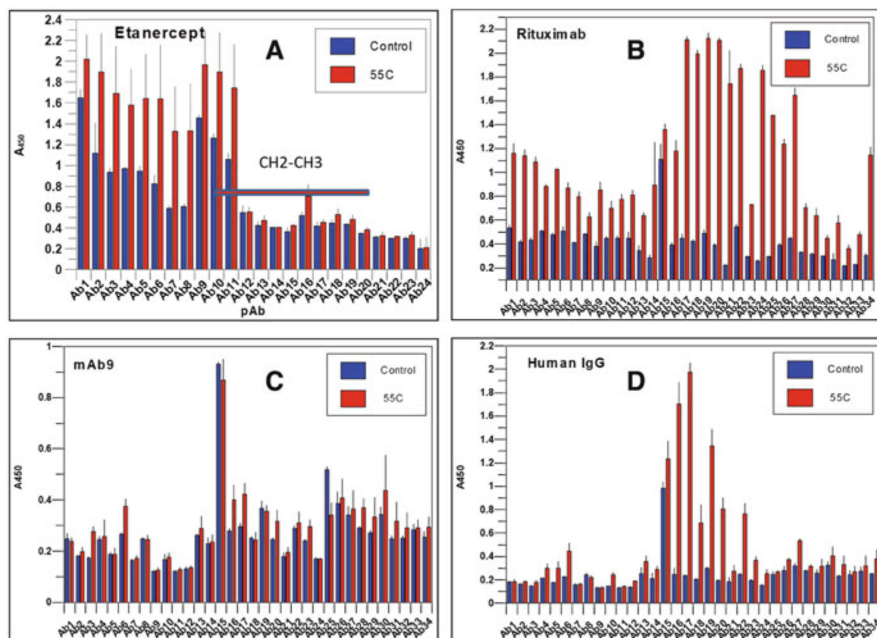


Fig. 14.12 HOS polymorphism from four different molecules with common CH2-CH3 domains under stress condition (Thermal stress at 55C)

potentially impact the HOS status of mAb molecules (Davies et al. 2015a). Because of the limit and scope of this chapter, the detailed findings in HOS changes under these stress will not be discussed in detail, further information can be found from a paper we published recently (Wang et al. 2017). One of the interesting finding in the most recent study of mAb HOS and immunogenicity showed that higher epitope exposure as measured by the PCA technology may be positively correlated with cytokine release. Another interesting finding is that the mAb refolding rate may be positively correlated with its cytokine release property, detailed findings will be discussed in a future publication.

The PCA Technology Adapted to Luminex xMAP® Multiplex Technology Platform

The Luminex xMAP technology is a magnetic beads-based multiplex system, up to 100 different addressable beads are available to be conjugated with different antibodies as initial capturing agent (Fig. 14.13).

For the PCA technology to adapt to the Luminex xMAP platform, 34 different antibodies are needed to cover the entire mAb molecule and these antibodies were

What is Luminex xMAP® Technology?

6.5 micron magnetic beads

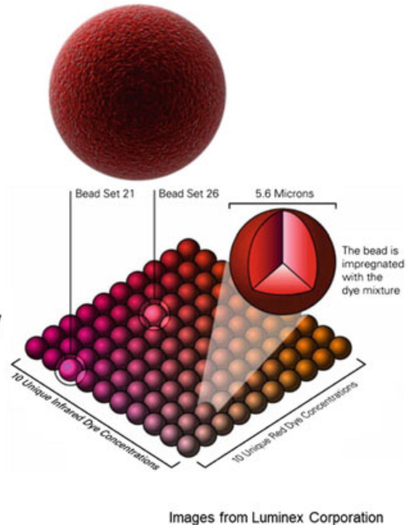
- **Liquid bead arrays**
  - Liquid kinetics = faster

Maximum of 100 different beads/well

- Bead color = spectral address
- A combination of red and infrared dyes distinguishes one bead set from another

Advantages of Luminex xMAP bead array

- Measures multiple analytes in a single well **simultaneously**
- Requires low sample volume
- High Sensitivity
- Fast, reproducible, quantifiable results



Biosimilar HOS: Multiplex vs. ELISA_18-SEP-2015

Fig. 14.13 Diagram of the Luminex xMAP technology

conjugated to 34 selected magnetic beads and tested for their capability in the detection of HOS changes and also compared with the results from the microplate-based ELISA. The testing indicated that the PCA technology can be easily adapted to the Luminex xMAP system with several distinct advantages. First, all the 34 antibodies and their captured mAb can be measured in a single well, this improve the throughput of the PCA technology tremendously. Secondly, much less samples are needed for the assay, therefore valuable samples could be saved. Thirdly, the Luminex xMAP demonstrated increased sensitivity and wider dynamic range as compared with the microplate-based assay, and finally the platform is highly automatic, large number of samples can be analyzed from a single microplate, Fig. 14.14 outlined the easy protocols of the assay.

In the initial evaluation of the xMAP technology, the sensitivity of the platform in the detection of the mAb HOS changes were tested using unfolded mAb induced from 8M urea treatment. In this test, different levels of unfolded mAb were spiked into the native mAb solution at the indicated level, and the new epitope exposure was measured by the Luminex instrument. As demonstrated in Fig. 14.14, across the 34 different antibodies, the Luminex xMAP can detect new epitope exposure as low as 0.05%, this equals to 1 ng of mAb with new epitope exposure in a 5 μg population, indicating the high sensitivity of the technology in the measurement of HOS changes for mAbs.

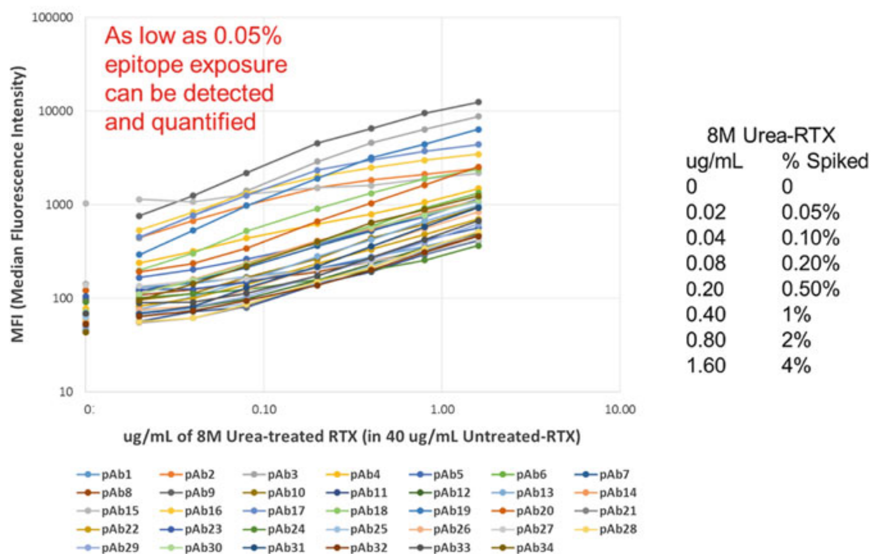


Fig. 14.14 Sensitivity of the PCA analysis for mAb HOS on the Luminex xMAP platform

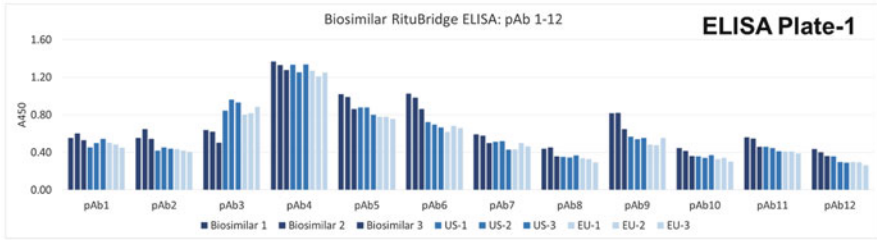
HOS Characterization of Biosimilar Rituximab with the Luminex Multiplex System

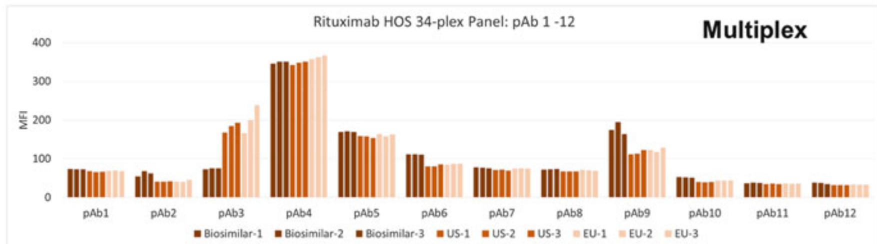
With the assay conditions on the Luminex xMAP platform optimized, samples prepared from Rituximab under different stress conditions were analyzed on the new platform and the results were compared with those from the microplate-based ELISA using the same set of samples. Figures 14.15 and 14.16 showed that both the Luminex xMAP and ELISA have very similar HOS profile for the control Rituximab as well as the oxidized molecule respectively.

In the further testing with additional stress conditions, it showed that the Luminex xMAP actually demonstrated wider dynamic range as compared to the ELISA. For the pH3-treated samples, the signal in the ELISA was saturated whereas on the Luminex xMAP, the sample's epitope exposure can still be quantified (Fig. 14.16).

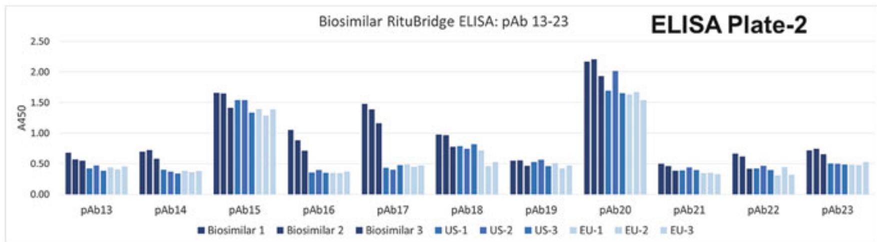
The Luminex-based PCA technology has been used successfully in the analysis of mAb HOS comparability with advantages in throughput, automation and significantly reduced cost. It is especially valuable in the analysis of bioprocess and formulation development samples where large numbers of samples are generated daily and the timely feedback on the process development could significantly impact the progress of the mAb of interest.

ELISA vs. Multiplex: pAb 1-12 (Variable Regions)





ELISA vs. Multiplex: pAb 13-23 (Constant Regions-1)



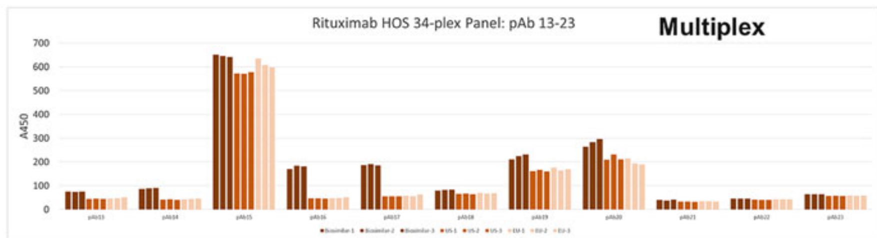


Fig. 14.15 (**a–c**) HOS comparability analysis with the microplate-based ELISA and the Luminex xMAP platform

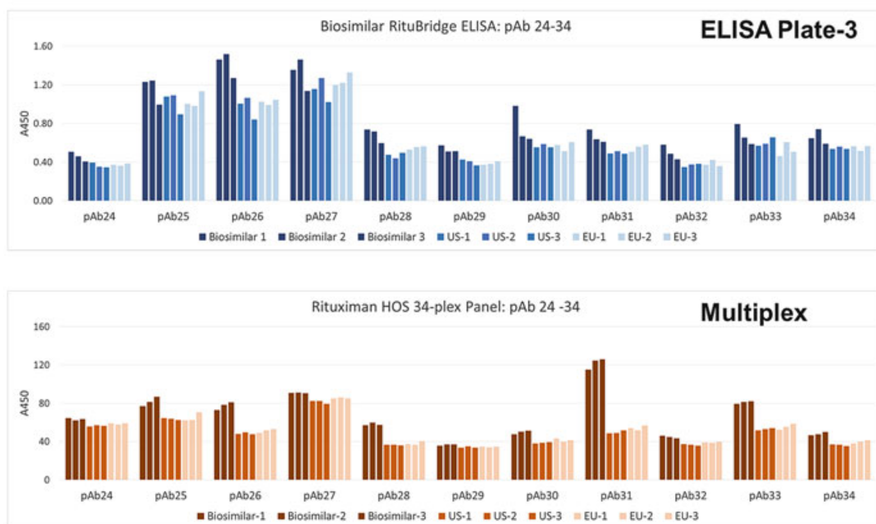


Fig. 14.15 (continued)

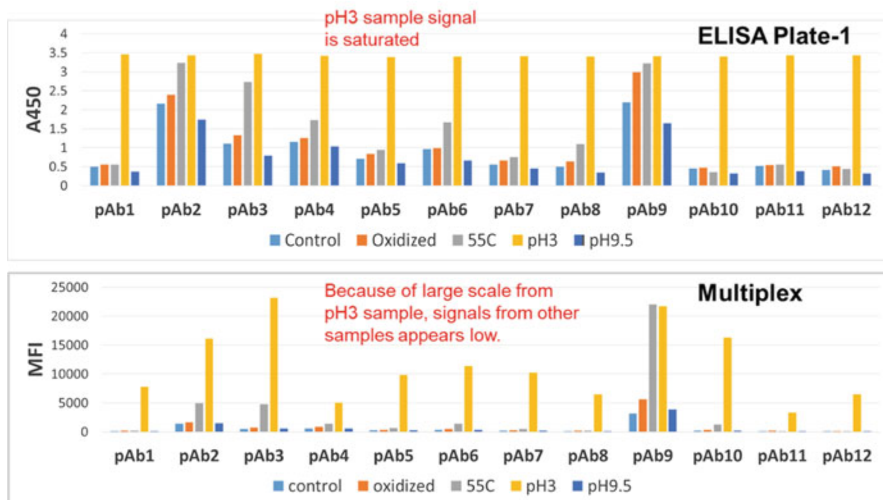


Fig. 14.16 Luminex xMAP technology demonstrated wider dynamic range and high automation compared with the ELISA-based mAb HOS analysis

Conclusion

The two PCA platforms discussed in this chapter, offer high sensitivity in the detection of new epitope exposure induced by physical as well as chemical conditions typically encountered during biosimilar development. The implementation of the PCA technology during biosimilar development can offer valuable insights on the quality of the bioprocess and formulation development and provide systematic and molecular level comparability of the biosimilar candidate as compared with the reference standards from the innovator. Since the PCA technology is a multifaceted analysis of the mAb molecule, HOS changes from almost all the stress conditions can be detected and quantified with high sensitivity and accuracy (Wang et al. 2017; Davies et al. 2015). This analysis, together with other biophysical technologies used for HOS analysis, will make sure the development of the biosimilar mAb is on the right track and eventually bring a successful product to the market.

References

Abes R, Teillaud J. Impact of glycosylation on effector functions of therapeutic IgG. Pharmaceuticals. 2010;3:12.

Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973;181(4096):8.

Anfinsen CB, Haber E. Studies on the reduction and re-formation of protein disulfide bonds. J Biol Chem. 1961;236:3.

Anfinsen CB, et al. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. PNAS. 1961;47(9):6.

Arosio P, et al. Aggregation stability of a monoclonal antibody during downstream processing. Pharm Res. 2011;28(8):11.

Arthur KK, Dinh N, Gabrielson JP. Technical decision making with higher order structure data: utilization of differential scanning calorimetry to elucidate critical protein structural changes resulting from oxidation. J Pharm Sci. 2015;104:6.

Beck A, et al. Characterization of therapeutic antibodies and related products. Anal Chem. 2014;85:715–36.

Berkowitz SA, et al. Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. Nat Rev Drug Discov. 2012;11:14.

Bessa J, et al. Immunogenicity of antibody aggregates in a novel transgenic mouse model. Pharm Res. 2015;32:16.

Buttel I, et al. Taking immunogenicity assessment of therapeutic proteins to the next level. Biologicals. 2011;39:10.

CBER/CDER. Guidance for industry on quality considerations in demonstrating biosimilarity to a reference protein product. Rockville: US Food and Drug Administration; 2015.

Chang J, Lu B, Li L. Conformational impurity of disulfide proteins: detection, quantification, and properties. Anal Biochem. 2005;342:8.

Chaudhari PS, Nath R, Gupta SK. Opportunities and challenges in biosimilar development. Bioprocess Int. 2017;15(5):10.

Davies M, et al. mAb higher order structure analysis with protein conformational array ELISA. Br J Pharma Res. 2015;7(6):12.

Davies M, et al. Biosimilar mAb in-process sample higher order structure analysis with protein conformational array ELISA. Br J Pharma Res. 2016;9(3):11.

DiPaola M. Analytical strategies in the development of biosimilars. BioPharm Int. 2017;30:10.
Gabrielson J, Weisis WF. Technical decision-making with higher order structure data: starting a new dialogue. J Pharm Sci. 2015;104:6.
Hermeling S, et al. Structure-immunogenicity relationships of therapeutic proteins. Pharm Res. 2004;21(6):7.
Hermeling S, et al. Structural characterization and immunogenicity in wild-type and immune tolerant mice of degraded recombinant human interferon alpha2b. Pharm Res. 2005;22(12):10.
Hermeling S, et al. Antibody response to aggregated human interferon alpha2b in wild-type and transgenic immune tolerant mice depends on type and level of aggregation. J Pharm Sci. 2006;95(5):13.
Jefferis R. Antibody therapeutics: isotype and glycoform selection. Exp Opin Biol Therap. 2007;7(9):13.
Jefferis R. Glycosylation as a strategy to improve antibody-based therapeutics. Nat Rev Drug Discov. 2009;8:9.
Jiang Y, et al. Technical decision making with higher order structure data: higher order structure characterization during protein therapeutic candidate screening. J Pharm Sci. 2015;104:6.
Jiskoot W, et al. Immunological risk of injectable drug delivery system. Pharm Res. 2009;26(6):12.
Jung S, et al. Physicochemical characterization of Remsima. MAbs. 2014;6(5):15.
Laat B, et al. Immune response against domain I beta2-glycoprotein I are driven by conformational changes. Arthritis Rheum. 2011;63(12):9.
Langer T, et al. Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. EMBO J. 1992;11(13):9.
Maas C, et al. A role for protein misfolding in immunogenicity of biophmaceuticals. J Biol Chem. 2007;282:8.
Mason D, Schoneich C, Kerwin B. Effect of pH and light on aggregation and conformation of an IgG1 mAb. Mol Pharm. 2012;9(4):17.
Meager A, et al. An assessment of biological potency and molecular characterics of different innovator and noninnovator interferon-beta products. J Interf Cytokine Res. 2011;31(4):10.
Mok CC, et al. Immunogenicity of anti-TNF biologic agents in the treatment of rheumatoid arthritis. Exp Opin Biol Therap. 2016;16:11.
Ohkuri T, et al. A protein's conformational stability is an immunologically dominant factor: evidence that free-energy barriers for protein unfolding limit the immunogenicity of foreign proteins. J Immunol. 2010;185:16.
Porter S. Human immune response to recombinant human proteins. J Pharm Sci. 2001;90(1):11.
Ratanji KD, et al. Immunogenicity of therapeutic proteins: influence of aggregation. J Immunotoxicol. 2014;11:11.
Reymond M, et al. Folding propensities of peptide fragments of myoglobin. Protein Sci. 1997;6:11.
Rosenberg A. Effects of protein aggregates: an immunologic perspective. AAPS J. 2006;8(3):8.
Schellekens H. Factors influencing the immunogenicity of therapeutic proteins. Nephrol Dial Translant. 2005;20:7.
Sharma B. Immunogenicity of therapeutic proteins. part 1: impact of product handling. Biotechnol Adv. 2007;25:8.
Thiagarajan G, et al. A comparison of biophysical characterization techniques in predicting monoclonal antibody stability. MAbs. 2016;8(6):10.
US FDA. Draft guidance for industry on quality condiserations in demonstrating biosimilarity to a reference protein product., in Fed. Regist. 2012.
Vermeer A, Norde W. The thermo stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. Biophys J. 2000;78:11.
Wang X, Tabita FR. Interaction of inactivated and active ribulose 1,5-bisphosphate carboxylase/oxygenase of rhodobacter sphaeroides with nucleotides and the chaperonin 60 (GroEL) protein. J Bact. 1992;174(11):5.
Wang X, et al. Potential aggregation prone regions in biotherapeutics. MAbs. 2009;1:14.

Wang X, Li Q, Davies M. Development of antibody arrays for monoclonal antibody higher order structure analysis. Front Pharmacol. 2013;4:8.

Wang X, Li Q, Davies M. Higher order structure comparability: case studies of biosimilar monoclonal antibodies. Bioprocess Int. 2014;12:6.

Wang X, et al. Antibody higher order structure polymorphism revealed by protein conformational array. Bioprocess Int. 2017;15(10):8.

Wei Z, et al. The role of higher-order structure in defining biopharmaceutical quality. Bioprocess Int. 2011;9(4):9.

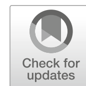
Weiss WF, Young TM, Roberts CJ. Principles, approaches and challenges for predicting protein aggregation rates and shelf life. J Pharm Sci. 2009;98:12.

Weiss WF, et al. Technical decision making with higher order structure data: perspectives on higher order structure characterization from the biopharmaceutical industry. J Pharm Sci. 2016;8:6.

Zheng K, Bantog C, Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. MAbs. 2011;3(6):9.

Zurdo J. Developability assessment as an early de-risking tool for biopharmaceutical development. Pharm Bioprocess. 2013;1:21.

Chapter 15
Protein Particulates and Biosimilar Development: Analytical Tools and Therapeutic Implications



Aaron B. Krueger and Matthew D. Brown

Abstract Particulate content of parenteral products represents one of the major challenges during the development and commercialisation of safe biotherapeutics. Indeed, the presence of particles is one the top 10 reasons for product recalls. The risk of immunogenicity and adverse clinical reactions in patients has resulted in subvisible particles becoming one of the major focus topics for regulatory agencies. The US FDA has issued several guidance documents in the recent past on subvisible particles, immunogenicity, quality and technical aspects of biosimilars. These guidance documents make the characterization of subvisible and submicron particles relevant and important for both novel biotherapeutics and biosimilar therapeutics. Significant advances have been made in analytical technologies, improving the detection, quantification, and characterisation of particles from the nm range up to 100 μm plus. With this improvement in analytical tools, there is an increasing expectation from regulatory agencies for sponsors to provide more robust subvisible particle characterisation along with risk assessment. Understanding the particulate content of biotherapeutics provides a unique challenge in the Biopharmaceutical industry. Although regulatory requirements for biotherapeutics filing, demand compliance with USP <788>, recent instances have demonstrated the serious consequences of performing only limited particle characterization. Characterizing the subvisible and submicron particles in biosimilars is extremely critical from an immunogenicity and safety perspective. In the coming years, characterization of subvisible particles will continue to play a crucial role in biosimilar development and approval.

A. B. Krueger
KBI Biopharma Inc., Boulder, CO, USA
e-mail: akrueger@kbibiopharma.com

M. D. Brown (✉)
Malvern Panalytical, Westborough, MA, USA
e-mail: matthew.brown@malvern.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_15

Keywords Protein aggregates · Subvisible particles · Immunogenicity · Submicron particles · Guidance · Light Obscuration · Flow Imaging Microscopy (FIM) · Nanoparticle Tracking Analysis (NTA) · Dynamic Light Scattering (DLS) · Resonance Mass Measurement (RMM) · Morphology-Directed Raman Spectroscopy (MDRS) · Microfluidic Resistive Pulse Sensing (MRPS)

Introduction

Significant advances in analytical technology over the past few years have improved the quantification and characterization capabilities for subvisible (1–100 μm) and submicron (≤1 μm) particles. As the technology continues to improve, so do the expectations of regulatory agencies for sponsors to characterize particles in these size ranges. However, multiple orthogonal methods are required to span the entire range and accurately characterize the particle profile. Each instrument has its own limitations based on detection method and properties of therapeutic protein products that must be well understood to generate high quality data.

The quantification of subvisible particles in injectable therapeutic protein products has been established in the US and European Pharmacopeia under Chapters 788 and 2.9.19, respectively. In accordance with these rules and procedures, limits have been defined for subvisible particulate content that were based on the risk of ≥10 μm and ≥25 μm particles blocking certain percentages of blood vessels in the lungs following intravascular infusion. This has marked a shift in the industry towards more convenient administrations, such as subcutaneous and intramuscular injections. The move has also opened a discussion on the relevance of the pharmacopeia chapters, specifically those relating to protein therapeutics. Over the past decade, a substantial amount of work has been published that investigates the propensity of proteins to aggregate and form subvisible particulates (Simler et al. 2012; Singh et al. 2010; Barnard et al. 2012), and the potential risk of immunogenicity due to presence of protein aggregates and subvisible protein particles in therapeutic protein products (Filipe et al. 2012; Fradkin et al. 2009; Fradkin et al. 2011).

It is generally accepted that protein particles form during degradation processes. These degradation processes may involve perturbation of protein structure or the formation of reversible native aggregates (Fig. 15.1), but the occurrence of small irreversible aggregates is considered to act as nucleation sites for the formation of submicron-sized protein particles. These submicron protein particles, even at very low concentrations, have been shown to promote the formation of larger particles (Shujun et al. 2013). This understanding highlights the importance of detecting, sizing, and characterizing particles from high micron size right down to the nanometer scale and soluble aggregates. Quantifying particles across the entire size range allows not only a risk assessment of the product in its current state, but also an assessment of future risk. Studies have shown that relying on SEC alone, is not sufficient to accurately quantify and size protein aggregates (Carpenter

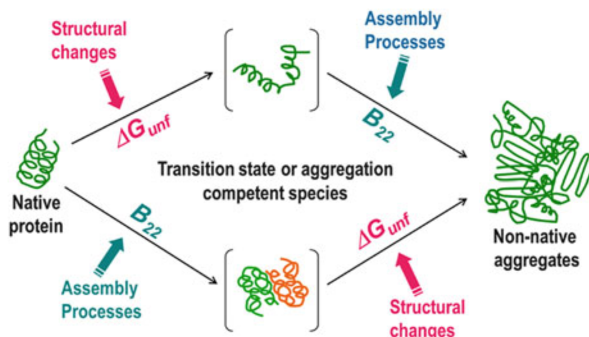


Fig. 15.1 Protein degradation pathways and the formation of subvisible particles. Native proteins can undergo structural changes to form misfolded intermediates or form native, self-associated aggregates. Either pathway can lead to the formation of non-native, irreversible aggregates

Fig. 15.2 Commonly accepted definitions of particulate matter, based on their source

et al. 2010), and orthogonal methods are important for a full assessment of product quality.

Traditionally, there was an assumption that particles in parenteral products were always protein particles. With the development of novel technologies, the capability now exists for companies to not only quantify particles, but also identify the composition and therefore, the source. Particles are generally classified as being Extrinsic, Intrinsic or Inherent (Fig. 15.2), depending on their source.

Inherent particles are derived from the product itself, and so are formed of protein aggregates and/or buffer components. Their presence is a key indicator of formulation and product stability. Intrinsic particles are derived from the process and container closure. Such particles may be shed from equipment or leached from product contact surfaces. Some of the most common intrinsic particles are stainless steel, glass, rubber, and silicone oil. Silicone oil has gained significant interest due to the development of pre-filled syringes, but aggregates of excipients is also becoming a major topic for understanding. Finally, extrinsic particles are from outside of the process and are indicative of a failure of containment and sterile fill/finish. These particles pose the greatest risk to sterility assurance (Bukofzer et al. 2015) and their presence may lead to disciplinary action from regulatory agencies. Of course, these particle classifications are not mutually exclusive, and particles may contain material from different sources, such as protein aggregates forming around silicone oil droplets.

More recently, an additional dimension to product stability has come in the form of product handling in the clinic. Whilst companies direct considerable cost and

resources to ensure the product satisfies shelf life and Critical Quality Attribute (CQA) requirements, limited data is available on stability of products at point of use. Nearly 70% of mAb-based therapeutics are administered via intravenous infusion systems (An 2011) which can vary in terms of IV setup. Diluent composition, IV bag headspace, and time have all been shown to affect the formation of particles, from submicron up to visible (Kumru et al. 2012). Many particles are shed from the polycarbonate material, and even with in-line filters in place, many particles are present in the final administered product (Pardeshi et al. 2017). Consequently, handling of the product at point of use can impact particle formation and, therefore, immunogenic risk. Two key aspects underpinning this issue are clear product handling guidelines and education of clinical staff in following these guidelines.

Companies are now being urged to use orthogonal approaches and techniques for assessing the CQA's of the protein therapeutic product. Nowhere is this more evident than when quantifying particles, where there is a strong emphasis to extend beyond USP <788> and standard pharmacopoeial methods. The FDA's "Guidance for Industry" on Immunogenicity (US FDA 2014) states "the use of any single method for assessment of aggregates is not sufficient to provide a robust measure of protein aggregation". Industry must now utilize multiple technologies to span the entire sub-visible/visible particle size range, from 2 μm up, ensuring the strengths and weaknesses of each technology are fully understood.

Immunogenicity of Biotherapeutics: Evidence and Therapeutic Consequences

From the onset, a persistent concern for biotherapeutic use has been unwanted immunogenicity. The FDA defines immunogenicity as "the propensity of the therapeutic protein product to generate immune responses to itself and to related proteins or to induce immunologically related adverse clinical events (US FDA 2014)". Some of the earliest use of biotherapeutics showed cases of associated immunogenicity. When insulin from porcine and bovine sources first began to be used to treat diabetes mellitus, immunogenic responses were observed. This included anti-drug antibody (ADA) formation and in some cases led to fatal anaphylactic reactions (Schernthaner 1993). With further purification of the products, incidents of adverse reactions were reduced and provided evidence that impurities, including aggregates, are linked with immunogenicity.

As early as the 1960s, nonclinical studies indicated a link between therapeutic protein aggregates and immunogenicity. Centrifugation studies with bovine gamma globulin revealed that removal of high molecular weight fractions was required to prevent immunological paralysis in mice (Dresser 1962; Claman 1963) and when heat-induced aggregates of human gamma globulin (HGG) were injected in mice a dose-dependent immune response was demonstrated (Gamble 1966). Further studies with human interfon-α (IFN-α) transgenic mice demonstrated that immune

tolerance to self-proteins can be broken by the presence of aggregates and that the type of aggregate is important for an immune response (Hermeling et al. 2005; Braun et al. 1997).

Whilst most biological products will elicit some immune response in patients, in most cases this has no relevant clinical consequences (Kessler et al. 2006). In the rare cases where there is a significant clinical impact, these can be severe or even lethal. It was discovered in the 1950s and 1960s that aggregated material in HGG products caused severe anaphylactic reactions in patients (Barandun et al. 1962), with the antibody response directed towards the high molecular weight species (i.e. the pellet upon ultracentrifugation) (Ellis and Henney 1969). Similar studies demonstrated anaphylactic responses of patients treated with Human Serum Albumin (HSA) preparations or pasteurized plasma solutions could be ascribed to an antibody response directed towards the aggregates, which were found to be at levels of 5–15% in those products (Ring et al. 1979). Therapeutic proteins of human origin, chosen with a desire of reducing unwanted immunogenicity, included hormones and clotting factors obtained from cadaver pituitary glands. Even though these were of human source, they retained a strong antibody response in patients, which was attributed to the high levels of impurities in the products (Milner 1985; Jacquemin and Saint-Remy 1998). In the 1960s, therapeutic human growth hormone (hGH) preparations contained between 40 and 70% aggregates and triggered antibody responses in up to 50% of patients. Reduction of aggregates to 5–10% through optimization of purification processes lead to a substantial reduction, but not elimination, of anti-drug antibody formation (Underwood et al. 1974; Moore and Leppert 1980). Furthermore, IL-2 containing high levels of small aggregates can result in high rates of immunogenicity, with as much as 60% of patients developing antibody responses, potentially leading to the appearance of neutralizing antibodies with concomitant treatment with IFN-α. However, in vitro prediction of neutralization did not translate in vivo neutralization (Rosenberg 2006; Prümmer 1997).

The most common biological consequence of immunogenicity is loss of efficacy, as has been described for interferon products (Patten and Schellekens 2003). Neutralizing antibodies are prevalent in interferon-beta products currently on the market and clinical data for Betaseron®, Rebif®, and Avonex® demonstrates levels of 27.8–47%, 5.3–35%, and 2–7.5%, respectively (Bertolotto et al. 2004; Grossberg et al. 2011). Analyses of these products demonstrate that there is a direct correlation between aggregate or particle content and clinical rates of immunogenicity, however many other factors may be affecting adverse immunogenicity. This underlies the discussion whether biosimilars should have the same high aggregate and particle content as the innovator products or should they meet current product quality expectations. Overall with these studies, it appears that for products that are immunogenic, trace amounts of particles or aggregates may play a role. Major clinical impact has been observed when natural proteins with essential biological activity are neutralized, as in the cases of thrombocytopenia (Neumann and Foote 2000) and pure red cell aplasia (Casadevall et al. 2002). Other clinical consequences, such as anaphylaxis, have become less common with the development of highly purified products.

The factors leading to an increased immunogenicity for a biotherapeutic product remain generally unclear or contradictory. The introduction of recombinant DNA technologies and sequencing the human genome ensured that most modern therapeutic proteins are at least human homologues. Typically, the human immune system reacts to the sequence variation of nonhuman proteins to elicit a response (Chaffee et al. 1992; Rosenschein et al. 1991; Grauer et al. 1994). However, immunogenicity can still occur with recombinant proteins with human sequences (Schellekens 2002) while failing to increase levels of immunogenicity with therapeutic proteins containing sequences different from naturally occurring human proteins (Girard and Gourmelen 1986; Kontsek et al. 1999). Because of this, most biotherapeutic products can still induce ADA formation and often in the majority of patients. It is extremely rare for a therapeutic protein after repeated administrations to have no reported cases of immunogenicity. Furthermore, product modifications such as de-glycosylation of glycoproteins (Karpusas et al. 1998; Gribben et al. 1990) or posttranslational modifications (Prümmer 1997; Antonelli et al. 1997) may lead to differences in immunogenicity rates. Similarly, PEGylated proteins may confer slight improvements in immunogenicity (Schellekens 2008; Veronese and Pasut 2005), yet cases exist where immunogenicity levels are elevated with a PEGylated product (Vadhan-Raj 2000). Formulation components may also contribute to the immunogenicity. As seen previously, HSA is a common stabilizing agent added to biotherapeutic formulations and is often associated with an increase in immunogenicity (Palleroni et al. 1997; Christie et al. 2015).

FDA Guidance on Subvisible Particles, Immunogenicity, and Biosimilar Development

Immunogenicity is currently viewed by industry experts as one of the major safety challenges associated with biotherapeutics. Assessment of product immunogenic risk is a key aspect of product development and approval. As stated by the FDA, “All therapeutic protein products should be evaluated for their content of and immune responses directed to incidental product components, including proteins and nonprotein components (US FDA 2014)”. Research on principles of immunological processes has shown that large molecular weight species containing repetitive antigens of native conformation have the greatest potential to illicit an immune response (Rosenberg 2006). In addition, antigenic proteins absorbed onto the surface of nano or micro sized particles can be particularly immunogenic (Lebron et al. 2007), potentially due to particulate antigens being rapidly captured and presented to the immune system by dendritic cells (Martin et al. 2001). Therefore, the presence of both protein and contaminant particles are considered as key contributors to the development of immunogenicity by parenteral products (Carpenter et al. 2009).

As a result of these reports, the FDA has become increasingly concerned with the safety and efficacy of therapeutic protein products as, in the past, subvisible particles between 0.1 and 10 μm were not being actively monitored. To begin addressing this concern, a new US Pharmacopeia (USP) monograph, USP <787>: Subvisible Particulate Matter in Therapeutic Protein Injections, was drafted and became effective in 2014. In this monograph, in addition to allowing smaller particle size detection reporting, several improvements such as reductions in required testing volumes were implemented. For example, while the previous USP <788> required a minimum volume of 25 mL to complete a test, the new Chapter allows for a volume as low as 1 mL. This small volume method not only saves costs by requiring less product, but also allows for the determination of vial-to-vial or syringe-to-syringe variability of particle counts due to no longer needing to pool containers for analysis. Furthermore, although the reporting of ≥10 and ≥25 μm size ranges is still required, the establishment of specific container-limits is permitted.

As well as USP <787>, a general information USP <1787>: Measurement of Subvisible Particulate Matter in Therapeutic Protein Injections, was drafted. This chapter recommends the collection of 2–10 μm (≥2 and ≥5 μm) subvisible particle concentrations, and gives guidance on orthogonal methods to characterize subvisible particles as inherent, intrinsic, or extrinsic, and specifically spells out the need to distinguish silicone oil from inherent particles, or other intrinsic particles. The FDA has also approved Guidance for Industry on the Immunogenicity Assessment for Therapeutic Protein Products that states: "[assessment] should be made of the range and levels of subvisible particles (2–10 μm) present in therapeutic protein products initially and over the course of shelf-life [...] As more methods become available, sponsors should strive to characterize particles in smaller (0.1–2 μm) size ranges (US FDA 2014)". Meanwhile, the EMA's Guideline on Development, Production, Characterisation and Specifications for Monoclonal Antibodies and Related Products states: "[the] formation of aggregates, subvisible and visible particulates in the drug product is important and should be investigated and closely monitored on batch release and during stability studies. In addition to the pharmacopoeial test for particulate matter, other orthogonal analytical methods may be necessary to determine levels and nature of particles (EMA 2008)".

As regulatory agencies start to request additional particle detection, specifically below 10 μm, it becomes imperative for companies to fully understand the orthogonal methods available and their limitations. This is particularly important as covering the entire particle sizing range requires the use of several different technologies (Fig. 15.3). As discussed in USP <1787>, particles in the ≤10 μm size range can be highly variable due to other inherent or intrinsic components, such as silicone oil. Therefore, combining orthogonal techniques to help identify and characterise these particles becomes essential for data analysis and interpretation.

When considering Biosimilars specifically, these products must satisfy the same release tests as any innovator parenteral product, USP <787> and USP <1787>. However, during Biosimilar development there is the need to demonstrate similarity

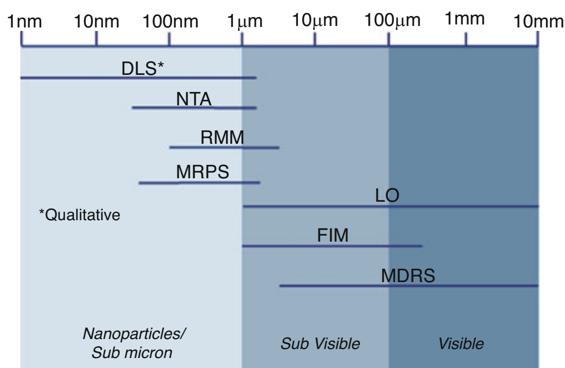


Fig. 15.3 Particulate matter size range and associated instrumentation. *DLS* dynamic light scattering, *NTA* nanoparticle tracking analysis, *RMM* resonant mass measurement, *MRPS* Microfluidic Resistive Pulse Sensing, *LO* light obscuration, *FIM* flow imaging microscopy, *MDRS* morphology directed Raman spectroscopy

and present such data to the regulatory agencies, adding the additional aspect of characterizing particulate content in order to assess similarity between the innovator and the biosimilar. The use of the term "totality of evidence" in assessing similarity data means no individual set of data or characterization is sufficient to ensure success or failure of a submission. However, a review of successful applications can provide an indication as to what data should be included to satisfy the similarity criteria. Table 15.1 summarizes the particle characterization methods used for similarity assessment of some biosimilars recently recommended for approval in US. Most cases used flow imaging microscopy methods and light obscuration to quantify particles above 2 μm, while Amgen's submission for ABP501 also include the use of DLS for qualitative assessment of submicron particles. Only the submission for GP2015 (June 6th, 2016), did not include subvisible particle analysis for similarity assessment; however, since GP2015 is a sterile liquid product it would still be required to comply with USP <788> and Ph Eur. 2.9. 19. Therefore, although analysis of subvisible particles appears to be an important component of the "totality of evidence" towards biosimilarity, the size ranges reported are not consistent and the ability to positively identify the composition of all sizes of particles is absent.

Instrumentation and Characterization of Subvisible Particles (2–100 μm): Light Obscuration and Flow Imaging Methods

Light Obscuration

Light Obscuration (LO) is a highlighted method in the USP Monographs for identification of subvisible particles with limits defined in several of the chapters. LO works on the principle of light blockage or light extinction. Typical instrumentation contains a syringe to automatically withdraw solution from a sample container

Table 15.1 Summary of subvisible particle measurement methodologies applied for similarity assessment of several biosimilar products that were recommended for approval by FDA advisory committees between 2015 and 2017

Product	Method(s)	Results, similarity criteria (Yes/No)	Methodology details	Rationale
ZARXIO® (filgrastim biosimilar)	FIM using Micro Flow Imaging (MFI)	Measurements of proteinaceous subvisible particles by MFI demonstrated a lower number of subvisible particles in ZARXIO® as compared to Neupogen®	Details not provided in ADCOM briefing document	Both products are marketed in prefilled syringes, adequate approach to measure and control subvisible particles. MFI can distinguish between silicone oil and other particle types
CT-P13 (infliximab biosimilar)	LO using HIAC	Biosimilarity CT-P13 vs US (highly similar)	The system suitability requirements were based on USP <788>	Subvisible particle count by MFI was used for similarity assessment. The % aggregates as measured by SE-HPLC and SEC-MALS were higher for CT-P13, however this approach demonstrated that the biosimilar had similar subvisible particle count like both US and EU reference products
	FIM using MFI	Biosimilarity CT-P13 vs US (highly similar)	The MFI 5200 model used can detect SVP in the 1–100 μm size range. The MFI equipment was calibrated using particle sizes of 2, 5, and 10 μm and system suitability requirements of ±10% were applied	

(continued)

Table 15.1 (continued)

Product	Method(s)	Results, similarity criteria (Yes/No)	Methodology details	Rationale
GP2015 (Etanercept® biosimilar)	Method details unavailable	N/A	N/A	The product GP2015 is proposed to be available in both prefilled syringe and autoinjector forms, a method to measure subvisible particles and differentiate silicone oil like particles should have certainly been in place. However, the method details were not shared in ADCOM materials
ABP 501 (biosimilar adalimumab)	LO using HIAC	Yes similar, Tier 3 criteria—qualitative comparison	$\geq$2 μm particles $\geq$5 μm particles $\geq$10 μm particles $\geq$25 μm particles measured and classified	Compendial test requirement for parenteral product, proposed use as a control strategy
	FIM using MFI: $\geq$5 μm particles	Yes similar, Tier 3 criteria—qualitative comparison	Details not provided in ADCOM briefing document	Both products are marketed in prefilled syringes, adequate approach to measure and control subvisible particles. MFI can distinguish between silicone oil and other particle types
	FIM using MFI: $\geq$5 μm non-spherical particles	Yes similar, Tier 2—quality range Range for Humira®: 0–197 Range for ABP 501: 24–172 (For Tier 2 attributes using a quality range, analytical similarity was demonstrated if at least 90% of the individual test lot values fell within the adalimumab derived quality range)	Details not provided in ADCOM briefing document (a filter to differentiate between spherical particle like silicone oil or air bubble and non-spherical particle would have been selected)	Since non-spherical particles are essentially protein like particles, the approach specifically identifies and categorizes such particles. Quality range criteria is applied to adequately control such non-spherical (protein like) particles in the biosimilar product

HOSPIRA EPOETIN® (Biosimilar to EPOGEN®/ PROCRIT®—Epoetin alfa)	FIM using MFI: (≥25 μm, ≥10 μm, ≥5 μm and ≥2 μm)	Levels five to tenfolder lower in Epoetin® Hospira Yes—qualitative comparison only, no assessment cirteria/quality range applied	Details not provided in ADCOM briefing document. The method can distinguish silicone oil from non-spherical particles. *The Epoetin® biosimilar is formulated without plasma derived human serum albumin (HAS), which could add significant subvisible particles to the product	The five to tenfold lower subvisible particles are probably due to use of a different formulation compared to reference product which is devoid of human serum albumin (HSA)
	Nanoparticle tracking analysis using NanoSight (0.1–1 μm)	Levels significantly lower in Epoetin® Hospira Yes—qualitative comparison only, no assessment criteria/quality range applied	Details not provided in ADCOM briefing document	
ABP 215, (biosimilar to Avastin® bevacizumab)	LO using HIAC	Yes similar, Tier 3 criteria—qualitative comparison	≥2 μm particles ≥5 μm particles ≥10 μm particles ≥25 μm particles measured and classified	Compendial test requirement for parenteral product, proposed use as a control strategy
	FIM using MFI: ≥5 μm particles	Yes similar, Tier 3 criteria—qualitative comparison	The system characterizes particles through liquid sampling, image acquisition, and image analysis. Cumulative particle counts per mL for ≥5 μm particles were reported	Both products are marketed in vials with a fill volume of 4 mL and 16 mL. Rubber stoppers for vials are generally siliconized. MFI can distinguish between silicone oil and other particle types. Adequate approach to measure and control subvisible particles

(continued)

Table 15.1 (continued)

Product	Method(s)	Results, similarity criteria (Yes/No)	Methodology details	Rationale
	FIM using MFI: $\geq 5\ \mu m$ non-spherical particles	Yes similar, Tier 2—quality range Range for Avastin: 0–3533 Range for ABP 215: 0–65 (For Tier 2 attributes using a quality range, analytical similarity was demonstrated if at least 90% of the individual test lot values fell within the bevacizumab derived quality range)	To quantify product-related particles that are likely proteinaceous and thus have a higher risk for immunogenicity, the MFI data were further analyzed for the concentration of $\geq 5\ \mu m$ non-spherical particles with an aspect ratio of <0.85	Since non-spherical particles are essentially protein like particles, the approach specifically identifies and categorizes such particles. Quality range criteria is applied to adequately control such non-spherical (protein like) particles in the biosimilar product
MYL-1401O (biosimilar to trastuzumab)	FIM using MFI: $\geq 5\ \mu m$ particles	Yes—highly similar *Particle counts were separately assessed in the following size ranges: $\geq 2\ \mu m$, $\geq 5\ \mu m$, $\geq 10\ \mu m$, and $\geq 25\ \mu m$. Similarity performed for all particles $\geq 5\ \mu m$	MYL-1401O ($n = 12$), US-Herceptin® ($n = 6$), and EU-Herceptin® ($n = 6$) were highly similar at 5 μm and above, however, in the 2–5 μm range, higher counts were observed for 4 lots of MYL-1401O lots. These particles were characterized and found to be non-proteinaceous and thus not of concern with respect to immunogenicity	Both products are marketed as lyophilized powders for reconstitution in vials. Rubber stoppers for vials are generally siliconized. MFI can distinguish between silicone oil (circular) and other non-circular particle types. Adequate approach to measure and control subvisible particles

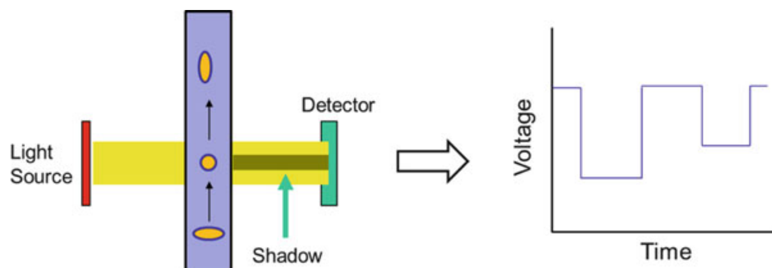


Fig. 15.4 Light obscuration analysis and determination of particle size and concentration

that passes the sample, at desired flow rates, through a sensor to count particles (Fig. 15.4). Under ideal particle concentrations, optimized by syringe choice and flow rate, individual particles are passed through the sensor one-by-one. As each particle passes the light-sensitive detector, a change in the voltage output from the detector is observed to count the particle and the magnitude of the change is used to determine the size of the particle. As such, a calibration curve using NIST-traceable size standards, usually polystyrene latex, is needed to calibrate the LO method to generate accurate sizes from the voltage changes.

LO is dependent on proper instrument preparation to ensure that the particle concentration does not exceed the typical 10,000–20,000 particles/mL. More stringent approaches include optimizing the test environment and sample preparation by preparing samples in and running the instrument in a laminar air flow hood or similar, utilizing appropriate testing methods for a particular product and employing the correct data analysis and subsequent interpretation.

LO can typically measure particle sizes between 1–2 μm up through 100–400 μm. The lower size is dependent on adequately obtaining signal-to-noise ratio of the particles over background whereas the upper limit depends on the dimensions of the sensor flow channel. An upper limit of 10,000–20,000 particles/mL exists, above which coincidence of particles becomes an issue. Coincidence is the traversing of two or more particles simultaneously through the sensor and is counted as a single, larger particle. This results in lower concentrations than actual with a distribution shifted to larger sizes.

The importance of sample preparation procedures must be emphasized. Protein solutions, especially those at high concentration or containing surfactants, tend to generate and retain air bubbles. If sample pooling is necessary for obtaining adequate volumes for USP testing, additional concerns must be highlighted. Combining samples can not only introduce air bubbles, but the mixing process can impact existing particle levels. Sufficient analyst training is necessary for minimizing inter-analyst and intra-analyst variation, and often requires consistent pipetting (gentle) or hand swirling to gently mix to obtain consistent results.

The main strength of LO is the simplicity and robustness; the instrument is easy to use with simple data analysis and contains few user-optimized settings. This has allowed this method to become widespread for use in biological formulations as a

standard characterization and release assay. Despite its adoption, the LO method has several well-known and important weaknesses. LO is relatively labor intensive and requires an analyst to manually run the samples, making the method both time consuming and low throughput. Furthermore, most particles observed in biopharmaceuticals differ in physical properties to calibration polystyrene latex beads and in matrix composition (buffer versus water) used to calibrate the instrument. As such, they differ in optical and morphological properties and may result in differences in both size and particle concentrations. Protein particles often have only slight differences in refractive indices than the matrix, making them nearly invisible (or under-sized) to the detector. However, divergences in reported sizes may not pose too much of a problem for routine analysis as reported size bins are generally large such as >10 or >25 μm. Similarly, problems may occur with samples that are not clear or have viscosities significantly different than water. Additionally, an equivalent circular diameter is assumed to calculate sizes from the voltage changes without any details provided on particle morphology. LO cannot determine particle morphology, shape, or composition and will not provide any information towards identifying the particle type or origin.

Flow Imaging Microscopy

USP <1787> Measurement of Subvisible Particulate Matter provides guidance on the strategies available for characterizing and identifying particle populations within therapeutic protein injections and emphasizes differentiating extrinsic and intrinsic particle populations from inherent proteinaceous particle populations. It is recognized that monitoring the subvisible particle populations, specifically in the 2–10 μm range, may be a key product quality attribute. A key purpose of USP <1787> is to provide guidance on applying analytical methods for the 2–10 μm size range, and its application is suggested for orthogonal characterization during product development, stability studies, investigations into root cause analysis of nonconformity, and other purposes not specific. Methods listed can be used to determine whether particles are inherent for a therapeutic protein product or from an extrinsic source, such as process-related or product container.

Although USP <1787> discuses many different, complementary methods, Flow Imaging Microscopy (FIM) is typically the first method attempted and its application is becoming more widespread and routine in biopharmaceutical development. Also known as Flow Imaging Analysis or Dynamic Image Analysis, this is an optical method that illuminates a sample under flow and captures digital images of the particles present (Fig. 15.5). Sample is drawn through a cell of defined dimensions and illumination is synced with image acquisition through microscope lens to generate high resolution images of the particles. Analysis of the stored images post-collection allows morphological characterization to generate particle size distributions or to identify unique or separate particle populations. FIM is typically used to identify and characterize particles between 1 and 300 μm.

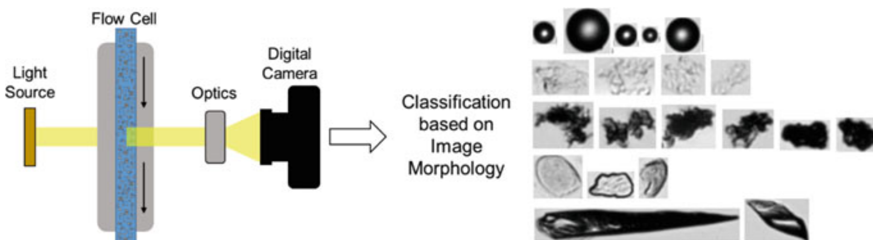


Fig. 15.5 Flow Imaging Microscopy analysis and determination of particle size and concentration

FIM has several key advantages over traditional LO methods. FIM has an increased sample volume and reduced analysis time, resulting in a medium throughput. Furthermore, because particles are directly observed in situ and their images (morphologies) saved, different analysis algorithms can be performed on a sample even after data collection is complete and allows specialized or proprietary morphology characterization software to be generated to a user's individual needs. Morphological features including aspect ratio, shape, particle contrast, and particle intensities can provide important insight towards identifying the source of identified particles, including the ability to differentiate silicone oil droplets from air bubbles from protein particles. Similarly, under the right conditions, morphological analysis can differentiate different sub-populations of particles, such as semi-transparent fibrous-like protein particles from dark, compact polysorbate particles. This is especially useful in a forensic approach where information about particle appearance provides crucial information towards identifying the source of the particles, such as identifying them as intrinsic, extrinsic, or inherent. Often this can serve to differentiate problems with the formulation, during process development, or with container compatibility during the development lifecycle of a biopharmaceutical product.

However, proper morphologic characterization can only be applied to particles greater than 4–5 μm. Further limitations include difficulty in visualizing particles with low optical contrast, such as translucent proteinaceous particles, and thereby results in under-counting their populations. Additionally, the results depend on the algorithms used for morphological analysis to classify and size particles. For instance, differences in how to classify a silicone oil droplet versus a protein particle may report different concentrations, although the trends are usually similar. Dilution is necessary in many cases and may affect the sample properties or particle profile, especially if particle formation is concentration-dependent. Stresses during dilution, including as sheer stresses during handling, pipetting, or mixing steps, may contribute to changes in particle concentrations. However, linearity for dilution is typically assessed when a dilution scheme is required for sample analysis. Finally, FIM is a destructive technique and samples cannot be reused or retained.

Instrumentation and Characterization of Submicron Particles: Nanoparticle Tracking Analysis and Dynamic Light Scattering

Regulatory agencies are starting to show considerable interest in submicron particle assessment (US FDA 2014). The presence of nanoparticles in protein formulations can provide very early indication for stability problems, and is the one of the most sensitive methods for tracking degradation processes. Instrumentation in this size range is starting to become quite widely incorporated into industry research, leading to an increase in the number of publications and case studies containing particles in this size range.

Dynamic Light Scattering

Dynamic light scattering (DLS) is an analytical technique used to measure the particle size distribution of protein formulations across the oligomer and submicron size ranges of approximately 1 nm to 1 μm. DLS is a well-established technology within the biotherapeutic industry, with a large body of application and literature knowledge, primarily due to the technique's wide working size and concentration ranges, low volume and minimal sample preparation requirements. Particle size is determined by quantifying Brownian motion, the random motion of particles in solution, through measurement of the diffusion coefficient. The light scattered from a solution of particles diffusing under the influence of Brownian motion will fluctuate with time (Fig. 15.6). Across long time intervals, the scattering trace appears to be representative of random fluctuations about a mean. When viewed on smaller time scales, however (inset in Fig. 15.6—Left), it is evident that the intensity trace is in fact not random, but composed of a series of continuous data points. By performing a correlation function on these intensity fluctuations, average particle size (Z-ave) and particle size distributions can be determined. Size measured by DLS is the hydrodynamic dimeter/radius, and is defined as the hydrodynamic radius of a solid sphere with the same diffusion coefficient as molecule of interest. Importantly, derivation of Z-ave and PdI values from correlation data are defined by ISO standards (ISO 22412:2008).

By default, DLS measurements provide intensity-based size distributions. Scattering intensity is proportional to the particle diameter to the power of six by the Rayleigh approximation. Therefore, size distribution by intensity is biased to larger sizes. This is a major advantage when trying to detect large nanometer scale particles in the presence of monomeric protein. DLS data is qualitative, not quantitative, and so is particularly suited to biocomparability studies when looking for relative differences between samples and batches (Jiang and Narhi 2006). As described earlier, large protein particles form as the result of smaller nanometer sized particles and soluble aggregates. The ability to monitor changes in the aggregated state of proteins, as well as detect the presence of large particles, makes DLS a flexible

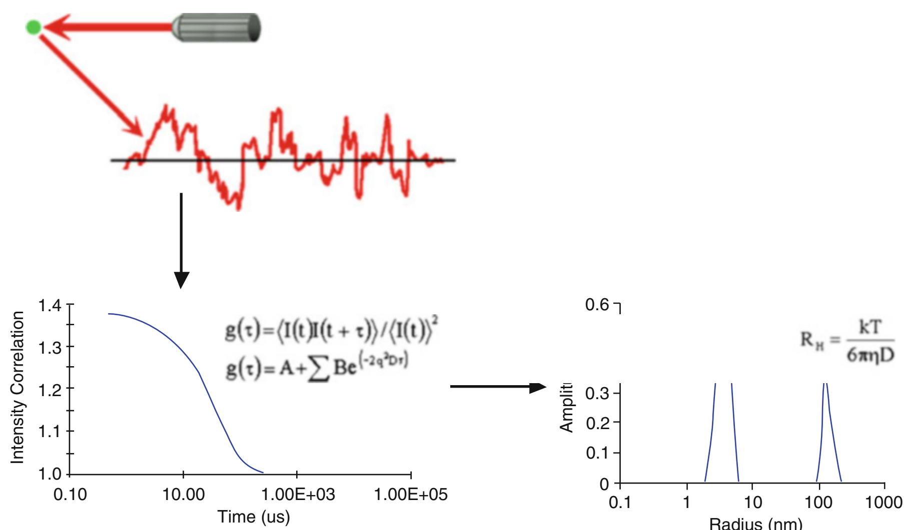


Fig. 15.6 Correlation of light intensity to determine particle hydrodynamic size by Dynamic Light Scattering

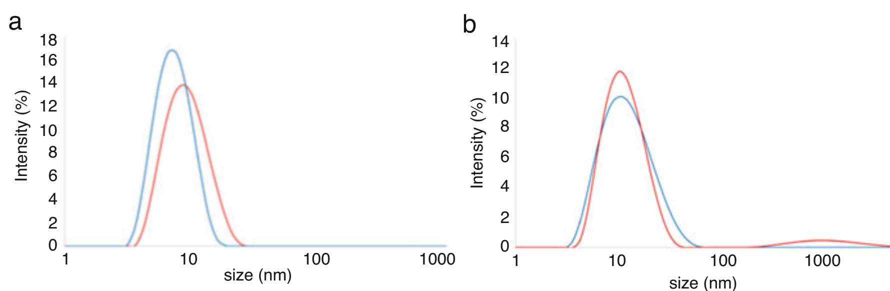


Fig. 15.7 Tracking changes in particles size (**a**) and detection of subvisible particles (**b**)

and versatile tool for assessing particle content of biotherapeutics (Nobbmann et al. 2007). Indeed, DLS has been used previously to detect formation of large protein aggregates in the presence and absence of silicone oil (Bai et al. 2016).

Uniquely amongst other particle sizing technologies described here, DLS can size and track changes in the presence of soluble aggregates and low molecular weight oligomers. Monitoring changes in the hydrodynamic size can provide very sensitive detection of early onset aggregation processes (Fig. 15.7). The ability to measure samples with minimal preparation allows rapid determination of aggregation state. Fig. 15.7a shows a comparison of two mAb batches, showing a clear difference in hydrodynamic size as a result of increased aggregation state.

Moving beyond monitoring changes in oligomeric state, DLS is also a highly sensitive method for detecting the presence of larger submicron-sized particles

(Panchal et al. 2014; Filipe et al. 2010). The Rayleigh approximation means that larger particles will scatter light with a much higher intensity, often several orders of magnitude higher than the monomer and low molecular weight oligomers. Consequently, for a rapid qualitative assessment for the presence of particles, few techniques can match the sensitivity of DLS. Figure 15.7b shows the size distribution for a protein at different time points during a storage stability study. The peak for the main size population shows minimal change during the time course, but the presence of much larger particles can be clearly detected. In fact, these particles may be present in such low quantities that other particle counting techniques may not detect them at all. However, this data is not quantitative, and cannot provide particle counts, which is an important directive from the regulatory agencies. Consequently, DLS is powerful support tool, but not a front-line technology for subvisible particle content assessment within the biopharmaceutical industry.

Considering the biosimilar field specifically, DLS has been included as part of Amgen's filing for ABP501, a biosimilar of Humira® (Liu et al. 2016; FDA AAC Brief 2016). Included in the submicron particle assessment category, DLS was used as a Tier 3 qualitative assessment of biosimilarity. Tier 3 attributes have the lowest risk to clinical outcome and are not suitable for numerical assessment, and so no specific criteria are set. However, the intensity-weighted size distribution profile was used to provide qualitative comparison between ABP 501 and Humira® from the US and Europe (Liu et al. 2016).

Nanoparticle Tracking Analysis

Despite the advantages of DLS in submicron particle analysis, one major drawback is the qualitative nature of the data, and the lack of particle content quantification of samples. To address this gap in submicron particle knowledge, a separate technology can be utilized. Nanoparticle Tracking Analysis (NTA), is a technique which allows the sizing and concentration measurement of nanoscale particles. NTA is similar to DLS, in that it determines hydrodynamic size through measurement of Brownian motion and diffusion coefficient. However, the key difference is that the scattered light from each particle is tracked and measured individually, providing high resolution data, and crucially for submicron applications, particle concentration can be determined.

For sample acquisition, a laser beam is directed into the sample chamber. The angle of incidence, and refractive index of the sample chamber and liquid sample is designed to compress the laser beam to a reduced profile and a high power density. The particles in the path of this beam scatter light in such a matter that they can be easily visualized via a long working distance, microscope objective fitted to an otherwise conventional optical microscope or equivalent optical train. The system then uses a digital camera, operating at typically 30 frames per second (fps) to capture a video file of particles moving under Brownian motion within a known field of view (Fig. 15.8). Different NTA instrument manufacturers provide sample

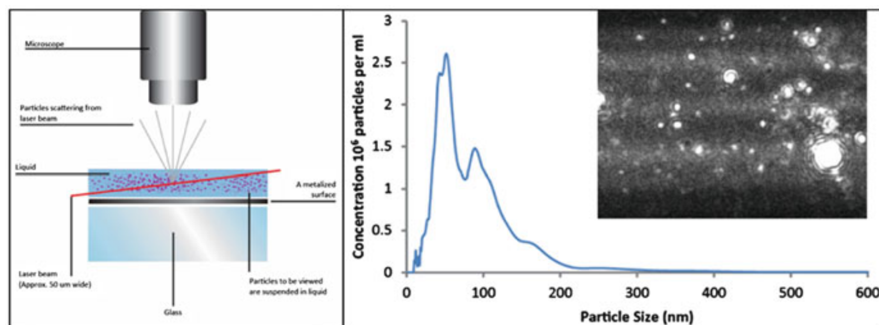


Fig. 15.8 Nanoparticle Tracking analysis and determination of particle size and concentration (Images and data shown are taken from a NanoSight NTA system)

chamber and optical train configurations that differ in orientation, laser unit, optical objectives, and custom software for data acquisition and analysis.

The lower size range of NTA excludes the ability to detect monomers and small aggregates, but as each particle is sized individually, protein aggregates starting around 50 nm can be sized and counted, with greatly enhanced resolution compared with DLS (Filipe et al. 2010). The concentration and size distribution profiles of nanometer scale particles can be determined in therapeutic proteins, both under standard conditions and following forced degradation experiments (Joubert et al. 2011). Studies with commercial biopharmaceutical products have utilized NTA for the assessment of submicron particle counts and concentration on storage and following agitation and freeze-thaw (Shujun et al. 2013; Bai et al. 2016). Of particular interest is the impact of submicron particles on the formation of micron-sized particles. Shujun et al. (2013) showed that thermally-induced submicron sized particles induce the formation of micron-sized particles during long term storage (Shujun et al. 2013). This suggests screening of submicron sized particles may serve as a useful predictor for the formation of larger particles on shelf life or during commercial manufacture. Given that the focus on USP methods and FDA regulations focus on larger, micron range particles, NTA could provide an important method for the identification of problematic products and formulations.

Novel Technologies for the Assessment and Characterization of Particles

With the increasing demands from regulatory agencies to further characterize particles in parenteral products, novel, advanced technologies and instruments are now available to provide previously unattainable product knowledge. With scrutiny of product safety and immunogenicity only likely to intensify in the coming years, these technologies are likely play an increasingly important role in monitoring product quality.

Resonance Mass Measurement

Resonant Mass Measurement (RMM) was designed to satisfy two key limitations in existing particle data. Firstly, the size range of 200 nm to 5 μm nicely fills the sizing gap between SEC and FIM, providing data on particles that were previously missed (Barnard et al. 2012; Weinbuch et al. 2013). Secondly, is the ability to unequivocally distinguish between silicone oil and other particle types, including protein aggregates. Silicone oil is a common component of pre-filled syringes and rubber stoppers, and whilst considered safe, the misidentification of silicone oil for protein (or other) particles can lead to misleading product data (Barnard et al. 2012). Indeed, in their "Guidance for Industry" the FDA recommend "All therapeutic protein products should be evaluated for their content of [...] incidental product components, including proteins and nonprotein components (US FDA 2014)".

The principle of RMM relies on having a mechanically resonant structure to which mass can be added or subtracted. This addition or subtraction shifts the structure's resonant frequency either up or down, providing the basis for buoyant mass measurement because frequency can be measured very precisely. A microfluidic channel is embedded inside the resonator to measure the mass of very small particles suspended in fluid. As a suspended particle passes through the structure, its buoyant mass causes a shift in the mass of the overall resonator, thereby shifting its resonant frequency (Fig. 15.9). It is the excursion in resonant frequency from the baseline that enables the measurement of the mass of a particle, and from there the size of the particle can be calculated.

The principles of the measurement also allow the distinction between particles that are negatively buoyant, i.e. more dense that the bulk liquid, and those that are positively buoyant, i.e. less dense than the bulk liquid. In the case of biotherapeutics, this allows the quantification of silicone oil droplets independently of other particles types. Whilst negatively buoyant particles show a drop in frequency, positively buoyant particles, such as oil droplets, have the opposite effect with a positive frequency shift (Fig. 15.9). Silicone oil is often used as a lubricant inside the

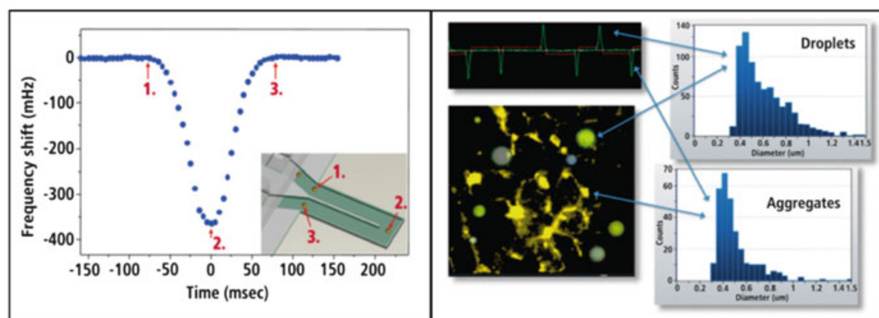


Fig. 15.9 Depiction of frequency shifts detected during resonant mass measurement analysis, and the distinction between protein aggregates and silicone oil

syringes and as a coating from vial stoppers, and so often leaches into the product over time. The issue is not primarily one of biocompatibility as the oil is generally considered to be safe. Reports have also shown that, for some products at least, the presence of silicone oil does not adversely affect protein stability (Bai et al. 2016). Of greater concern is the tendency of the oil droplets to be misperceived by some measurement methods as protein aggregates (Barnard et al. 2012; Bai et al. 2016; Weinbuch et al. 2013; Strehl et al. 2012), therefore compromising the accuracy of the results. RMM can distinguish silicone oil droplets from protein aggregates through buoyancy measurements, and therefore this distinction is unequivocal in the size range analyzed (Weinbuch et al. 2013). Multiple reports have demonstrated the use of RMM to study the content of subvisible particles in both the nanometer and low micron size range (Haji et al. 2016), and in particular, the ability to distinguish between protein aggregates and silicone oil droplets (Barnard et al. 2012; Bai et al. 2016; Weinbuch et al. 2013). A study on the impact of silicon oil droplets on the stability of peginterferon beta-1a used RMM to demonstrate that the majority of positively buoyant particles are less than 1 μm, meaning they would be missed by MFI, whilst silicon oil is the predominant particle below 1 μm (Bai et al. 2016). This ability to count particles and distinguish between silicon oil and other particle types is an important aspect of biosimilarity assessment, as it considers both product stability and container closure. Without such a distinction, it is possible to misinterpret biosimilarity data.

Morphology-Directed Raman Spectroscopy

Optical microscopy has long been used to characterize particulates present in biotherapeutic formulations. Information such as particle size, shape, and transparency characteristics, can be used to group particulates into distinct classes (i.e. aggregates, silicone oil, contaminants etc.). In fact, light microscopy is a standard requirement for USP <788>. However, the ability of a microscope to provide explicit identification is limited to the two-dimensional images it collects. Spectroscopy provides a means to unequivocally identify the composition of the particles, and therefore their source. Although FTIR is often used for material identification purposes, Raman has significant advantages including reduced water sensitivity, improved small particle detection and identification of common particles in protein formulations (Saggu et al. 2015). The addition of Raman spectroscopy to an automated microscopy system provides an identification method for the verification of particle chemistry, and therefore the potential to enumerate, characterize, and identify particulates in native formulations, as well as those immobilized on a filter substrate. This range of capabilities directly addresses FDA quality requirements for parenteral products, in terms of particulate count per unit volume, but also the source of contaminant particles. Identification of the source of contaminant particles is an important part of trouble-resolution during commercial manufacture of biotherapeutics. In 2017, a Celltrion site in South Korea, which manufactures

Pfizer's Inflectra®, a biosimilar of Remicade®, was issued a Form 483 from the FDA. One highlighted observation was a failure to investigate foreign matter in the drug product in a timely manner and that the source of the particles was not fully evaluated (FDA 2018). Therefore, there is a clear directive for companies to identify the source of particles in sterile products.

Morphology-Directed Raman Spectroscopy (MDRS) allows measurement of particle size and shape, while also using Raman spectroscopy to chemically identify the particles. Analysis is performed in three stages: automated imaging of particles to construct number-based distributions based on size and shape; particles of interest, based on size or shape, are analyzed by Raman spectroscopy to chemically identify the particles; and spectra are compared with either internal, proprietary libraries or commercial libraries of known materials to provide positive identification of particles of interest.

As described earlier, particles in parenteral products are typically defined as inherent, intrinsic and extrinsic. Raman spectroscopy can identify many of the typical particle types in biotherapeutics. Examples of these types of particles and contaminants are shown in Fig. 15.10. When identifying protein aggregates, the Phenylalanine peak at 1005 cm^{-1} is typically present, but bands assigned to tyrosine and β-sheet can also be utilized under certain conditions (Pardeshi et al. 2017). Typically, Raman spectra collected for protein particles using MDRS are generally weak, containing little useful structural information. This is because the morphologies of the particles, typically sheets or thin particles, do not lend themselves to strong Raman signal based on the thickness of the particles. Particles

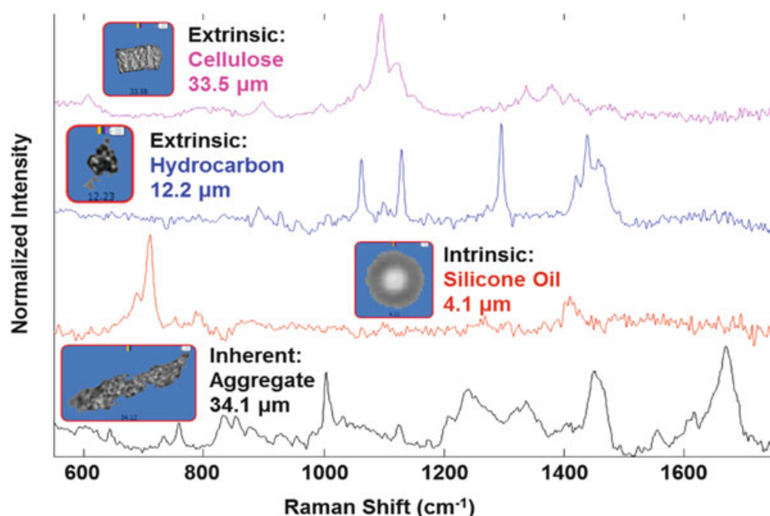


Fig. 15.10 Particle images and the respective Raman spectra, and therefore identification, collected by MDRS

containing silicone oil, surfactant degradants, and polymers typically have much stronger and distinctive spectra.

There are multiple reports utilizing Raman spectroscopy for identification of particles within parenteral products, including proteinaceous silicone oil and polycarbonate (Pardeshi et al. 2017; Saggu et al. 2015; Cao et al. 2009). Of interest is the formation of particles through the degradation of excipients. Although most inherent particles may be considered proteinaceous, polysorbate 20 can be hydrolyzed into insoluble particles, such as lauric acid. The distinction between protein particles or free fatty acid cannot be made based on morphology alone, and requires Raman spectra (Saggu et al. 2015). Although efforts of biopharmaceutical companies focus on particle generation during manufacture and storage, particles have been shown to be introduced during administration and dosing. Following dilution into an IV saline bag and passed through an IV infusion set, particles were identified in a mAb-based product. By using MRDS, these particles were identified as polycarbonate, and so were likely to have been shed from the IV line and IV connectors. Therefore, particle analysis activities should not be limited to the development and Fill/Finish activities, but rather expanded to include drug handling and clinical administration (Pardeshi et al. 2017).

Microfluidic Resistive Pulse Sensing

Instruments based upon the Coulter Principle or Resistive Pulse Sensing (RPS) have long been a staple of particle analysis. The Coulter Principle dictates that as a particle transverses through a constriction with an electric current, displacement of the electrolyte solution produces a change in the impedance proportional to the electrolyte solution volume displaced. A particle's size is calculated from this displacement. Recent instrumentation that incorporates microfluidics (microfluidic resistive pulse sensing, MRPS) has increased the versatility of traditional RPS instruments by expanding the dynamic range of analysis into the submicron range. Newer technologies allow detection of particle sizes between 50 nm and 2 μm, depending on which cartridge is chosen. This is a particular range of interest as it is truly orthogonal to both NTA and RMM, relying on completely different detection principles to identify particles in the same size range. Like NTA and RMM, MRPS is a high-resolution method able to individually count particles. Unlike NTA and RMM, MRPS has few limitations on samples types and can analyze particles that do not rely on optical contrast, a limitation with NTA, or relies on differential buoyancy, oft times a limitation with RMM. Sample analysis with MRPS can be completed in minutes, able to identify a maximum of roughly 10,000 particles/s (dilution may not be necessary), and requires a minimum volume of 3 μL. Similar to NTA and RMM, MRPS has been shown to be useful in a wide range of applications, including monitoring aggregation in biopharmaceutical products, analysis of excipients, and characterization of nanomedicines such as liposomes and exosomes. However, because of the use of the microfluidic systems, the system is prone to blockage

and care must be taken to remove particles >2 μm. For proper particle sizing, a calibration step is required using size standards suspended in the sample buffer. Furthermore, data analysis is more intensive and requires a more experienced user to operate. Although experiments can be achieved in a relatively short amount of time, sample preparation and data collection is a labor-intensive process and requires a user to manually run the samples. As such, the method is typically low-to-medium throughput.

Examples of Particle Characterization in Biosimilarity and Biocomparability Studies for Licensed Products

ABP-501: Amgen

ABP 501 is Amgen's biosimilar to Adalimumab, previously marketed as Humira© by Abbvie. ABP 501 was approved by the FDA in 2016, for the treatment of multiple inflammatory diseases. The FDA advisory committee meeting notes, and a study (Liu et al. 2016) outline in detail the analytical and methodological approach that was taken to demonstrate similarity. Particle and aggregate analysis was performed using LO, FIM, DLS and Field Flow Fractionation. These categories were assigned as tier 3 categories, meaning no numerical acceptance criteria, except for FIM counts of non-spherical particles. This latter category was assigned as tier 2, with an acceptance range of 0–197 ≥5 μm non-spherical particles per mL. The LO data shows a broad range of particle concentration data, for example, 4560–31,000 particles ≥2 μm/mL, for US supplied adalimumab (Liu et al. 2016). This highlights the challenge of assessing particle concentration as part of a biosimilarity assessment. Given such a broad range of particle concentrations for the reference material, the target criteria for the biosimilar must also be similarly broad. A comparison of FIM counted ≥5 μm non-spherical particle data (24–172 particles/mL) with LO ≥5 μm particle counts (1000–7673 particles/mL) strongly suggests that the predominant particle species in these samples is silicon oil. No assessment of submicron particle concentration was performed, with the only comparison being qualitative assessment by DLS.

With the advancement of manufacturing equipment and understanding of the processes, one might expect the particle counts for biosimilars to be lower than innovator products, which have often been on the market for several years. However, at present there is no requirement for biosimilar companies to produce products with lower particle counts than the products they are copying.

FDA Study on Multi-Dose Erythropoietin: Why Characterization Is Important-Immunogenicity and Safety

A recent study by the FDA has highlighted the importance of extended particle characterization to fully assess product safety and immunogenic risk (Kotarek et al. 2016). Approved in 2012, Peginesatide® (Affymax Inc.) was withdrawn voluntarily from the market after less than 1 year, due to severe clinical complications. Marketed as an erythropoiesis-stimulating agent (ESA), Peginesatide® has no amino acid homology to erythropoietin (Epo), and therefore was effective in patients with pure red cell aplasia and anti-Epo antibodies. However, 49 cases of anaphylaxis were reported, including 7 fatalities. The associated hypersensitivity rate was 3.5 cases per 1000, higher than the pre-market clinical trial data yielding a rate of 0.84 cases per 1000. Data available for 32 of the anaphylaxis cases, identified 30 of these had prior exposure to some form of ESA treatment, suggesting the effect was specific to Peginesatide®. Based on these clinical data, the focus of the investigation between pre- and post-market effects was to identify differences in product quality.

The product itself was manufactured and approved as both a single use vial (SUV) and a multi-use vial (MUV), which differed slightly in terms of the formulation. Although clinical trials focused on the SUV, only the MUV was marketed. Review of product release documentation showed that it met all approved release specifications, including USP <788>. As mentioned above, USP <788> specifies particulate limit for SVPs above 10 μm. However, biotherapeutic products often contain many particles below 10 μm (Carpenter et al. 2009). As part of the research study, submicron particles between 50 nm and 1 μm were quantified by NTA and those above 10 μm by flow imaging. Between 100 nm and 1 μm, NTA showed that the MUV contained significantly more particles than the SUV, a result that was confirmed with DLS data. Above 10 μm, MUV contained a higher concentration of particles and a broader size range compared with SUV, although all values were below the limit specified in USP <788>. The cause of the difference has not been identified, although the difference in formulation has been highlighted as the most likely cause. These data suggest that more sensitive monitoring of submicron and subvisible particles could distinguish differences in product characteristics, even if they satisfy USP <788> limits. As a closing statement, the paper concludes: "This case illustrates the potential value of using methods beyond those described in USP <788> to characterize subvisible particles in biological therapeutics".

This is a particularly high-profile case that highlights the risk of submicron and subvisible particles in parenteral products. However, as the range of available technologies has increased, companies now have the capability to perform more vigorous characterization of particle content on their products, both during development and post marketing. In the future, such product knowledge will become important to demonstrate product safety and satisfy regulatory expectations.

Conclusion

Immunogenicity and immunogenic risk continues to play a central role in the concerns that direct the regulatory guidance. As the rate of development and approval of biosimilars starts to increase, there is a growing interest in the methods and criteria used to assess similarity. Given the link between protein aggregates and subvisible particles in the development of immunogenicity, this will undoubtedly mean particle characterization data will remain a topic of interest. Whilst the focus on submicron and subvisible particles applies to all biotherapeutic products, biosimilars pose some unique challenges. Firstly, biosimilars are often developed and commercialized through collaborations between large pharmaceutical companies and smaller contract/development companies. As a result, the level of oversight may be different than for a product developed exclusively in-house. Secondly, there is the question of whether a biosimilar should have the same particle content as the innovator product, or whether it should be assessed based on best practise and modern standards. The particle characterization tool box that is now available to biopharmaceutical development teams is considerably more capable, providing a wealth of information about the particles present in a product across a much wider sizing range and also includes the ability to chemically identify the source of the particles. As these methods become more widely available, should biosimilar products be targeting lower particles counts, and therefore potentially reducing the immunogenic risk? As ever, the driver behind these decisions are the regulatory agencies. However, as more data becomes available to both development scientists and regulatory agents, the focus on particle content of biotherapeutics will not subside.

References

An Z. Therapeutic monoclonal antibodies: from bench to clinic. Hoboken: Wiley; 2011.

Antonelli G, et al. Interferon antibodies in patients with infectious diseases. Biotherapy. 1997;10(1):7–14.

Bai S, et al. Evaluation of incremental siliconization levels on soluble aggregates, submicron and subvisible particles in a prefilled syringe product. J Pharm Sci. 2016;105(1):50–63.

Barandun S, Kistler P, Jeunet F, Isliker H. Intravenous administration of human γ-globulin. Vox Sang. 1962;7(2):157–74.

Barnard JG, Babcock K, Carpenter JF. Characterization and quantitation of aggregates and particles in interferon-β products: potential links between product quality attributes and immunogenicity. J Pharm Sci. 2012;102(3):915–28.

Bertolotto A, et al. Immunogenicity of interferon beta: differences among products. J Neurol. 2004;251(2):ii15–24.

Braun A, et al. Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-α) in normal and transgenic mice. Pharm Res. 1997;14(10):1472–8.

Bukofzer S, et al. Industry perspective on the medical risk of visible particles in injectable drug products. PDA J Pharm Sci Technol. 2015;69(1):123–39.

Cao X, et al. Raman microscopic applications in the biopharmaceutical industry: in situ identification of foreign particulates inside glass containers with aqueous formulated solutions. Appl Spectrosc. 2009;63(7):830–4.
Carpenter JF, et al. Overlooking subvisible particles in therapeutic protein products: gaps that may compromise product quality. J Pharm Sci. 2009;98(4):1201–5.
Carpenter JF, et al. Potential inaccurate quantitation and sizing of protein aggregates by size exclusion chromatography: essential need to use orthogonal methods to assure the quality of therapeutic protein products. J Pharm Sci. 2010;99(5):2200–8.
Casadevall N, et al. Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. N Engl J Med. 2002;346(7):469–75.
Chaffee S, et al. IgG antibody response to polyethylene glycol-modified adenosine deaminase in patients with adenosine deaminase deficiency. J Clin Invest. 1992;89(5):1643–51.
Christie M, et al. The role of protein excipient in driving antibody responses to erythropoietin. J Pharm Sci. 2015;104(12):4041–55.
Claman HN. Tolerance to a protein antigen in adult mice and the effect of nonspecific factors. J Immunol. 1963;91(6):833–9.
Dresser DW. Specific inhibition of antibody production: II. Paralysis induced in adult mice by small quantities of protein antigen. Immunology. 1962;5(3):378–88.
Ellis EF, Henney CS. Adverse reactions following administration of human gamma globulin. J Allergy Clin Immunol. 1969;43(1):45–54.
EMA. Guideline on development, production, characterisation and specifications for monoclonal antibodies and related products. EMA, editor. 2008.
FDA. US Food and Drug warning letter 320-18-28. 2018. https://www.fda.gov/ICECI/EnforcementActions/WarningLetters/ucm594395.htm. [Cited 23 May 2018].
FDA AAC Brief. 12 July 2016 Arthritis Advisory Committee Meeting Briefing Document, ABP 501, a proposed biosimilar to Humira®. 2016. https://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/arthritisadvisorycommittee/ucm510295.pdf. [Cited 2018 May 23].
Filipe V, Hawe A, Jiskoot W. Critical evaluation of Nanoparticle Tracking Analysis (NTA) by nanosight for the measurement of nanoparticles and protein aggregates. Pharm Res. 2010;27(5):796–810.
Filipe V, et al. Immunogenicity of different stressed IgG monoclonal antibody formulations in immune tolerant transgenic mice. MAbs. 2012;4(6):740–52.
Fradkin AH, Carpenter JF, Randolph TW. Immunogenicity of aggregates of recombinant human growth hormone in mouse models. J Pharm Sci. 2009;98(9):3247–64.
Fradkin AH, Carpenter JF, Randolph TW. Glass particles as an adjuvant: a model for adverse immunogenicity of therapeutic proteins. J Pharm Sci. 2011;100(11):4953–64.
Gamble CN. The role of soluble aggregates in the primary immune response of mice to human gamma globulin. Int Arch Allergy Immunol. 1966;30(5):446–55.
Girard F, Gourmelen M. Clinical experience with somatonorm. Acta Paediatr. 1986;75(s325): 29–32.
Grauer A, et al. Neutralizing antibodies against salmon calcitonin. The cause of a treatment failure in Paget's disease. Deutsche Medizinische Wochenschrift (1946). 1994;119(14):507–10.
Gribben JG, et al. Development of antibodies to unprotected glycosylation sites on recombinant human GM-CSF. Lancet. 1990;335(8687):434–7.
Grossberg SE, et al. Frequency and magnitude of interferon β neutralizing antibodies in the evaluation of interferon β immunogenicity in patients with multiple sclerosis. J Interf Cytokine Res. 2011;31(3):337–44.
Haji AM, et al. An albumin-free formulation for *Escherichia coli*-derived interferon beta-1b with decreased immunogenicity in immune tolerant mice. J Interf Cytokine Res. 2016;36(3): 192–203.
Hermeling S, et al. Development of a transgenic mouse model immune tolerant for human interferon beta. Pharm Res. 2005;22(6):847–51.
Jacquemin MG, Saint-Remy JM. Factor VIII immunogenicity. Haemophilia. 1998;4(4):552–7.

Jiang Y, Narhi LO. Applying selective biophysical techniques in assessing the comparability of protein therapeutics - case studies. Am Pharm Rev. 2006;9:34–43.

Joubert MK, et al. Classification and characterization of therapeutic antibody aggregates. J Biol Chem. 2011;286(28):25118–33.

Karpusas M, et al. The structure of human interferon-β: implications for activity. Cell Mol Life Sci. 1998;54(11):1203–16.

Kessler M, Goldsmith D, Schellekens H. Immunogenicity of biopharmaceuticals. Nephrol Dial Transplant. 2006;21(Suppl 5):v9–v12.

Kontsek P, Liptáková H, Kontseková E. Immunogenicity of interferon-alpha 2 in therapy: structural and physiological aspects. Acta Virol. 1999;43(1):63–70.

Kotarek J, et al. Subvisible particle content, formulation, and dose of an erythropoietin peptide mimetic product are associated with severe adverse postmarketing events. J Pharm Sci. 2016;105(3):1023–7.

Kumru OS, et al. Compatibility, physical stability, and characterization of an IgG4 monoclonal antibody after dilution into different intravenous administration bags. J Pharm Sci. 2012;101(10):3636–50.

Lebron JA, et al. Nonclinical safety assessment of vaccines and the evaluation of novel adjuvants and delivery systems. In: Vaccine adjuvants and delivery systems. Hoboken: Wiley; 2007. p. 403–20.

Liu J, et al. Assessing analytical similarity of proposed amgen biosimilar ABP 501 to adalimumab. BioDrugs. 2016;30(4):321–38.

Martin F, Oliver AM, Kearney JF. Marginal zone and B1 B cells unite in the early response against T-independent blood-borne particulate antigens. Immunity. 2001;14(5):617–29.

Milner RD. Growth hormone 1985. Br Med J. (Clin Res Ed). 1985;291(6509):1593–4.

Moore WV, Leppert P. Role of aggregated human growth hormone (hGH) in development of antibodies to hGH. J Clin Endocrinol Metab. 1980;51(4):691–7.

Neumann TAK, Foote M. Megakaryocyte growth and development factor (MGDF): an Mpl ligand and cytokine that regulates thrombopoiesis. Cytokines Cell Mol Ther. 2000;6(1):47–56.

Nobbmann U, et al. Dynamic light scattering as a relative tool for assessing the molecular integrity and stability of monoclonal antibodies. Biotechnol Genet Eng Rev. 2007;24:117–28.

Palleroni AV, et al. Interferon immunogenicity: preclinical evaluation of interferon-alpha 2a. J Interf Cytokine Res. 1997;17(Suppl 1):S23–7.

Panchal J, et al. Analyzing subvisible particles in protein drug products: a comparison of dynamic light scattering (DLS) and resonant mass measurement (RMM). AAPS J. 2014;16(3):440–51.

Pardeshi NN, et al. Microparticles and nanoparticles delivered in intravenous saline and in an intravenous solution of a therapeutic antibody product. J Pharm Sci. 2017;106(2):511–20.

Patten PA, Schellekens H. The immunogenicity of biopharmaceuticals. Lessons learned and consequences for protein drug development. Dev Biol. 2003;112:81–97.

Prümmer O. Treatment-induced antibodies to interleukin-2. Biotherapy. 1997;10(1):15–24.

Ring J, Stephan W, Brendel W. Anaphylactoid reactions to infusions of plasma protein and human serum albumin. Role of aggregated proteins and of stabilizers added during production. Clin Allergy. 1979;9(1):89–97.

Rosenberg AS. Effects of protein aggregates: an immunologic perspective. AAPS J. 2006;8(3):E501–7.

Rosenschein U, et al. Streptokinase immunogenicity in thrombolytic therapy for acute myocardial infarction. Isr J Med Sci. 1991;27(10):541–5.

Saggu M, Liu J, Patel A. Identification of subvisible particles in biopharmaceutical formulations using raman spectroscopy provides insight into polysorbate 20 degradation pathway. Pharm Res. 2015;32(9):2877–88.

Schellekens H. Bioequivalence and the immunogenicity of biopharmaceuticals. Nat Rev Drug Discov. 2002;1:457.

Schellekens H. How to predict and prevent the immunogenicity of therapeutic proteins. In: El-Gewely MR, editor. Biotechnology annual review. Amsterdam: Elsevier; 2008. p. 191–202.

Schernthaner G. Immunogenicity and allergenic potential of animal and human insulins. Diabetes Care. 1993;16(Suppl 3):155–65.

Shujun B, et al. Effects of submicron particles on formation of micron-sized particles during long-term storage of an interferon-β-1a solution. J Pharm Sci. 2013;102(2):347–51.

Simler BR, et al. Mechanistic complexity of subvisible particle formation: links to protein aggregation are highly specific. J Pharm Sci. 2012;101(11):4140–54.

Singh SK, et al. An industry perspective on the monitoring of subvisible particles as a quality attribute for protein therapeutics. J Pharm Sci. 2010;99(8):3302–21.

Strehl R, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: a novel multiparametric image filter for sub-visible particles in microflow imaging analysis. Pharm Res. 2012;29(2):594–602.

Underwood LE, Voina SJ, Van Wyk JJ. Restoration of growth by human growth hormone (Roos) in hypopituitary dwarfs immunized by other human growth hormone preparations: clinical and immunological studies. J Clin Endocrinol Metab. 1974;38(2):288–97.

US FDA. Guidance for industry: immunogenicity assessment for therapeutic protein products, US FDA, editor. 2014.

Vadhan-Raj S. Clinical experience with recombinant human thrombopoietin in chemotherapy-induced thrombocytopenia. Semin Hematol. 2000;37:28–34.

Veronese FM, Pasut G. PEGylation, successful approach to drug delivery. Drug Discov Today. 2005;10(21):1451–8.

Weinbuch D, et al. Micro-flow imaging and resonant mass measurement (Archimedes) – complementary methods to quantitatively differentiate protein particles and silicone oil droplets. J Pharm Sci. 2013;102(7):2152–65.

Chapter 16
Biological Activity Assays for Antibody Therapeutics

Xu-Rong Jiang and Anthony Mire-Sluis

Abstract The measurement of biological activity is required for therapeutic antibodies at release of a new manufacturing batch and throughout the product life cycle. However, the applicable regulatory guidelines provide relatively little advice with respect to the number and types of bioassays to be employed. For example, 21CFR610.10 states that potency tests "shall consist of either in vitro or in vivo tests, or both, which have been specifically designed for each product so as to indicate its potency in a manner adequate to satisfy the interpretation of potency" where potency is "the specific ability of capacity of the product, as indicated by appropriate laboratory tests or by adequately controlled clinical data obtained through the administration of the product in the manner intended, to effect a given result." ICH Q6B lists animal-based biological tests, cell culture-based biological tests, biochemical tests and ligand and receptor binding tests as examples of biological tests. However, neither guideline provides any assistance on how to select the most appropriate bioassay format(s) for a given product. Since biological activity is determined based on individual product characteristics and mechanism of action, there is considerable decision-making involved when selecting, developing, switching, and maintaining appropriate bioassays. This chapter discusses the types of biological assays, the current practices and regulatory expectations regarding the potency test format. A key question for antibody therapeutics that acts as agonists or antagonists of a ligand-receptor interaction, is the question of cell-based versus binding assays. Development and validation considerations for biological assays will also be discussed. This chapter will particularly focus on special considerations for biosimilar antibody therapeutics and provide a strategy for the biological testing methods, plan, and statistical assessment for the assessment of functional

X.-R. Jiang (✉)
BioVentures, AstraZeneca, Gaithersburg, MD, USA

Development Quality Biologics, AstraZeneca, Gaithersburg, MD, USA
e-mail: xu-rong.jiang@astrazeneca.com

A. Mire-Sluis
Development Quality Biologics, AstraZeneca, Gaithersburg, MD, USA
e-mail: anthony.mire-sluis@astrazeneca.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_16

biosimilarity. Key differences in the application of a battery of in vitro biological activity assays between a biosimilar and a novel biologic, and case studies will be discussed.

Keywords Biological activity assays · Bioassays · Potency assays · Receptor binding · Antibody-dependent cell-mediated cytotoxicity (ADCC) · Complement-dependent cytotoxicity (CDC) · Reference standard · Assay control · Method development and qualification · Method validation · Method transfer · Assay variability · Accuracy · Precision · Statistical assessment · Training · Parallelism · Data Analysis

Introduction

Monoclonal antibodies are in general complex large molecules, exhibiting many different physicochemical attributes and can exhibit more than one biological activity (both Fab and Fc functions). To ensure their quality and the robustness of process of production requires assessment of several physicochemical attributes including structure, purity, biological potency, stability and consistency batch to batch. In order to establish a risk based approach to assure consistency, characterization of the protein product should be undertaken during product development to establish what are the product attributes and their criticality, to create an overall control strategy.

Physicochemical analyses rely on the unique characteristics of a protein, and use multiple methods to characterize a structural or functional attribute. It is often the case that each test only provides detailed information about a single characteristic, although multi-attribute methods have been developed that can replace several existing types of tests. Even though the data from these assays can help provide the potential for an impact of differences in these attributes on the biological activity (PK, safety and/or efficacy), they are yet unable to predict the biological activity of the vast majority of biological products and therefore bioassays are an essential part of the characterization of the biological activity (Mire-Sluis 2001).

Biosimilar products are intended to be as close in all the areas described above as possible to the reference product, with any differences 'not having any meaningful clinical impact (i.e. safety and efficacy)'. Therefore, it is essential during biosimilar development that its potency and biological activity be established as such. Most of considerations for the development of bioassays to use both for the biosimilarity exercise as well as for eventual lot release are not significantly different from those for an innovator program but are essential if the program is to be a success. The specific considerations on how to use bioassays to illustrate biosimilarity (such as using the same assay to compare side by side, statistical aspects etc.) are described in this chapter.

Definition of Bioassay and Potency Assays

In the context of this chapter on biosimilar monoclonal antibodies, a bioassay is an analytical procedure utilizing a biological reporter system (resulting in a biological/functional response), the purpose of which is to measure the amount of active analyte or effective constituent in a biological product, i.e. to determine its biological potency.

The term bioassay should not be confused with a potency assay. Potency is the ability of a material to exert its intended activity and may not necessarily have to be measured in a biological system. Bioassays used for quality can illustrate the batch-to-batch consistency of biological potency of a product as well as defining the actual potency of each lot of product. The amount of product required to provide the optimal therapeutic biological activity in humans, as reflected in the therapeutic dose, is determined by clinical trials. This issue is extremely important as the biological potency measured for quality purposes and therapeutic dose are two very different issues. For example, antibody products that are intended to block the binding of one protein to another can have their potency measured in a binding assay. However, it is often the case that binding alone is not the sole biological endpoint of the product and a cell based format may provide a more relevant assay e.g., prevention of ligand binding to its receptor on the cell surface and induction of Fc activity. In this case, both a receptor based binding assay and a prevention of ligand induction of cell activity bioassay could be used. Cell based assays can be replaced as a potency assay with binding assays such as immunoassays and receptor-ligand methods although they do not measure the ability of a protein to induce a 'functional' biological response. Therefore, a thorough correlation between a bioassay and binding assays is required to show the latter can replace the former as a suitable potency assay, to be discussed later in the chapter.

Types of Biological Activity Assays for Antibody Therapeutics

Biological activity assays can be carried out in vivo or in vitro. The most appropriate method for assessing biological activity is to compare the biological activity of a sample to that of a well-characterized potency reference standard (Sasardic and Mire-Sluis 2000). Where possible, it is preferable to use an assay with a biological role that correlates with a clinical response. Although, it is not always necessary or possible to mimic therapeutic activity for a potency assay that is used to assess quality or efficacy, a justification is needed to the regulatory authorities. Biological activity assays for a biosimilar assessment may or may not have been conducted by the originator. It is critical that both the reference product and the biosimilar are tested in the same assay.

Binding Assays

Ligand Binding Assay

Ligand binding assay (LBA) is an assay, or an analytic procedure, that relies on the binding of ligand or target protein molecules to receptors, antibodies or other macromolecules. A detection method is used to determine the presence and extent of the ligand-receptor complexes formed, and this is usually determined electrochemically or through a fluorescence detection method. This type of analytic test can be used as potency assay to test the biological activity of therapeutic antibodies.

As further complement to LBA, cell-based binding assays (CBA) acts as an indispensable part in determining the exact mechanism of action (MoA) for antibody products. Regulatory agencies expect that cell-free LBA outcomes should be eventually translated into a cell-based format, so that antibody potency and bioactivity can be better assessed in a more biologically relevant environment. The cell-based system can offer many advantages, such as preservation of the native target form, and higher sensitivity towards physiochemical changes (i.e. glycosylation patterns and conformational alternations).

Competitive Ligand Binding

Competitive ligand binding (CLB) assay refers to how therapeutic antibodies exert their biological function by preventing a ligand from binding to its receptor on the cell surface. For these products, a CLB assay offers direct measurement of the product's inhibition of ligand binding to its intended target receptor and may be suitable for potency testing.

The most common type of binding assay is the Enzyme Linked Immuno-Sorbent Assay (ELISA), which can be developed relatively quickly and typically offers robust performance. With the advancement of technology, various "homogeneous" immunoassays have been developed and successfully utilized for potency measurement in QC settings. Examples are Time Resolved Homogeneous Fluorescence Resonance Energy Transfer assays, Amplified Luminescence Proximity Homogeneous assays (such as AlphaLISA) and Proximity Based Electrochemiluminescence Immunoassays. These homogeneous immunoassays eliminate the need for wash steps, and the simple "mix and go" procedures result in decreased assay time and potential analyst error. In some cases, superior signal-to-noise ratio and better overall assay performance, as compared to traditional ELISA, may be achieved. However, custom protein conjugation may be required, and assay performance is highly dependent on the quality of these critical reagents (tagged proteins, donor and acceptor beads, etc.). In addition to immunoassays, Surface Plasmon Resonance (SPR) assays have also been utilized to measure product binding to its intended target. In an SPR assay, protein-protein interaction is detected in real time through changes in mass due to adsorption at the chip surface. Data generated can be used

to calculate the binding constant; therefore, SPR assays can be particularly useful during product development. Although SPR assays have not been used as widely as QC methods for potency measurement but have been adopted sometimes for product characterization, they have been particularly used for biosimilar development as part of biosimilarity assessment in comparison with the reference products.

Bioassays

The choice of bioassay depends upon the nature of the monoclonal antibody, its intended therapeutic use and whether biological activity is only measurable in whole animals. It is obviously desirable to develop in vitro biological activity assays where possible because they offer distinct advantages over using live animals. One also has to consider which assay to choose during the biosimilar antibody development lifecycle, since biosimilars require earlier and more extensive biological characterizations than that of the innovative antibody development. It must also be stressed that one cannot be assured that the removal and use of cells from tissues or the use of clonal cell lines in in vitro formats represents what is occurring in vivo and thus most, if not all, bioassays are a surrogate marker for biological activity.

In Vivo Bioassays

Earliest attempts to measure biological activity often took the form of an in vivo bioassay, where protein was administered to animals and the response in those animals measured. The use of such assays in monoclonal antibody development has occurred, as an example, to understand the overall consequences blocking of ligands, or to understand if an antibody exhibits Fc functions in its activity, in vivo. However, it is difficult to reduce inter-animal variability in estimates of potency, and in vivo bioassays are expensive and labor intensive. In order for an in vivo assay to provide valid estimates of biological potency, a large number of animals are required to account for this variation. A great deal of care and expense is required to decrease any variation through breeding, housing and feeding of animals. A balance must also be maintained between the large number of animals that could be used to provide several data points for potency estimates and humane, ethical and economic pressures to reduce the use of laboratory animals for assays.

It can be argued that testing in vivo provides biological potency tests more relevant to the clinical use of biologicals because a “whole body” approach takes into account bioavailability, serum half-life, toxicities etc. However, this argument is incorrect because biological assays are not intended to mimic the biological activity of a product in the clinical situation. As described, bioassays are intended to be used for quality control and illustrate the batch-to-batch consistency of biological potency of a product (Thorpe et al. 1997). Bioassays are key to the analytical biosimilarity assessment between the reference product and biosimilar product, as well as for the comparability study during biosimilar development.

However, there may be a case to be made for in vivo testing where a combination of physicochemical and biological tests cannot detect differences known to impact on in vivo activity. Such issues could involve complex glycosylation relevant to biological half-lives or modified monoclonal antibodies.

Attempts to avoid the requirement for live animals testing has led to the production of many different formats for the in-vitro estimation of biological potency for a wide range of monoclonal antibodies.

In Vitro Tissue Based Bioassays

The use of in vivo assays as described, can be useful for characterization of a monoclonal antibody. However, one can use an approach that is more stable, yet retains some of the advantages of in vivo assays by developing in vitro bioassays where cells or tissues from animals are cultured in the laboratory and used as responders to the test protein. Assays for monoclonal antibodies that had an impact on the hematopoietic system used cells from the blood or bone marrow. Monoclonal antibodies that act on solid tissues, such as interfering with growth factors and hormones, require assays that involve the removal of the specific tissue on which they act and its homogenization into single cells that can then be cultured and exposed to protein in vitro (Mire-Sluis and Thorpe 1995). However, donor-to-donor variability still occurs in these systems and pure populations of target cells are difficult to achieve.

In Vitro Cell Line Based Bioassays

Using clonal cell lines that respond to specific ligands is a significant improvement as a source of materials for bioassays. The cellular response of ligand dependent cell lines can take a variety of forms, but is most often proliferation or inhibition of proliferation, expression of cellular markers or enzymes, cytotoxicity, or anti-viral activity. The use of murine cell lines increases specificity in some cases, as they may not respond to proteins that are species restricted in their activity.

Taking advantage of recombinant DNA technology has allowed for the cloning of specific receptors and their expression on previously non-responsive cell lines. This can create a specific, responsive cell line for almost any protein with a cellular receptor, without the need to screen a wide range of existing cell lines or tumor cells for responsiveness.

Reporter Gene Based Bioassays

While transfected receptor cell lines can offer selective responsiveness, such lines are still prone to the variability that occurs during the extended periods required for some induced biological function to appear (e.g., cell division, maturation or

cell death). Therefore, the development of bioassays that identify the activation of the genes involved in that function can be much more rapid and robust. The format of these assays is to introduce a plasmid containing a promoter (or rather a relevant region) known to be involved in the expression of genes induced by a test ligand. The promoter region is linked to a reporter gene that subsequently is expressed on ligand binding to its receptor.

The earliest forms of such assays, termed reporter gene assays, used luciferase expression as a marker for gene activation induced by test ligands. This enzyme catalyzes a reaction that results in light formation detectable by luminometers. In recent years, even more sensitive reporter gene systems have been devised, including green fluorescent protein and beta-galactosidase. Due to the shorter time required for significant expression of reporter genes, 2 h as opposed to days for standard bioassays, the assays appear less affected by extraneous influences and are therefore less variable and more precise.

Biological Assay Selection Based on Their Therapeutic Mechanism of Actions (MoA)

Agonistic MoA

Agonistic biotherapeutics, e.g. cytokines, growth hormones, agonistic mAbs, exert pharmaceutical activity by directly binding and activating a cellular receptor. Cell-based assays should be used whenever possible for this class of biotherapeutics. For agonistic mAbs biotherapeutics, if efforts to identify and/or develop a cell-based assay are unsuccessful during the early stages of drug development, regulatory input should be sought for using an alternative non-cell-based potency assay in the interim while a reliable cell-based potency assay will continue to be developed to support later studies.

Antagonistic MoA

Monoclonal Antibodies Binding to Soluble Ligands

A mAb biotherapeutic belonging to this class exerts its pharmacological effect by binding directly to a humoral target (a ligand presents in body fluid), which prevents the interaction of the ligand with the target receptor and obstructs its biological function. Since drug binds to the ligand in circulation and does not involve direct interaction with cells, a non cell-based binding assay suitably reflects the therapeutic MoA and may be deemed acceptable for potency assay.

Anti-receptor MAbs

An antagonistic mAb targeting a receptor directly blocks the biological function of cellular receptor. For example, the therapeutic antibody prevents the receptor specific ligand from binding to it and inhibits the downstream ligand mediated activation of signaling pathway. Since this category of biotherapeutics functions through an inhibitory mechanism and the drug-target interaction involves cells, either cell-based or binding assays may be suitable for potency assay. However, it is important to emphasize that for this class of products, the selection of a proper potency assay format requires a thorough understanding of the structural attributes of the cell surface receptor. If the antagonist targets a hetero-oligomeric receptor, a cell-based assay may be the most appropriate format for potency assessment. In a competitive ligand binding (CLB) assay, the multiple subunits of the purified recombinant receptor may not retain their native structure, function and integrity and would thereby fail to represent the drug-target interaction as it occurs in vivo. Similarly, if the target receptor requires a co-receptor to exert its biological function, a cell-based assay would be a better choice compared to the CLB assay. In a CLB assay, it is not feasible to couple the receptor and the co-receptor in a conformation that resembles their native orientation and structure when associated with the cellular membrane (Hu et al. 2015).

Soluble Receptors/Receptor Fusion Proteins

Soluble receptors or receptor fusion proteins, such as etanercept (Peppel et al. 1991) and abatacept (Korhonen and Moilanen 2009), function as an antagonist by binding the ligand and preventing its interaction with the cognate cellular receptor. When drug binds to the ligand in circulation and does not involve direct interaction with cells, a binding assay suitably reflects the therapeutic MoA may be deemed acceptable for potency assay. However, a cell-based potency assay may be needed when the target ligand is cell membrane-associated.

Multiple or Novel MoA

Besides the above-mentioned classes, “novel” biotherapeutics with diverse “modalities” and more complicated therapeutic MoAs have emerged recently for drug development. These include biotherapeutics with multiple functional domains such as antibody-drug conjugates (ADCs), antibodies with effector function, multi-specific molecules (a biotherapeutic binding to multiple drug targets), oligonucleotides, gene therapies and cell-based therapies, etc. This chapter will focus on defining the potency assay selection for biotherapeutics with MoAs mediated by Fab and Fc domains, such as ADCs and antibodies with effector function. Oligonucleotides, gene therapies, and cell-based therapies are not covered here, although the basic

principle of designing assays based on the MoAs would still apply for these treatment remedies. For multi-specific biotherapeutics, separate potency assays may be needed for measuring biological responses specific to each functional domain. There are exceptions when only one potency assay is required for development if the potency of the respective domain function can be measured in one multiplex assay. Compared to the potency assay specific to each functional domain in separate assays, the multiplex approach monitors the simultaneous interaction of drug with its respective targets in one single assay and thereby closely mimics the therapeutic MoAs. This would allow more precise potency assessment since additional steric hindrance is observed when drug binds to multiple targets instead of one target.

Antibody Drug Conjugates (ADCs)

ADCs are unique immunoconjugates that couple a cytotoxic drug to a monoclonal antibody through a peptide or small molecule linker for targeted cancer therapy (Sievers and Senter 2013). An ADC binds to a specific tumor marker on cancer cells via mAb to deliver the cytotoxic payload or drug through cell membrane internalization for therapeutic intervention. For ADC molecules, a cell-based potency assay using read-out associated with cell death or proliferation that reflect the therapeutic MoA is a scientifically justified potency platform. In addition, a binding potency assay is also required that can measure either ADC binds to a specific tumor marker or the binding to both naked Ab portion and cytotoxic payload portion of the ADC.

Effector Function of MAbs

Therapeutic antibodies may manifest their clinical efficacy through effector functions such as antibody-dependent cell-mediated cytotoxicity (ADCC), antibody dependent cellular phagocytosis (ADCP), and complement-dependent cytotoxicity (CDC) (Jiang et al. 2011; Wu et al. 2015). MAbs with ADCC and CDC function induce target cell lysis following specific binding of antibody variable region to the target antigen on cell surface and the interaction of antibody Fc with either FcγRs on effector cells or complement (Jiang et al. 2011). Since target cell lysis is the end-point resulting from the drug's pharmaceutical activity, cell-based effector function assays which directly measure target cell lysis would be reflective of the relevant MoA for this category of antibody biotherapeutics. However, such effector function assays often suffer from relatively higher variability compared to regular cell-based assays due to the employment of two cell types (i.e. target and effector cells) (Schnueriger et al. 2011). If a cell-based effector assay is not feasible to serve as potency assay, other assay formats such as a cell-based binding assay or a non cell-based CLB assay may be acceptable to serve as the surrogate assay. However, cell-based effector assay should be used for biological characterization to bridge with the binding assay.

Biological Assays in Biosimilar Product Life-Cycle Management

Biological assays of various types can be used at the very early stages for establishing the quality target product profile (QTPP) of the reference products as well as developing biosimilar products to match the QTPP of reference products. These can provide information of the types of activities shown by the reference product. During product development, bioassays are invaluable for investigations aimed at characterizing the biological activities of the product and for stability, dosing and formulation studies. After product development, the assays are used to show batch-to-batch consistency and product (final form) stability under the proposed storage conditions as well as defining the biological potency of the product (Mire-Sluis et al. 1996).

Different bioassays may be needed for different purposes at different stages of biosimilar monoclonal antibody product development. Biological assays used early in product identification and characterization may need to be designed to maximize biological relevance and information content and could be less precise, accurate and rugged as well as being different from assays used later in product development to determine potency, stability and batch-to-batch consistency.

For example, the impact of glycosylation on a biosimilar monoclonal antibody either increased or decreased in comparison to the reference product can be studied by exploring the impact on product potency using in vivo, binding, antibody dependent cell-mediated cytotoxicity (ADCC), or complement dependent cytotoxicity (CDC) assays. The function of other structural characteristics of a product (e.g., size) should also be understood and included in the design of comparability studies.

For the development of both biosimilar and innovative molecules, biological assays, whether binding or cell-based functional assays, should be run concurrently so that a developer can select the appropriate potency assay in later development and have a body of data to support that choice. For example, if a bioassay is deemed too variable or not optimal for quality control testing the manufacturer should consider switching back to the binding assays. However, when replacing a bioassay with a binding or other surrogate assay, data must be gathered to demonstrate a strong correlation between the assays. Therefore, you should develop a cell-based assay as soon as possible because of the time required to do so and to gain experience with the cell line. This also allows ample time to gather correlation data between the assays, which should mitigate the regulatory risk associated with a poorly justified method. A parallel path can be built between the bioassay and a binding assay that justifies the selection of the final method.

The final release potency method should be “locked down” and in place prior to the pivotal trial. This would include the final internal controls of the assay (e.g. assay suitability) as well as the ranges in which the assay can be performed appropriately. This presents some advantages for either an innovator or biosimilar product. This provides a good deal of experience with the final potency method before submission, which offers a true estimate of method performance and success rate. In addition,

fewer validations and bridging studies can be performed if a small number of methods are used throughout the clinical phases. However, it may be necessary to use multiple potency assays until a clear understanding of the product attributes and MoA has been achieved. Ultimately, the most well characterized, precise bioassay reflective of the mechanism of action is generally selected as the lot release potency assay to support commercialization of the product. Regardless of the final assay format, appropriate design, validation, and analysis are necessary if an assay is to provide reproducible and meaningful data.

Cell Based Assay Optimization and Method Remediation

For a cell-based bioassay, one of the most important factors is to choose a cell line that responds well to the drug and that response should be durable. The cell line should thus be stable, meaning that cell growth and response to the drug are consistent over time. This requires an understanding of the cellular growth patterns and receptor expression kinetics. A developer should determine how cell responsiveness and receptor expression are affected by passage number, cell density, feeding schedules and days in culture. Establishing these cell traits during development can help ensure a consistent and robust cellular response to the drug. The output used to measure cellular activity (e.g., fluorescence, luminescence) should be quantitative and indicative of a robust cellular response. Therefore, a primary goal in cell line and output selection should be to maximize the signal-to-noise ratio of the response.

Carefully controlled bioassays are technically demanding, relying heavily on the competence of staff carrying out the assays to accurately and reproducibly dilute and pipette solutions. However, automation of both bioassays and immunoassays have been particularly successful, but the capital investment is large. Therefore, the design of any bioassay must take into account factors that introduce variability and the analysis of bioassays must test for variability if results are to be statistically valid. A titration of the test material has to be made and compared to a titration of a reference material, with particular attention paid to comparisons of the linear portion of the dose-response curve. At least three points on the linear portion of the dose response curve are required to compare sample and reference curves.

Proper assay design also integrates multiple strategies to minimize variability and bias. There should be as few handling steps and reagents as possible to minimize dilution or technical errors. Most bioassays and immunoassay use microtiter plate that are particularly prone to position effects that can result in variability of data. To reduce the effects of position within microtiter plate assays, randomization of the position of sample titration curves within plates is recommended, as is the inclusion of a standard reference preparation on each plate; again, preferably in different positions. The use of coded duplicates in the assessment of variability and bias is particularly valuable.

A Reference Standard and Assay Control should be established as early as possible for continuous trending of assay performance. When performing a bioassay for biosimilar, an in-house Reference Standard, often from one of the GMP drug substance batch, needs to be qualified and established against pre-defined acceptance criteria. Bioassay always uses the Reference Standard to establish system suitability and assay acceptance criteria. Concurrently, bioassay also needs to use an "Assay Control", which can be a lot of reference product. Thus, one always runs both biosimilar and reference product, side-by-side, and use the same passage of cells. The system suitability, or acceptance criteria, of a bioassay should be sufficient to ensure that the assay remains in control between runs. System suitability criteria often include requirements for cell viability, cell count, passage number, the signal-to-noise ratio, internal control potency, and parallelism, but may include any parameter that is determined to be important in minimizing inter-assay variability. Additionally, several statistical tools can be used to improve assay robustness.

These studies are similar to validation testing, but they are not as protocol driven and are performed at earlier phases in development to demonstrate that they are suitable for use. Proper robustness studies are also key to method transfer and performance trending because they establish the method variability that may exist between runs without detriment to the results. An experienced biostatistician can aid in experimental designs to determine the component variance analysis, or the factors in the assay that contribute most to variability. Assay performance can be improved significantly by understanding and controlling for these factors.

Proper analyst training is also of great importance in delivering consistent and reliable assay results. Because an analyst is generally one of the most significant sources of bioassay variability, the focus of training should be to limit this source of variability to whatever extent possible. Implementation of these practices will yield a bioassay that is well controlled and usable as a quality control release assay.

When designing an immunoassay, one has to consider whether proteins on a plastic surface or in solution have the same affinity as those on a cell or in solution. In addition, are the receptors/ligand oriented in the same way? This can be an issue when considering the hydrophobicity of full length receptors. Reproducibility of plate coating is also important to assess.

During immunoassay development, it is also important to judge if binding of a monoclonal antibody to a receptor/ligand on a plate reflects functionality. Binding to a protein on a plastic plate does not always mean the product is functional (there could be sticky degradants, charged variants etc.)—the same applies to a monoclonal antibody binding to its ligand on a plate that might not always assure the ligand is then neutralized in vivo.

Method Validation

The appropriate validation of any assay used for the characterization and release of biosimilar monoclonal antibodies is critical. Even though there are general regulatory guidelines for assessment of the validity of an assay, details contained

in the US Pharmacopoeia are specific to bioassays. However, it is up to the assay developer to use these guidelines and develop in house protocols based on sound scientific principles and the nature of the assay. Assay characteristics with associated acceptance criteria such as reproducibility, robustness, signal to noise ratio etc. should be contained in a predefined validation protocol. When more than one biosimilar monoclonal antibody is produced in a facility, specificity should be part of the validation criteria.

The fundamental condition for any assay validity however is the condition of biosimilarity of sample and reference standard; that is, the dose–response relations (i.e. slope, asymptotes etc.) for the sample and the reference standard should be identical. During assay development, substantial information about dose–response curves should be collected in order to select an optimal dose range for potency estimation assays. After such data are available, analysis in terms of log doses is often found preferable to analysis in terms of absolute units.

The optimal assay range is often chosen in a linear (or linear under suitable transformation) part of the log dose–response relation. In such a situation, the condition of biosimilarity becomes a condition that the log dose–response line for the sample should be parallel to that for the reference standard, i.e. a parallel line assay. Provided at least three or ideally more doses of each preparation are included in the assay, the conditions of linearity and parallelism of the log dose–response lines can be tested in the individual assay. Moreover, the slope of the line or other characteristics of the responses may also provide information about conformity or otherwise with the previously determined complete dose–response relation.

Various assumptions about the statistical nature of the assay response data must be satisfied if estimates derived from such analyses, and the tests for the conditions of linearity and parallelism given by such analysis, are to be valid.

One must always assume that the "experimental units" providing the response represent a random selection from a defined population of such units. For example, results obtained for cells in a microtiters plate applied earlier may be different from those applied later. Units may differ because of a temperature, oxygen or humidity gradient across the plate. If an assay extends over several microtiters plates such differences between wells become even greater so these and other factors must become part of the definition of the "experimental unit".

During assay validation, one should also assess the precision of an assay. Precision is a measure of assay variability as it illustrates how similar the results of an assay are when several estimates of potency are provided. This should not be confused with accuracy, which is a measure of how close an assay result is to the 'correct' result. An assay can be very precise, but give the wrong answer i.e. is inaccurate.

Assay repeatability is the precision of the assay internally (intra assay variability)–how repeated estimates within a single assay compare with each other. Intermediate precision is the precision of the assay when performed on different occasions or by different analysts, but within the same laboratory (inter-assay variability); i.e. How the results of independent bioassays compare to each other. Lastly, reproducibility is the precision of the assay when performed by different laboratories. This is measured by collaborative study between laboratories.

Critical Reagents

Reagents that are deemed critical through assay qualification studies should be well characterized and tightly controlled, monitored, and thus ‘qualified’. Whenever feasible, critical reagents should not be single-sourced, meaning that they should be available from more than a single vendor. This precaution will prevent an inability to perform assays if one source is suddenly unable to provide the quantity or quality of the reagent required. The stability of critical reagents has to be assessed so a shelf life can be applied that ensures new material is made before the original material becomes ineffective.

With all cell line based assays, careful evaluation of the stability of the cell line should be carried out. Cell lines can often lose their biological responsiveness over time, so it is important to have a well-characterized cell bank and some idea of how long a line can be passaged before its response becomes compromised. Therefore, cell lines should normally be cultured for a pre-specified period or number of cell doublings and then replaced with an early passage of cells.

Different batches of fetal calf serum used to maintain cell cultures can greatly affect the performance of bioassays and should be carefully screened prior to use. Some batches of sera can provide excellent maintenance of cell lines (i.e. cells grow rapidly), but result in poor bioassays with high backgrounds or low stimulation indices. Screening of sera should include both performance in cell maintenance and in bioassays.

For both immunoassays and cell based assays, the type of microtiter plate and its materials of construction cannot be underestimated. Changing from one manufacturer to another can impact the ability of adherent cells to stick to the plate, the same for antigen or antibody coatings in immunoassays.

For immunoassays, receptor or ligands used are usually critical reagents, as are any conjugated antibodies used in the assay. Each should be well characterized before use.

Whenever critical reagents are identified during assay development, it is necessary to ensure enough time is provided between replacing the reagents to assure the assay still performs as expected. A suitable critical reagent replacement protocol should be in place to follow when the time comes, with associated predefined acceptance criteria.

Legacy Potency Assay and Assay Replacement/Comparability

Although it might not appear relevant for a biosimilar product for its initial approval, one must always consider that the original bioassay could be replaced. For example, a cell based assay with an immunoassay, or to a more robust cell line, as one gains more experience with the product—even if the innovator was using a cell based assay at the time of the biosimilar approval.

Regardless of product type or history, the replacement of a bioassay in potency testing is not possible without a strong body of data that strongly correlates product activity between assays. It is advisable to start putting that together early in development, devising a parallel path that provides extensive experience with all assays and includes testing of multiple lots and product variants. The data must be combined with a strong knowledge of the MoA to demonstrate that product potency is well represented by the surrogate assay(s).

An interesting aspect of the replacement of a bioassay in potency testing is the fact that a robust and precise bioassay is required for the effort to end in success. It is unrealistic to expect you could replace a poor bioassay with a quality surrogate assay. That is because the results of a surrogate assay will not be accurate and reliable unless they were correlated to a bioassay that possessed those features in the first place. Given the value of a quality bioassay, this could represent an interesting dilemma to any company considering the use of a non cell-based assay for potency testing of biosimilar monoclonal antibodies.

When replacing one assay with another, one should test a set of samples side by side to justify that the assay is able to detect changes in potency if the product differs in some way. These samples can include: several lots of product (bulk and FDF, SKU's), existing product variants (aggregates, oxidized etc.), temperature degradation over time, freeze/thaw, light exposure degradation over time, proteolytic degradants, pH exposure, glycosylation variants and any relevant in process materials.

In addition to lot release, stability testing has to be considered when changes to a bioassay are being proposed. For stability testing, a surrogate assay may not need to be the most sensitive assay for product change, but it must be able to detect all aspects of change that are important for potency. Of course, extensive characterization studies must be performed to determine what changes occur to a product over time and the impact of each on its potency. It may then be possible to combine this knowledge with risk assessments to ensure that the surrogate assay provides necessary coverage for accurate and reliable potency testing.

To assure regulatory acceptance of any assay comparability plan, a comparability protocol should be written outlining:

- Detailed description of both assays
- Statistical plan
- Sample plan
- Testing plan
- Data to be presented
- Acceptance criteria

Discussing any bioassay replacement plan with regulators ahead of time is advisable.

Method Tech Transfer

Transferring a method between sites often occurs as one moves from the clinical manufacturing site to the commercial scale site. Whilst there is an increasing movement in the rapid development of biosimilar monoclonal antibodies to do both clinical and commercial manufacturing at the same site (thus alleviating the need to transfer assays) for approval, method transfer would still have to occur if import testing was to be required as products are approved globally.

Therefore, a company should have specific protocols for how to transfer an assay to ensure it performs in the expected way between sites and continues to do reproducibly over time. Having drift between assay sites can cause considerable issues if not monitored for and addressed in a timely manner.

A suitable protocol should be created with the necessary contents similar to assessing comparability between to assays as described above (e.g. sampling and analysis plan, acceptance criteria etc.). It is highly recommended that both the sending and receiving site execute the protocol on the same samples and results examined carefully for any bias, even if acceptance criteria are met. A drive to simply use existing sending site data as the comparator to the receiving site can cause issues if the same samples are not executed around the same timeframe (i.e. sample stability can shift results).

The type and number of samples should be based on the known inherent variability of the assay and should be derived in consultation with a statistician.

Special Considerations of Biological Assays Comparing Innovative Biologics and Biosimilar

Regulatory Expectations and Current Practices on Potency Test

The regulatory pathway for biosimilar medicines is a unique and thoughtful process. It is designed to help ensure the development and approval of high-quality biosimilar medicines. Approved biosimilar medicines should have no clinically meaningful differences in terms of safety and efficacy from the relevant reference product, based on the totality of evidence from analytical, nonclinical, pharmacokinetic, and clinical studies. The totality of evidence represents a new approach by the FDA to the development of a new biologic product.

Robust analytical testing, including comparative structural and functional characterization, should be employed to establish high biosimilarity of the biosimilar and the reference product. Nonclinical testing will be used to evaluate the toxicity and safety profiles of the biosimilar. Comparative human pharmacokinetic and pharmacodynamic studies and clinical immunogenicity assessment will also need to be established. If residual uncertainty exists, comparative trials may be required (based on recent approvals, the FDA has required studies).

Functional assays can serve multiple purposes in the characterization of protein products. These assays act to complement physicochemical analyses and are a qualitative measure of the function of the protein product. Depending on the structural complexity of the protein and available analytical technology, the physicochemical analysis may not be able to confirm the integrity of the higher order structures. Instead, the integrity of such structures can usually be inferred from the product's biological activity. If the clinically relevant MoAs are known for the reference product, the functional assays should reflect these MoAs. Multiple functional assays should, in general, be performed as part of the analytical biosimilarity assessments. The assessment of functional activity is also useful in providing an estimate of the specific activity of a product as an indicator of manufacturing process consistency, as well as product purity, potency, and stability.

If a reference product exhibits multiple functional activities, a set of appropriate assays designed to evaluate the range of relevant activities for that product should be performed. For example, with proteins that possess multiple functional domains expressing enzymatic and receptor-mediated activities, one should evaluate both activities. For products where functional activity can be measured by more than one parameter (e.g., enzyme kinetics or interactions with blood clotting factors), the comparative characterization of each parameter between products should be assessed.

It is recognizable that some types of biological assays have potential limitations, such as high variability, that might preclude detection of small but significant differences between the proposed biosimilar product and the reference product. Because a highly variable assay may not provide a meaningful assessment as to whether the proposed product is highly similar to the reference product, efforts should be made to develop bioassays that are less variable, more sensitive to changes in the functional activities of the product. In addition, in vitro bioactivity assays may not fully reflect the clinical activity of the protein. For example, these assays generally do not predict the bioavailability (pharmacokinetics and biodistribution) of the product, which can affect pharmacodynamics and clinical performance. Also, bioavailability can be dramatically altered by subtle differences in glycoform distribution or other posttranslational modifications. Thus, these limitations should be taken into account when assessing the robustness of the quality of data supporting biosimilarity and the need for additional information that may address residual uncertainties.

How to Select a Potency Assay

Potency assays play a pivotal role in determination of potency of protein products. As required by U.S. regulation, an assessment of potency is required for the licensure of the biopharmaceuticals defined in 21 CFR 601.2. Ideally, a potency assay should reflect the product's MoA, be sensitive to changes in product critical quality attributes, and stability indicating. The potency test should be validated

as per ICH Q2 (R1). Developing a robust, sensitive, and relevant potency assay represents a substantive challenge both in planning and execution. Selecting the best potency assay format (i.e., in vivo or in vitro) should be based on scientific knowledge of the product-target interactions, therapeutic effect elicited through the product-target interaction, and the assay performance itself based on the status of assay's validation and qualification. System suitability and assay specification acceptance criteria are usually set as a numerical range and should be adjusted throughout the product development to reflect the manufacturing and clinical experience.

Several regulatory and guidance documents are published by the Food and Drug Administration (FDA), the International Committee on Harmonization (ICH), the United States Pharmacopeia (USP) (2012a, b), and the European Pharmacopeia (Ph. Eur.) (European Directorate for the Quality of Medicines 2004) to cover different aspects of bioassay validation. Despite the availability of these documents, often there are questions related to implementation and interpretation of these guidelines. Assay validation demonstrates that the assay, when performed per the SOP, is adequately precise and accurate for use in product release and stability studies.

The legacy potency assay employed by the originator product can always serve as a starting point for potency assay selection for biosimilar development. Nonetheless, it is possible that fast, homogeneous and precise bioassays reflective of product's MoA can be used to replace the variable legacy potency assay. In this case, a method "bridging" or "comparative" study may be needed to demonstrate the equivalent performance of two methods in detecting changes impacting bioactivity and demonstrating similar stability indicating properties.

Reference Standards

The design of these assays and calculation of relative potency for a product rely heavily on Reference Standards. This is not to be confused with 'reference *product*' to which the biosimilar is compared to the innovator (reference) product. Selecting and establishing the right material to serve as the Reference Standard is important. The biological response of a test sample is directly compared against the Reference Standard in a potency assay. Thus, the Reference Standard is ideally generated from a similar manufacturing process as the test sample and with known stability data under intended storage conditions. Moreover, the Reference Standard should be evaluated thoroughly through multiple runs in the potency assay ($n > 10$) to establish a "normal" range for EC_{50}, hill slope, and upper and lower asymptotes when the assay uses a 4-PL data-fitting model commonly used for potency assay evaluation. When the Reference Standard is deemed appropriate for a given assay, allocate sufficient quantities of material for future assays. It is likely that the material will be used not only for assay development and validation, but also for sample testing when its shelf life allows. When the current lot is close to depletion, retain some samples for use in a bridging study to compare with the new Reference Standard.

If there is a suitable, publicly available, and well-established Reference Standard for the protein product, a physicochemical and/or functional comparison of the proposed product with this standard may also provide useful information. Although studies with such a Reference Standard may be useful, they do not satisfy the BPCI Act's requirement to demonstrate the biosimilarity of the proposed product to the U.S.-licensed reference product. For example, if an International Standard for calibration of potency is available, a comparison of the relative potency of the proposed product with this potency standard should be performed. As recommended in ICH Q6B, an in-house Reference Standard(s) should always be established, qualified and used for control of the manufacturing process and product.

An International Reference Standard, when applicable, can be obtained from a nationally or internationally recognized source. Alternative material, including material generated in-house, may be qualified and designated as a Primary or Working In-House Reference Standard. In-house primary Reference Standard material must be prepared from lot(s) representative of production and clinical batches, and qualified following established procedures that include characterization testing requirements and specifications/assay acceptance criteria, as well as stability testing procedures. A process is established for succession planning of Reference Standards.

In summary, analytical studies carried out to support the approval of a proposed product should not focus solely on the characterization of the proposed product in isolation. Rather, these studies should be part of a broad comparison that includes, but is not limited to, the proposed product, the reference product, applicable Reference Standards, and consideration of relevant publicly available information.

Number of Lots Required for Physiocochemical and Functional Biosimilarity Studies

Extensive and robust comparative physicochemical and functional studies should be performed to evaluate whether the proposed product and the reference product are highly similar. A meaningful assessment as to whether the proposed product is highly similar to the reference product depends on, among other things, the capabilities of available state-of-the-art analytical assays to assess, for example, the molecular weight of the protein, complexity of the protein (higher order structure and posttranslational modifications), degree of heterogeneity, functional properties, impurity profiles, and degradation profiles denoting stability. Physicochemical and functional characterization studies should be sufficient to establish relevant quality attributes including those that define a product's identity, quantity, safety, purity, and potency. The product-related impurities, product-related substances, and process-related impurities should be identified, characterized as appropriate, quantified, and compared with multiple lots of the proposed product to multiple lots of the reference product, to the extent feasible and relevant, as part of an assessment of the potential impact on the safety, purity, and potency of the product (Food and Drug Administration (FDA) 2012).

In general, at least 15 lots of reference product and 10 Drug Substance lots of the proposed biosimilar product are required to be used for analytical biosimilarity assessment—although this is highly dependent on the methods use and their variability (Tsong *et al.* 2017). Therefore, it is strongly recommended that a statistician be consulted to select the appropriate, statistically valid, number of lots to show biosimilarity.

Addressing Assay Variability

A successful bioassay suitable for validation and final-product lot release may take multistage development and fine-tuning to reach a final design. Although many roadblocks can present on the way to a robust bioassay, controlling variables at early stage assay development and careful quality control in assay performance are key to a meaningful potency test to ensure product quality.

Here we focus on potential aspects to consider when building a consistent potency assay that is suitable as a release test.

Cell Type/Cell Line Selection

To develop a cell-based potency assay, there are many factors that need to be considered. Firstly, determine which cells are appropriate. If possible, select a type that is relevant to a product's MoA and is known to respond well to the product. For instance, when developing a mAb that binds to a cancer cell marker and subsequently leads to growth inhibition of target cells, screen several malignant cell lines that express that marker. The most responsive cell line should be selected, although one must consider stability of the cell line to the response above sensitivity. There is no benefit in having a cell line that exhibits a strong response if it is either highly variable or loses reactivity too quickly.

Primary cells in general should not be used because of their potential for lot-to-lot, donor-to-donor variability. However, in some cases where primary cells must be used, consider appropriate approaches to minimize cell heterogeneity. That can be done by securing a large lot of cells or isolating a subpopulation when feasible. Ready-to-use frozen cells can be helpful in reducing assay variability.

Peripheral blood mononucleated cells (PBMCs) are commonly used in bioassays for product characterization. But PBMCs lack consistency in potency tests in general, primarily because only a subset of cells generates the response of interest. Furthermore, the percentage and activity of different subpopulations of PBMCs vary from run to run and between lots. Instead, there should be attempts to isolate a desired cell population and use the "purer" cells in a potency assay.

Selected cell lines need to be extensively characterized. Information about cloning history, genetic stability, gene copy number, growth characteristics, and passage limits all should be established. At a minimum, evaluate passage limits and

vial-to-vial consistency in the potency assay. In addition, create and store phase-appropriate cell banks. It is not unusual to use a research-grade cell bank for early phase potency assay development. However, when a product progresses to phase 2–3, it is critical that you make and fully characterize a cell bank generated under a more controlled laboratory environment. Whenever a new bank is generated—in addition to the standard purity and identity testing—test cells from that new bank in the assay to ensure that the assay parameters are comparable with the current bank.

Lastly, ensure that cells are in the necessary physiological state and behave in the potency assay as expected. For suspension cells, establish the minimum and maximum cell density for culture maintenance. Spent media should not be used as it may impact cell growth and metabolism, and cause unwanted cell selection. It is also important to not under or over trypsinize adherent cells as it can potentially damage the cell membrane. Cells should not be allowed to grow over confluent to prevent potential cell transformation.

Procedural Accuracy

Because of the inherent, non-robust nature of potency assays, a robust potency assay requires the use of well-defined and accurate procedures. From a stock solution, both a Reference Standard and test sample are diluted over multiple steps to the final working dilution (concentration) range tested in the assay. In addition, a potency assay involves pipetting cell suspension onto 96-well microplates and mixing with other reagents. Without accurate pipetting, there is no solid foundation for a robust potency assay.

We do not discuss pipetting techniques at length here, but rather offer a few quick points to consider. First, work with a volume that is close to each pipette's calibration volume. Second, use prewet tips to increase consistency. Third, except for cell suspension, all reagents should be at room temperature for accurate pipetting. Last, use reverse pipetting when dealing with viscous liquids.

Incubation temperature and time should be well controlled. By contrast with an assay performed in an R&D environment, a potency assay must have a well-defined range for acceptable incubation temperature and time. Many good laboratory practice (GLP) or GMP laboratories have incubation chambers (incubator, refrigerator, or freezer) for 37 °C, refrigerated, or frozen conditions but no chambers for room temperature. As a result, plates are placed on the bench top for room-temperature incubation. This "room temperature" can range from 20 to 35 °C, even 15–40 °C. Fluctuations across the range of temperatures can significantly affect assay outcomes. For incubation steps that are performed at room temperature, using an incubator set at 20–25 °C can reduce assay variability. As for incubation time, do not use a wide range of times for critical incubation steps, if possible. For example, a 60 $\pm$ 10 min time window is much better than 1–2 h.

Consistent washing steps are essential for controlling assay background and precision between replicate wells. Whether using manual washing or an automated plate washer, be consistent and allow only one washing step method in the

procedure. When an assay requires manual washing, ensure that all analysts wash plates in a similar way—working through the plate at the same orientation, adding wash buffer at similar speed, and washing adjacent rows at similar intervals. When using a plate washer, make sure the same setting is used every time.

Proper and timely calibration and maintenance of equipment also can contribute to procedural accuracy. All equipment used in GMP assays should be validated for their intended use.

Assay Training

For an assay that is not completely automated, the analyst is the largest source of assay variability. This is especially true for a bioassay that involves multiple dilution steps and manipulation of test sample, cells, and reagents. Onsite training can be conducted when transferring an assay to a different laboratory. This training provides the personnel from the sending and the receiving laboratories an opportunity to observe each other. Cross-training allows analysts to identify steps that might not be documented in an assay's standard operating procedures but are important to assay performance. On many occasions, the sending lab SME can provide information about equipment or reagents that differ between the sending and receiving laboratories.

When an assay is performed infrequently, a periodic requalification program can familiarize analysts with assays and prevent potential assay failure due to long gaps between assay performance. The frequency of requalification depends on the complexity of the assay and the proficiency of the analyst. Generally, if an analyst has not run a given assay for 6 months, a requalification run should be performed before performing a GMP release test.

Data Analysis

The design of a bioassay that reports a relative potency value for a test sample against the Reference Standard takes into account run-to-run variability to some degree. Some assays are still highly variable despite thorough evaluations of the sources of variability. That is possibly attributed to the wide and unpredictable biological response being measured. For such assays, averaging final potency results from two or three independent setups or runs can be a useful approach to reduce the risk of the assay results being influenced by random factors. This strategy has been adopted by many scientists developing potency assays, especially for effector assays such as an antibody-dependent cell-mediated cytotoxicity (ADCC) assay or a complement-dependent cytotoxicity (CDC) assay. In such cases, the assay is qualified or validated based on two or three runs, the same as described in governing documents.

Assay Troubleshooting

There are several approaches for troubleshooting a non-robust bioassay. Dissecting a complex bioassay to individual steps is sometimes very helpful. When the response from cells plated in a 96-well microplate is measured after incubation with a number of reagents, evaluate the response after each step, if possible, to identify the problematic step in the procedure. Starting from a base plate with cells only often provides some clues such as position/edge effect or uneven cell seeding or growth.

A design of experiments (DoE) study is a useful tool to evaluate multiple variables systemically. You can perform DoE at the assay development stage to identify optimal assay conditions or for assay troubleshooting. For example, ligand concentrations, incubation time, and cell density all can be incorporated into one DoE, rather than be part of separate evaluations. DoE enables assessment of the impact from related experimental conditions that cannot be achieved by changing variables one at a time.

Data trending should be implemented to monitor performance of a potency assay. Key factors that could potentially affect assay outcomes, such as operator, cell seeding and harvest density, passage number, material lots, and equipment identification should be recorded. Other assay parameters such as EC_{50} values, hill slopes, and upper to lower asymptote ratios can also be trended. Those data often can answer questions such as (1) What has changed from when the assay was running well? (2) Is there a trend? And (3) What is the most likely root cause for the assay failure?

Data trending also helps detect data shift or drift before a system suitability failure or out-of-specification or out-of-trending event. Once a trend has been identified, preventative actions should be taken to prevent assay failure.

Setting limits goes hand-in-hand with data trending. For example, when a trend shows that an assay does not work well once cells have been cultured for more than 20 passages, then set a cell passage limit in that assay protocol. Knowing method limits such as cell passages, specific reagent lots expiry, and analyst-specific parameters is valuable and helps exclude potential factors that could introduce variability.

Critical Quality Attributes and Their Relationship to Potency

The critical quality attributes of a reference product, how each impacts safety, efficacy, pharmacokinetics, and overall quality, is fundamentally important to producing a high-quality biosimilar. A biologic drug is extremely complex and typically has more than 1 hundred features or "attributes." Some of these attributes are important to the different ways the body can recognize proteins and are therefore critical to the safety, efficacy, and pharmacokinetics of the drug. These are known as "critical quality attributes."

An understanding of which attributes are important to each function for each product is important to obtain the best possible match. Some attributes work individually to drive a biological function, and some work in a composite manner. A biosimilar will not be exactly like its reference product and some features will not match, but the critical quality attributes need to match so that the biosimilar medicine and the original biologic work in the same way, that is, have the same biological function for every patient. Similarly, a structural match is also desirable, but it is feasible to not have a precise structural match while still preserving function. However, an understanding of how physicochemical attributes especially ones such as glycoforms and charged variants impact potency, should be part of biosimilar product characterization. For each biological therapeutic, the differences between the reference product and the biosimilar in the key critical attributes should be the primary focus as they are expected to drive the "potential for biological differences". Any bioassays that can help resolve the residual uncertainty would help define the clinical development of the proposed biosimilar.

Biological Dose-Response Modeling, Parallelism and Data Analysis

The potency of a biological therapeutic is often determined relative to a Reference Standard, such as via parallel line analysis. Measurement of relative potency is only meaningful if the test sample behaves as a dilution or concentration of the Reference Standard, and exhibits a parallel relationship to the Reference Standard. Such similarity is called parallelism. Graphically, parallelism is observed where the dose-response curve of the sample is a horizontal shift of that of the Reference Standard on the logarithmic dose axis. The amount of shift represents the logarithm of relative potency (USP 2012b). As a necessary sample acceptance criterion for bioassay, there is a need to assess parallelism before the results of a bioassay are interpreted. The requirement for the evaluation of parallelism appears in both the United States Pharmacopeia (USP) (USP 2012a, b) and European Pharmacopeia (EP) (European Directorate for the Quality of Medicines 2004).

Parallelism is a necessary condition for the relative potency of a bioassay to be meaningful. Difference and equivalence tests are two major statistical methods used for parallelism testing, with the latter being recommended for use in the revision of USP Chapters <1032> and <1034>. The recommendation is largely motivated by the criticism that the difference test may reject parallelism even for insignificant differences when the sample size is large or the assay is too precise, and that it fails to reject parallelism of non-parallel curves when sample size is small or the assay is imprecise. Therefore, the method rewards assays of small sample sizes and large variability, and thus does not offer adequate protection to consumer's risk. From a compliance perspective, the equivalence test may be the preferred method because it makes the control of consumer's risk possible, and encourages

the manufacturer to improve its assay so as to provide better protection to the producer's risk. However, implementation of this method can be challenging for laboratories that lack experience in statistical analysis and software development. Development of such a parallelism testing enabling tool is important for a laboratory to be compliant.

A customized assay analysis template that is incorporated into a fully GMP compliant software package (Yang et al. 2012). The template automates USP-recommended parallelism testing method based on a 4PL model, and it is simple to use. It makes the implementation of the USP guidance both practical and feasible. A case study demonstrates that the equivalence test can fail non-parallel samples and pass parallel samples. The tool can easily be generalized to bioassays with other types of non-linear response data such as 5-parameter logistic function. Overall, we show that an equivalence approach for parallelism testing, as recommended by USP, can be implemented in a simple, QC-friendly, compliant, and validatable manner.

Analytical Biosimilarity Assessment

FDA currently recommends the use of a statistical approach to evaluate quality attributes of proposed biosimilar products that is consistent with the risk assessment principles set forth in the International Conference on Harmonization Quality Guidelines Q8, Q9, Q10, and Q11. Consistent with these principles, FDA recommends an analytical biosimilarity assessment that is based on a tiered system in which approaches of varying statistical rigor are used (Tsong et al. 2017; Christl 2015).

One approach to determining the tier to which a particular quality attribute would be assigned will depend upon a criticality risk ranking of quality attributes with respect to their potential impact on activity, PK/PD, safety, and immunogenicity with quality attributes being assigned to tiers commensurate with their risk.

- For quality attributes with the highest risk ranking (Tier 1), equivalency testing would be recommended and generally would include assay(s) that evaluate clinically relevant mechanism(s) of action of the product for each indication for which approval is sought.
- For assessing quality attributes with lower risk ranking (Tier 2), FDA recommends the use of quality ranges (mean $\pm$ X σ, where X should be appropriately justified.
- For the lowest risk ranking (Tier 3), FDA recommends an approach that uses raw data/graphical comparisons

In addition to criticality, other factors should be considered in assigning quality attributes and assays to a particular tier using this approach. This could include, but is not limited to, the levels of the attribute in both the reference product and proposed biosimilar product (as determined by the biosimilar sponsor's testing), the sensitivity of an assay to detect differences between products, if any, and

an understanding of the limitations in the type of statistical analysis that can be performed due to the nature of a quality attribute. Therefore, while many attributes may be considered high risk, not all would need to be included in Tier 1 testing. FDA recommends that sponsors submit their proposal for ranking attributes that will be assessed in each tier to gain agreement from the Agency prior to performing the statistical assessment.

FDA also recommends that sponsors carefully assess their analytical biosimilarity plan to identify and address any other factors that could potentially impact the ability to demonstrate that a biosimilar product is highly similar to the reference product. For example, considering the ages of the biosimilar and reference product lots tested, optimizing assays, and pre-specifying the criteria under which wider biosimilarity acceptance criteria for a particular assay would be considered appropriate.

However, it should be noted that while a statistical approach to evaluate quality attributes of a biosimilar product may be considered in support of the demonstration that the biosimilar product is highly similar to the reference product, the determination if a biosimilar product is highly similar to the reference product will also be based upon the totality of the evidence relevant to the assessment.

Case Studies

Therapeutic antibodies rely on two types of functionalities to achieve clinical efficacy: target-specific binding by the Fab (antigen-binding fragment) domain and immune-mediated effector functions—such as ADCC and CDC—via interaction of the Fc domain with receptors on various cell types. The Fc portion of a therapeutic antibody may therefore have an important role in its mechanism of action through its influence on either ADCC or CDC. Based on their putative mechanism of action, therapeutic antibodies can generally be classified into three categories, from which their potential for Fc functionality can be ranked (Jiang et al. 2011).

The reference product is a Class I antibody that recognizes and binds to cell-bound antigen and the Fc effector functions, ADCC and CDC, are part of the MoAs. The analytical biosimilarity assessment is accomplished through extensively characterizing the physicochemical and biological properties of reference products and the proposed biosimilar product. The case study here focuses only on biological characterization with the application of a battery of in vitro biological activity assays.

Tier 1 biological assays measure quality attributes with the highest risk ranking and evaluate clinically relevant MoAs of the product. Equivalency testing will be employed for these assays that include the following:

- ADCC
- CDC

Tier 2 biological assays assess quality attributes with lower risk ranking, and quality ranges (mean ± 3SD) will be examined. These Tier 2 biological assays include:

- Apoptosis
- Cell surface antigen binding
- C1q binding
- FcγRI binding
- FcγRIIa binding
- FcγRIIb binding
- FcγRIIIa (F/V) binding
- Neonatal Fc receptor (FcRn) binding

ADCC Assay

An ADCC assay measures the biological activity of an antibody against tumor target cells. In order to induce cell lysis of the tumor target cells, Fab and Fc regions of the antibody need to bind to cell surface antigen on tumor target cells and CD16 on NK effector cells, respectively.

An engineered CD16-expressing human NK cell line (NK92-CD16) was used as effector cells while antigen-expressing cell line was used as target cells. CytoTox-Glo™ cytotoxicity assay kit (Promega) was applied and ADCC activity was quantified by measuring the ratio of EC_{50} Ref Std/EC_{50} Test Sample × 100% in luminescence (Molecular Devices, SpectraMax®L). The relative potency was calculated via software SoftMax® Pro 5.4.5 (Molecular Devices). Figure 16.1 is representative dose response curves of ADCC for the proposed biosimilar product vs. reference product.

An equivalence test was performed applying ± 1.5SD. Results show that the reference products and biosimilar product were statistically equivalent, indicating that the ADCC activity of the proposed biosimilar product is equivalent to that of the reference products.

CDC Assay

CDC measures complement mediated cellular cytotoxicity. The reference product can mediate CDC by binding to cell surface antigen on tumor target cells and recruiting complement complex via C1q binding from serum to target cells. For CDC assay, human serum was used for C1q source and antigen-expressing tumor cells were sued as target cells. CytoTox-Glo™ cytotoxicity assay kit (Promega) was applied and CDC activity was quantified by measuring the ratio of EC_{50} Ref Std/EC_{50} Test Sample × 100% in luminescence (Molecular Devices,

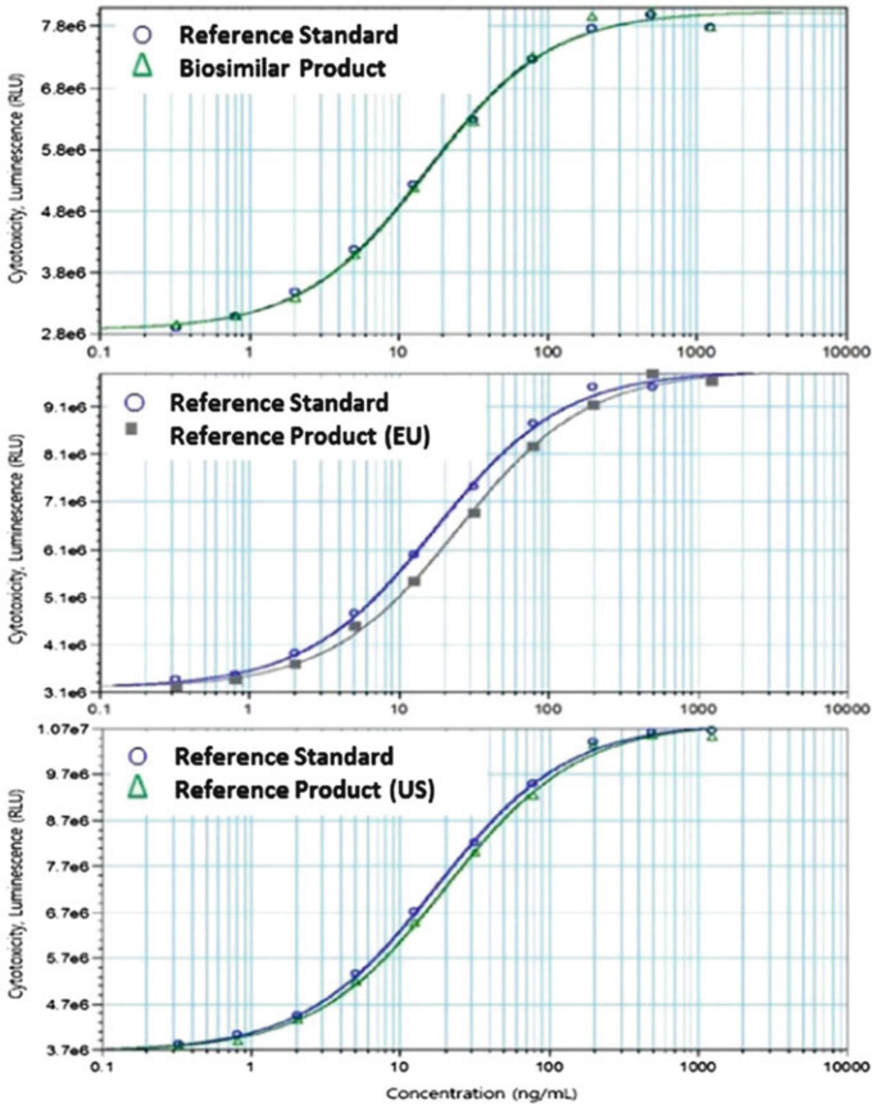


Fig. 16.1 Representative ADCC dose response curves for the proposed Biosimilar Product, Reference Product (EU) and Reference Product (US)

SpectraMax®L). The relative potency was calculated via software SoftMax® Pro 5.4.5 (Molecular Devices). Figure 16.2 is representative dose response curves of CDC for the proposed biosimilar product vs. reference product.

Equivalence test was performed applying ± 1.5SD. Results show that the reference products and biosimilar product were statistically equivalent, indicating

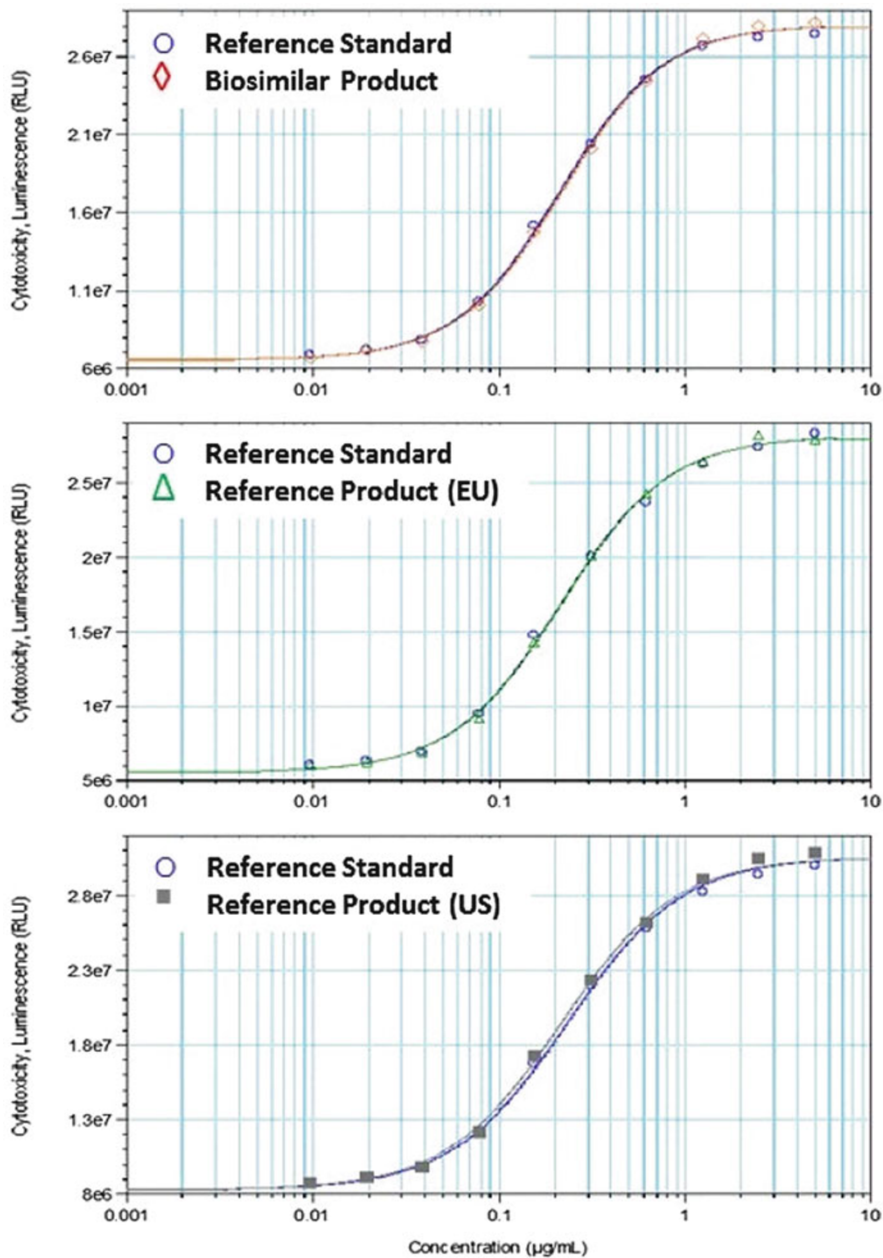


Fig. 16.2 Representative CDC dose response curves for the proposed Biosimilar Product, Reference Product (EU) and Reference Product (US)

that the CDC activity of the proposed biosimilar product is equivalent to that of the reference products.

Apoptosis Assay

Antibody induced apoptosis was determined using antigen-expressing tumor cell line and Caspase 3/7-Glo™ assay kit (Promega). Apoptosis activity was quantified by measuring the ratio of EC_{50} Ref Std/EC_{50} Test Sample $\times$ 100% in luminescence (Molecular Devices, SpectraMax®L). The relative potency was calculated via software SoftMax® Pro 5.4.5 (Molecular Devices). Figure 16.3 is representative dose response curves of Apoptosis for the proposed biosimilar product vs. reference product.

The biosimilarity range for apoptosis assay was set to be 94.9–105.3% which represents the mean $\pm$ 3SD of data from 28 different batches of reference products. All results (94.9–99.2%) of proposed biosimilar batches were within the biosimilarity range of reference products.

Cell Surface Antigen Binding Assay

Binding activities of the proposed biosimilar product and reference products to cell surface antigen were determined by a flow cytometric method. The relative binding activity was calculated by comparing to a Reference Standard by measuring the ratio of EC_{50} Ref Std/EC_{50} Test Sample $\times$ 100% in fluorescence (Beckman Dickinson). The relative potency was calculated via software SoftMax® Pro 5.4.5 (Molecular Devices). Figure 16.4 is representative dose response curves of cell surface antigen binding for the proposed biosimilar product vs. reference product.

The biosimilarity range for antigen binding assay was set to be 94.2–111.9% which represents the mean $\pm$ 3SD of data from 25 different batches of reference products. The range of the antigen binding activities of the proposed biosimilar were 92.6–100.6% whereas reference products (EU) and (US) were 97.9–109.4%.

From method qualification of the cell surface antigen binding assay, intermediate precision was 8.1%. Thus, 1.7% difference between lower limit of biosimilarity range and the lowest value of proposed biosimilar (94.2% and 92.6% each) might be due to assay variability. In addition, ADCC, CDC and apoptosis activities of proposed biosimilar product, which represent the MoAs of the reference products, were equivalent or similar to the reference products (see above). Furthermore, cell surface antigen binding activity of all GMP batches of the proposed biosimilar were close to 100% (100.0–100.6%).

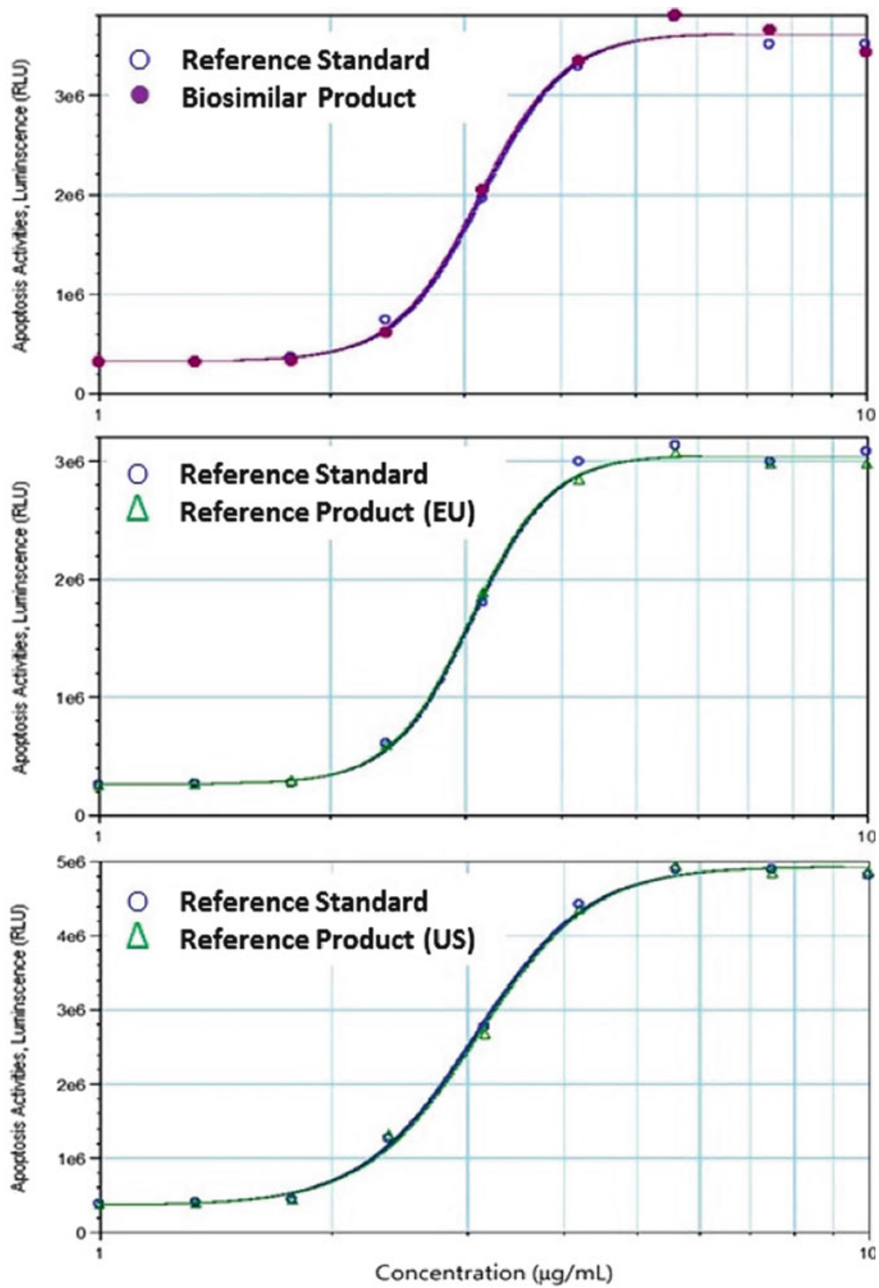


Fig. 16.3 Representative Apoptosis dose response curves for the proposed Biosimilar Product, Reference Product (EU) and Reference Product (US)

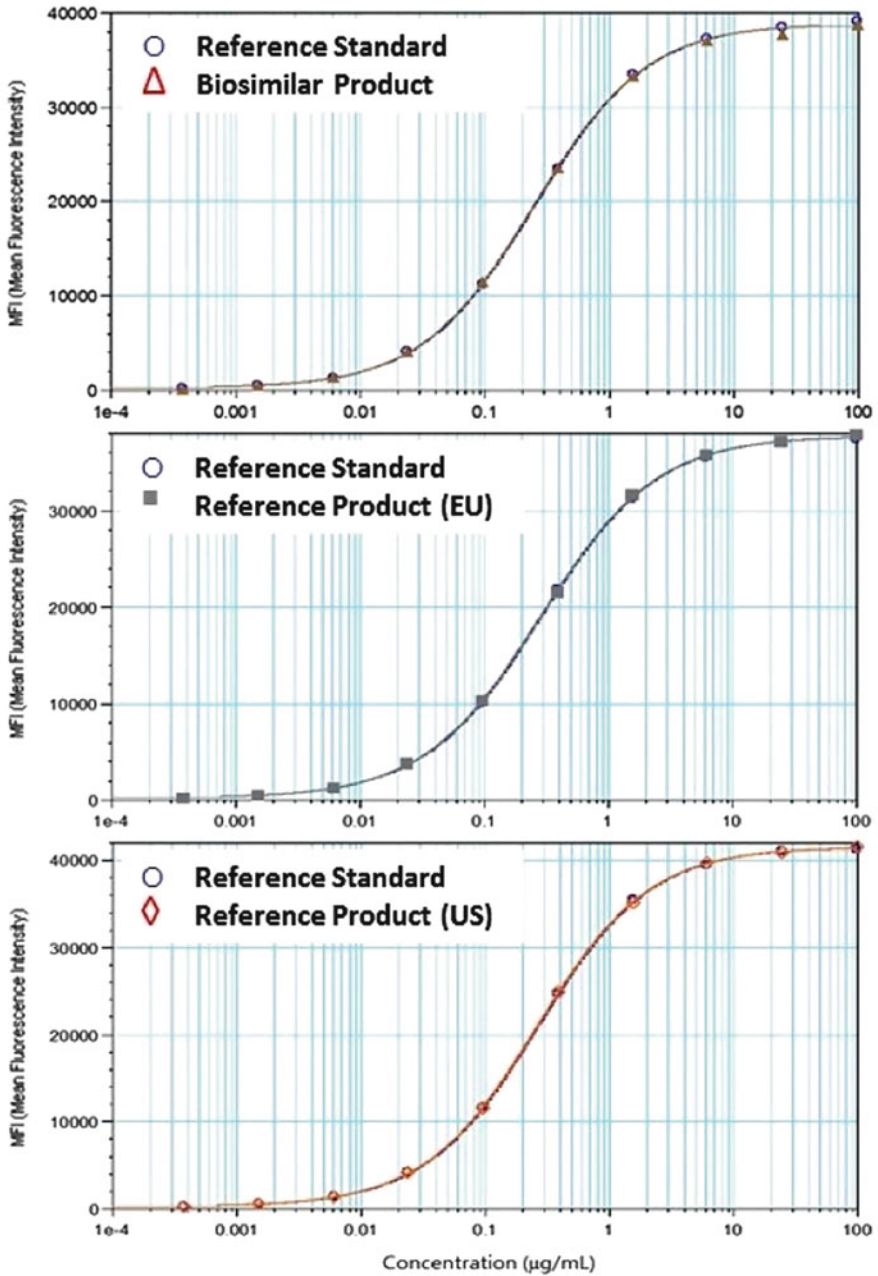


Fig. 16.4 Representative cell surface antigen binding dose response curves for the proposed Biosimilar Product, Reference Product (EU) and Reference Product (US)

C1q Binding Assay

C1q binding plays an important role in CDC function. The binding activity was measured by ELISA using a microplate reader (Molecular Devices). The relative binding activity was calculated by comparing to a Reference Standard. Figure 16.5 is representative dose response curves of C1q binding for the proposed biosimilar product vs. reference product. The biosimilarity range of reference products was 86.1–105.4% and all data of the proposed biosimilar product were within the biosimilarity range.

FcγR I Binding Assay

As the antibody binds to the immobilized FcγR I on the chip surface, the accumulation of protein results in a change of the refractive index. To generate binding curves, each test sample was serially diluted from 4000 to 125 nM. The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

The biosimilarity range for the reference products was 80.3–118.9%. The proposed biosimilar product showed a range of 95.7–113.9%, well within the biosimilarity range of the reference products.

FcγR IIa Binding Assay

The FcγR IIa binding activity was determined using Biacore™ T200 (GE Healthcare). The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

The biosimilarity range for the reference products was 90.2–107.9%. The proposed biosimilar product showed a range of 94.7–107.0%, well within the biosimilarity range of the reference products.

FcγR IIb Binding Assay

The FcγR IIb binding activity was determined using Biacore™ T200 (GE Healthcare). To generate a binding curve, each test sample was serially diluted from 20,000 to 625 nM. The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

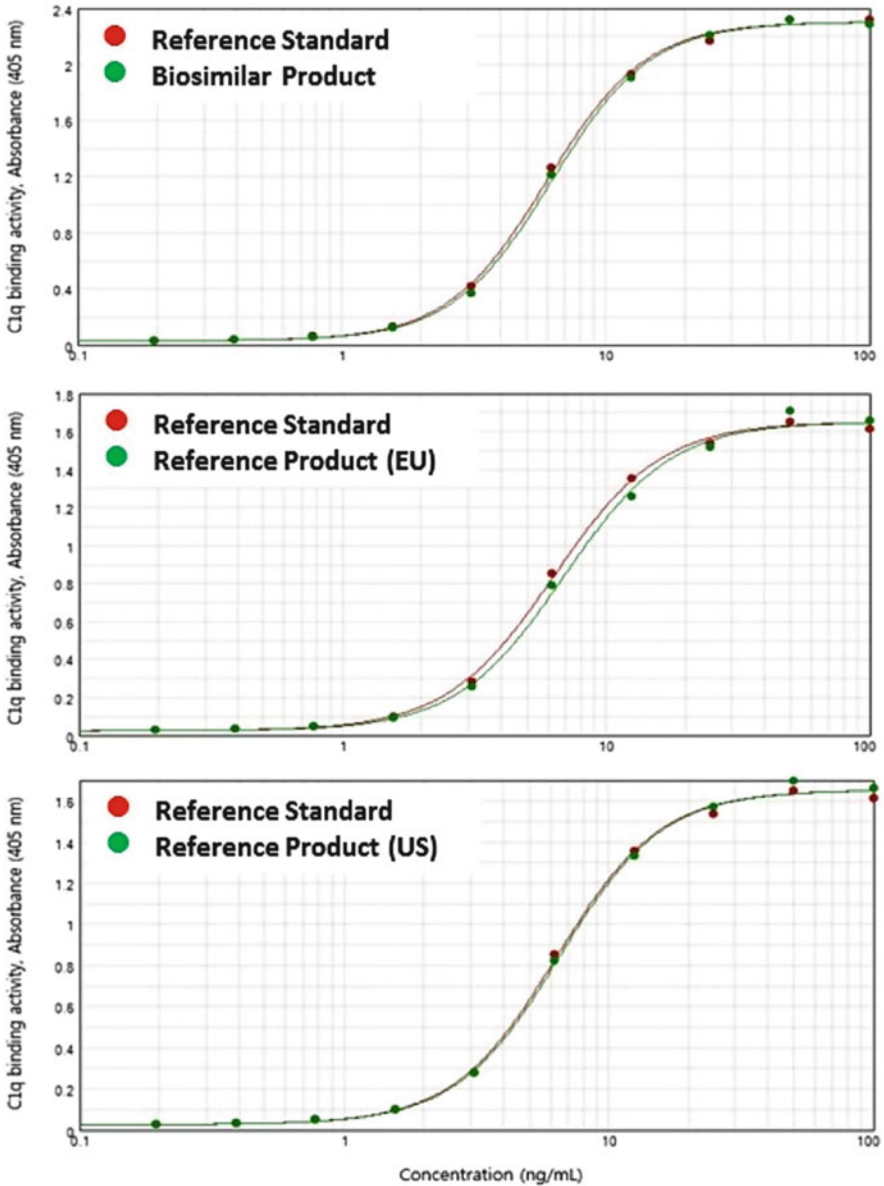


Fig. 16.5 Representative C1q binding dose response curves for the proposed Biosimilar Product, Reference Product (EU) and Reference Product (US)

The biosimilarity range for the reference products was 90.7–104.6%. The proposed biosimilar product showed a range of 96.1–101.3%, well within the biosimilarity range of the reference products.

FcγR IIIa-Phe Binding Assay

The FcγRIIIa-Phe binding activity was determined using Biacore™ T200 (GE Healthcare). To generate a binding curve, each test sample was serially diluted from 4000 to 125 nM. The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

The biosimilarity range for the reference products was 79.3–114.7%. The proposed biosimilar product showed a range of 94.9–123.3%. Three batches of the proposed biosimilar product had a slightly higher binding affinity than that of reference products. However overall range of the proposed biosimilar product was overlapped with that of reference products, and the observed difference is unlikely to affect the biological functions of the product, as ADCC, CDC and apoptosis activities of the proposed biosimilar product are equivalent or similar to that of reference products.

FcγR IIIa-Val Binding Assay

The FcγR IIIa-Val binding activity was determined using Biacore™ T200 (GE Healthcare). To generate a binding curve, each test sample was serially diluted from 2000 to 62.5 nM. The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

The biosimilarity range for the reference products was 77.5–111.6%. The proposed biosimilar product showed a range of 88.7–114.0%. Two batches of the proposed biosimilar product had a slightly higher binding affinity than that of reference products. However overall range of the proposed biosimilar product was overlapped with that of reference products, and the observed difference is unlikely to affect the biological functions of the product, as ADCC, CDC and apoptosis activities of the proposed biosimilar product are equivalent or similar to that of reference products.

FcRn Binding Assay

The FcRn binding activity was determined using Biacore™ T200 (GE Healthcare). To generate a binding curve, each test sample was serially diluted from 4000 to 125 nM. The binding affinity is evaluated using BIA evaluation software (GE Healthcare). The relative binding activity was calculated by comparing to a Reference Standard.

The biosimilarity range for the reference products was 78.6–113.2%. The proposed biosimilar product showed a range of 99.4–132.3%. However, FcRn binding activity was considered for Tier 3 analysis. In addition, clinical studies have already demonstrated PK equivalence between proposed biosimilar product and the reference products. Therefore, considering overlapped range and method intrinsic variation, it is concluded that the proposed biosimilar product was similar to the reference products in FcRn binding.

Summary of the Case Studies

A series of Fab-related biological assays (Apoptosis and antigen binding assays), as well as Fc-related biological assays (ADCC, CDC, C1q binding, Fcγ RI, Fcγ RIIa, Fcγ RIIb, FcγR IIIa-Phe and FcγR IIIa-Val binding) and the FcRn binding assays were all performed in order to assess the biosimilarity of functional properties between the proposed biosimilar product and the reference products. All results except antigen binding were within the equivalence margin or biosimilarity range of the reference products. Even though the antigen binding activity of the biosimilar product was slightly lower than that of reference products, major functional bioactivities such as ADCC, CDC and apoptosis of the biosimilar product did not show any significant difference from that of the reference products. Overall, the biosimilar product is highly similar to the reference products in biological functions.

Concluding Remarks

The development of monoclonal antibodies has been an ongoing process for many years. The rapid rate at which they are produced to clinical grade requires reassurance obtained through the rigorous testing applied to such biological products. Bioassays and assays based on physicochemical principles address different aspects of the characteristics of biologicals. The data produced by these different types of procedures complement each other to provide a spectrum of information on the substance and different batches of product. Although some of this may overlap, the different assay types provide data that relate to different properties of the molecule in question. Advances in physicochemical and biological analytical sciences enable protein products to be characterized extensively in their physicochemical and biological properties. These analytical procedures have improved the ability to identify and characterize not only the desired product but also product-related substances and product- and process-related impurities. Advances in manufacturing science and production methods, as well as advances in analytical sciences, may enhance the likelihood that a proposed product can be demonstrated to be highly similar to a reference product by better targeting the reference product's physiochemical and functional properties.

A relatively common misconception is that bioassays are so variable and imprecise that the results obtained are not usable for quantitative purposes and thus it is argued that they serve limited purpose. Although this might apply if a bad choice of assay is made, it can be avoided by careful selection of bioassay methodology, format and analysis. Therefore, carefully designed, validated and correctly analyzed bioassays can provide suitably quantitative information. In fact, bioassays have been able to detect differences in activity of biosimilar monoclonal antibodies that could not be readily predicted by physicochemical testing alone, especially as it relates to Fc functionality. Multiple functional assays should, in general, be performed as part of the analytical biosimilarity assessments. The assessment of functional activity is also useful in providing an estimate of the specific activity of a product as an indicator of manufacturing process consistency, as well as product purity, potency, and stability.

References

Christl L. Overview of regulatory pathway and FDA's guidance for development and approval of biosimilar products in US. Presented at FDA ODAC meeting, 7 Jan 2015, Silver Spring, Maryland. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM436387.pdf.

European Directorate for the Quality of Medicines. Chapter 5.3. Statistical analysis. *European pharmacopoeia*. Strasbourg: European Directorate for the Quality of Medicines; 2004. p. 473–507.

Food and Drug Administration (FDA). Scientific considerations in demonstrating biosimilarity to a reference product, 2012.

Hu J, Wala I, Han H, Nagatani J, Barger T, Civoli F, Kaliyaperumal A, Zhuang Y, Gupta S. Comparison of cell-based and non-cell-based assay platforms for the detection of clinically relevant anti-drug neutralizing antibodies for immunogenicity assessment of therapeutic proteins. J Immunol Methods. 2015;419:1–8.

Jiang XR, Song A, Bergelson S, Arroll T, Parekh B, May K, Chung S, Strouse R, Mire-Sluis A, Schenerman M. Advances in the assessment and control of the effector functions of therapeutic antibodies. Nature Rev Drug Discov. 2011;10(2):101–11.

Korhonen R, Moilanen E. Abatacept. A novel CD80/86-CD28 T cell co-stimulation modulator, in the treatment of rheumatoid arthritis. Basic Clin Pharmacol Toxicol. 2009;104:276–84.

Mire-Sluis AR. Progress in the use of biological assays during the development of biotechnology products. Pharm Res. 2001;18(9):1239–46.

Mire-Sluis AR, Thorpe R. Quantitative cell line based bioassays for cytokines. J Immunol Methods. 1995;187:191–9.

Mire-Sluis AR, Gaines-Das R, Gerrard T, Padilla A, Thorpe R. Biological assays: their role in the development and quality control of biological medicinal products. Biologicals. 1996; 24:351–61.

Peppel K, Crawford D, Beutler B. A tumor necrosis factor (TNF) receptor-IgG heavy chain chimeric protein as a bivalent antagonist of TNF activity. J Exp Med. 1991;174:1483–9.

Sasardic D, Mire-Sluis AR. Biological standardization and control at the World Health Organization: successes, current issues and future challenges. Biologicals. 2000;28:47–53.

Schnueriger A, Grau R, Sondermann P, Schreitmueller T, Marti S, Zocher M. Development of a quantitative, cell-line based assay to measure ADCC activity mediated by therapeutic antibodies. Mol Immunol. 2011;48(12–13):1512–7.

Sievers EL, Senter PD. Antibody-drug conjugates in cancer therapy. Annu Rev Med. 2013;64: 15–29.

Thorpe R, Wadhwa M, Mire-Sluis AR. The use of bioassays for the characterisation and control of biological therapeutic products produced by biotechnology. Dev Biol Stand. 1997;91:79–88.

Tsong Y, Dong X, Shen M. Development of statistical methods for analytical similarity assessment. J Biopharm Stat. 2017;27(2):197–205.

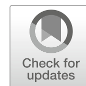
USP. Chapter <1032>. United States Pharmacopeia (USP): the national formulary. Rockville: USP Pharmacopeia Convention, Inc.; 2012a.

USP. Chapter <1034>. United States Pharmacopeia (USP): the national formulary. Rockville: USP Pharmacopeia Convention, Inc.; 2012b.

Wu YL, Li J, Kim HJ, Liu X, Liu WY, Akhgar A, Bowen MA, Spitz S, Jiang XR, Roskos L, White W. A novel antibody dependent cellular cytotoxicity (ADCC)-based reporter neutralizing antibody (NAb) assay in support of benralizumab clinical development. AAPS J. 2015;17(6):1417–26.

Yang H, Kim HJ, Zhang L, Schenerman M, Strouse R, Jiang XR. Implementation of parallelism testing for four-parameter logistic model in bioassays. PDA J Pharm Sci Tech. 2012;66:262–9.

Chapter 17
Statistical Considerations for Demonstration of Analytical Similarity



Harry Yang, Richard K. Burdick, Aili Cheng, and Richard O. Montes

Abstract Analytical similarity is the foundation for demonstration of biosimilarity between a proposed biosimilar product and a reference product. For this assessment, the FDA has recommended a tiered system in which quality attributes are categorized into three tiers commensurate with their risk ranking. Different approaches of varying rigor have been recommended to analyze the tiered quality attributes. Two of these approaches are the equivalence test of means, and a quality range approach that requires individual biosimilar lot values to fall in a range based on the reference product lots. However, lack of knowledge of the reference product such as target specifications, process changes, and sources of bulk materials used to produce the final product lots makes it extremely challenging to set the acceptance criteria for both of these approaches. Further confounding the issue is that there is limited published literature on the subject and the FDA draft guidance published in September 2017 was withdrawn in June 2018. In this chapter, we provide an in-depth discussion of practical issues concerning analytical similarity and statistical remedies. Focus is on (1) Statistical criteria for equivalence of means testing and quality ranges, (2) Sample size considerations; and (3) Statistical strategies to mitigate risk of correlation among the reference products lots. Finally, a list of goals is provided to be considered when developing criteria to demonstrate analytical similarity.

H. Yang (✉)
Statistical Sciences, MedImmune, LLC, Gaithersburg, MD, USA
e-mail: yangh@medimmune.com

R. K. Burdick
Burdick Statistical Consulting, Longmont, CO, USA
e-mail: RickBASU@aol.com

A. Cheng
Pharmaceutical Sciences and Manufacturing Statistics, Pfizer, Andover, MA, USA
e-mail: aili.cheng@pfizer.com

R. O. Montes
Biosimilars Pharmaceutical Sciences, Pfizer, Lake Forest, IL, USA
e-mail: richard.montes@pfizer.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_17

Keywords Analytical similarity · Biosimilars · Correlation impacts · Effect size · Equivalence testing · Quality range · Risk mitigation

Introduction

The advance in regulatory guidance on biosimilars, including the passage of the Biologics Price Competition and Innovation Act of 2009, has created abbreviated licensure pathways for biological products shown to be biosimilar or interchangeable with a reference product. Common to these regulator guidelines is the stepwise approach which starts with the assessment of critical quality attributes (CQA) used to characterize the biosimilar products. Furthermore, the FDA recommends a tiered system in which quality attributes are categorized into three tiers commensurate with their risk ranking. Different statistical approaches of varying rigor are then used to analyze the tiered quality attributes. These analytical assessments, taken together with PK/PD data and clinical efficacy and safety evaluations, provide the totality of evidence in support of licensure of the product as a biosimilar to the reference product. In this chapter, we discuss several challenging issues concerning the demonstration of analytical similarity. Also covered are statistical alternatives to methods discussed by the regulatory agencies.

Background

Biotechnology-derived drug products hold great promise for treating various diseases such as cancer and inflammatory diseases. Examples of these products include monoclonal antibodies, proteins, and peptides. In recent years, the expiration of originator biologic drug patent protection has spawned the development of alternative versions of the products known as biosimilars. Biosimilars are becoming one of the fast-growing sectors and investments in these products are at an all-time high. The establishment of regulatory licensure pathways for biosimilars in various regions provides added motivation for biosimilar development. A common thread of these pathways is demonstration of analytical similarity. However, because of their large size, complex structure, and complicated manufacturing process, such demonstration poses a host of unique scientific and statistical challenges.

When compared to small molecule drugs, biological products have much more complex structures. For example, aspirin has 21 atoms with a molecule weight of approximately 18 Da. By contrast, a typical IgG monoclonal antibody (mAb) has about 20,000 atoms and molecule weight of approximately 150 KD (or 150,000 Da). A generic or identical copy of a small molecule drug can be made through chemical synthesis, based on the known structure of the originator drug. However, it is nearly impossible to create an exact version of a biological drug product for the following reasons:

1. Biologic products are produced in living cells and prone to additional physico-chemical structural changes during production. As a result, a biologic drug is a mixture of various isoforms, many or all of which have biological activities.
2. Due to technological advances, the manufacturing process of the biosimilar is likely different from that of the originator drug in several aspects, including cell lines and formulation. These differences make it even more challenging to produce an exact mixture of proteins.
3. Manufacturing processes are proprietary, and thus the sponsor of a biosimilar has limited knowledge of the originator production processes.

The challenges noted above result in limited availability and historical knowledge of reference product lots. As noted by Berkowitz (2017), the greatest source of variability of bioprocessing comes from the cell culture process because of the complex and sensitive linkage of cell growth to its physical and chemical environment inside a bioreactor. Berkowitz also notes the average measures of CQAs may be biased due to repeated measurements of the same lots.

For biologic manufacturing, there is a constant drive for improving product quality and production efficiency, through adopting new technology, utilizing the same technological platforms, and sourcing materials from vendors of higher quality standards. Thus, changes in the originator product process force the biosimilar sponsor to match a dynamic as opposed to a static process. To compound the issue further, age differences between the sourced reference product (RP) lots and the biosimilar test product (TP) lots introduce additional variation in the comparison. It is quite possible that one may succeed in demonstrating similarity between the TP and RP based on one set of RP lots, but fail the demonstration when a different set of RP lots is used in the analysis.

Acquisition of lots by the biosimilar sponsor is also problematic. Ideally, a valid comparison between the TP and the RP can only be made based on random lots produced from different bioreactor runs (or drug substance lots). Since RP lots used for a biosimilar development are usually procured within a fixed time period, some or all of the lots may be sourced from the same bulk drug substance lots. As a result, they do not reflect the full range of manufacturing variations. This situation also creates correlation among the drug product lots, which complicates the statistical analysis of the data. Although correlation may be potentially mitigated through appropriate statistical analysis (see Section Mitigation of Correlation Effects), the knowledge of the historical records of the RP lots is usually lacking, making such statistical adjustments difficult.

Stepwise Approach

Although the requirements for demonstrating biosimilarity vary from region to region, there is a common theme in these requirements that has led to adoption of a stepwise approach to data generation and the evaluation of residual uncertainty

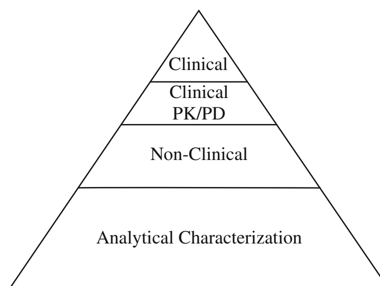


Fig. 17.1 A stepwise approach to demonstration of biosimilarity

(EMA 2014; FDA 2015). As shown in Fig. 17.1, the approach begins with assessment of analytical characterization, which is the foundation for demonstration of similarity between the proposed biosimilar and RP. The assessment includes analytical comparisons of structural and functional attributes between the biosimilar TP and RP.

If uncertainties about the safety of the proposed biosimilar arise from the structural and functional characterization, animal studies may be conducted before initiation of comparative clinical studies, which include assessment of immunogenicity and pharmacokinetics or pharmacodynamics to establish similarity in safety and efficacy. The extent to which the clinical studies are carried out needs to be scientifically justified. Overall demonstration of analytical similarity is the foundation of obtaining biosimilar designation of the proposed biosimilar. The stepwise approach provides a process to determine the nature and scope of each subsequent study. The approval of biosimilar applications is based on the totality of the evidence submitted in the application.

Analytical Similarity Assessment

Consistent with the recent regulatory initiatives of adopting risk-based and lifecycle principles in product and process development (FDA 2004; ICH 2005 (Q9), 2009a (Q8), b (Q10), 2012 (Q11)), the (FDA 2017) recommends a tiered approach for analytical similarity assessment. As the first step, all CQAs are assessed for risk and categorized into Tiers 1, 2, or 3, based on their impact on the safety, efficacy, and quality of the proposed biosimilar product. Table 17.1 provides definitions of the tier designations.

Tools such as failure mode and effects analysis (FMEA) can be used for the risk analysis. For example, one may quantify the impact of the CQAs on activity, PK/PD, safety, and immunogenicity through a risk score which is based on both severity and probability of occurrence. Figure 17.2 shows an example of such a scoring system.

The overall risk score is defined as the product of severity and occurrence. A CQA is judged to be of high, moderate, or low risk based on the criterion listed in Table 17.2. However, per the FDA guidance (FDA 2017), other considerations

Table 17.1 A summary of the risk assessment tiers

CQA	Definition
Tier 1	High impact on activity, PK/PD, safety, and immunogenicity
Tier 2	Moderate impact on activity, PK/PD, safety, and immunogenicity
Tier 3	Low impact on activity, PK/PD, safety, and immunogenicity, and attributes not amenable to statistical analysis

Severity						
	5	5	10	15	20	25
	4	4	8	12	16	20
	3	3	6	9	12	15
	2	2	4	6	8	10
	1	1	2	3	4	5
	0	1	2	3	4	5
		Occurrence				

Fig. 17.2 A scoring system based on severity and occurrence

Table 17.2 Definition of high, moderate, low impact

Risk	Criterion
High	Risk score $\geq$10
Moderate	4$\leq$ scores $\leq$9
Low	Risk score $\leq$3

such as level of attributes, assays used for assessing the attribute, and types of the attributes/assays need to be provided when determining the appropriate tier.

Table 17.3 presents a list of CQAs and methods for characterization of the TP, GP2015 (Erelzi®), a proposed biosimilar to Enbrel (etanercept) by Sandoz (FDA 2016).

Based on a risk analysis, the CQAs TNF-α binding and neutralization, were deemed to be Tier 1 CQAs.

EMA Statistical Approach

The European Medicines Agency (EMA) wrote a reflection paper discussing statistical methodology for comparative assessment of quality attributes (EMA 2017). As a reflection paper, it is not intended to offer practical guidance to drug developers, but rather to provide high level discussion of concepts and a blueprint for development of such guidance.

The current version of the paper emphasizes the strengths and limitations of several statistical approaches that have been applied in related comparability studies. The paper makes it very clear that descriptive statistics are generally not enough to support claims of comparability or similarity. In addition, the paper describes issues that should be considered by sponsors as they plan an appropriate statistical assessment of their data. Practical guidance to industry based on feedback from the reflection paper is the ultimate goal of this process.

Table 17.3 Critical quality attributes and their methods for GP2015 (Erelzi®), a biosimilar to Enbrel

Quality attribute	Methods
Primary structure	• Reduced peptide mapping with ultraviolet (UV) and mass spectrometry (MS) detection
	• Mass analysis of peptides (ESI-MS)
	• Amino acid analysis
	• Intact mass (MALDI-TOF-MS)
	• Peptide mapping coupled with tandem mass spectrometry (MS/MS)
	• Disulfide bridging
Protein content	• UV/Vis spectroscopy
Higher order structure	• Far and near UV circular dichroism
	• Differential scanning calorimetry
	• Hydrogen/deuterium exchange
	• FTIR
	• 1D-NMR
	• X-ray crystallography
High molecular weight species/aggregates	• Size exclusion chromatography (SEC-HPLC)
	• Size exclusion chromatography (SEC-MALLS)
	• Analytical ultracentrifugation
	• FFF-MALLS
	• 2D-DIGE (charge and size)
Fragments	• CE-SDS
	• SEC
Charge and hydrophobic variants	• CZE
	• Reversed phase chromatography (RPC-HPLC)
Glycosylation and glycosylation site occupancy	• *Np* = P-HPLC-MD (N-glycans–overall, TNRF portion and Fc portion)
	• MALDI-TOF (O-linked glycan analysis)
	• AEX, WAX and RP-HPLC of labeled N or N and O glycans (Sialic acid analysis)
	• Boronate affinity chromatography (glycation)
In vitro potency assays	• TNF-α neutralization assay reporter gene assay
	• TNF-β neutralization assay reporter gene assay
	• Cell based apoptosis inhibition assay
Binding assay–TNF-α	• Surface plasmon resonance

(continued)

Table 17.3 (continued)

Quality attribute	Methods
Binding assay–Fc and complement	• FcγRIIIa V and F type binding affinity (SPR)
	• FcγRI binding (SPR)
	• FcγRIIa binding (SPR)
	• FcγRIIIa binding affinity (SPR)
	• FcRn binding affinity (SPR)
	• C1q binding assay (ELISA)
Bioassay/mechanism of action exploration	• ADCC (NK cell line as effectors and engineered target cell expressing high levels of mTNF)
	• CDC (target cell stably transfected with a constitutively membrane-associated form of TNF-α)

A workshop was held in May 2018 titled, "Workshop on the reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development". Presentations of that workshop are available on EMA's dedicated website: http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/events/2017/09/event_detail_001507.jsp&mid=WC0b01ac058004d5c3.

FDA Statistical Approach

The FDA published a draft guidance on the statistical approaches for analytical similarity in September 2017 (FDA 2017). In response to industry concerns, the FDA withdrew the draft guidance in June of 2018. FDA indicates that it plans to reissue the draft guidance after considering changes to the document that will "promote a more efficient pathway for the development of biosimilar products".

Although withdrawn, it is useful to study this document because it represents the basic structure of the analytical similarity, and some approaches to the problem. It is recommended for sponsors to develop an analytical similarity assessment plan and discuss it with the Agency. The analytical similarity plan includes risk ranking of the quality attributes, the determination of the statistical methods for different tiers of quality attributes based on the risk ranking, the statistical analysis plan, and finalization of the analytical similarity assessment plan.

The draft guidance recommends that the quality attributes be categorized into three tiers according to their risk ranking with respect to their potential impact on activity, PK/PD, safety, or immunogenicity. Different statistical assessment methods are recommended for CQAs in different tiers

1. Equivalence testing for Tier 1,
2. A quality range approach for Tier 2, and
3. Side-by-side graphical comparison for Tier 3.

The equivalence test is the most rigorous of the statistical assessments, and is the only approach that protects type I error or patient's risk regardless of sample size. The quality range approach is based on determining the number of biosimilar TP lots that fall within an expected range based on the RP. For Tier 3, examination of the data using graphical comparisons is considered appropriate. In the next four sections, we focus on the Tier 1 and Tier 2 approaches.

Equivalence Test for Tier 1 CQAs

Per the FDA's recommendation, similarity of a Tier 1 CQA is established by testing the following hypotheses

$$\begin{aligned} &H_0 : \mu_T - \mu_R \leq -\delta \text{ or } \mu_T - \mu_R \geq \delta \\ &H_A : \quad -\delta < \mu_T - \mu_R < \delta \end{aligned} \tag{17.1}$$

where μ_T and μ_R are the means of TP and RP, respectively, and δ is the equivalence margin. The equivalence margin is a difference considered to have no clinical impact. If this information is not available, FDA has recommended in its guidance document to use $\delta = 1.5 \times \sigma_R$, where σ_R represents the standard deviation of the RP. Since σ_R is unknown, it is estimated using sampled RP lots sourced from the market. FDA Approach to Define the Equivalence Margin for Tier 1 CQAs provides a discussion of the rationale used by the FDA in selecting this equivalence margin.

Testing of the hypotheses (17.1) is usually carried out using two-one-sided t-tests (TOST). Based on the union-intersection principle, if each individual test is conducted with a type I error of α, the overall type I error remains α since the null hypothesis must be rejected in both cases in order to demonstrate equivalence (Berger and Hsu 1996). The test procedure was originally proposed by Schuirmann (1987) and is operationally equivalent to assessing if the $100(1 - 2\alpha)\%$ confidence interval of the difference $\mu_T - \mu_R$ is entirely contained within the range from $-\delta$ to $+\delta$.

Let $Y_{iT}(i = 1, \ldots, n_T)$ and $Y_{iR}(i = 1, \ldots, n_R)$ be the responses of TP and RP, respectively, from random samples of n_T test lots and n_R reference lots (not necessarily equal). It is assumed that Y_{iT} are independently and identically distributed (i.i.d.) normal random variables with mean μ_T and variance σ_T^2, and Y_{iR} are i.i.d. normal random variables with mean μ_R and variance σ_R^2. As will be described later, the assumption of independence is often not satisfied when drug product lots are sourced from the same drug substance lots.

Let $\overline{Y}_T, \overline{Y}_R, S_T^2$, and S_R^2 represent sample means and variances of the reference and test products, respectively, calculated using the following formulas for $j = T$ or R.

$$\overline{Y}_j = \frac{\sum_{i=1}^{n_j} Y_{ij}}{n_j} \tag{17.2}$$

$$S_j^2 = \frac{\sum_{i=1}^{n_j} \left(Y_{ij} - \overline{Y}_j\right)^2}{n_j - 1}. \tag{17.3}$$

Under the equal variance assumption $\sigma_T^2 = \sigma_R^2$, the TOST uses the following two test statistics to test the hypotheses (17.1)

$$T_L = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) + \delta}{\sqrt{S_P^2 \left(1/n_T + 1/n_R\right)}} \tag{17.4}$$

$$T_U = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \delta}{\sqrt{S_P^2 \left(1/n_T + 1/n_R\right)}} \tag{17.5}$$

where S_P^2 is the pooled sample variance given by

$$S_P^2 = \frac{\left(n_T - 1\right) S_T^2 + \left(n_R - 1\right) S_R^2}{n_T + n_R - 2}. \tag{17.6}$$

The null hypothesis in (17.1) is rejected at a significance level of α if

$$T_L > t_{1-\alpha:n_T+n_R-2} \text{ and } T_U < -t_{1-\alpha:n_T+n_R-2} \tag{17.7}$$

where $t_{1-\alpha:n_T+n_R-2}$ is the t-distribution quantile with degrees of freedom $n_T + n_R - 2$ and area $1 - \alpha$ to the left.

Equivalently, one may perform the test of hypotheses in (17.1) by computing the two-sided $100(1 - 2\alpha)\%$ confidence interval of the mean difference $\mu_T - \mu_R$

$$\overline{Y}_T - \overline{Y}_R \pm t_{1-\alpha:n_R+n_T-2}\sqrt{S_P^2 \left(1/n_T + 1/n_R\right)}. \tag{17.8}$$

If the confidence interval falls in the range from $-\delta$ to δ, then the null hypothesis in (17.1) is rejected and equivalence is demonstrated.

If equal variances cannot be assumed, the appropriate $100(1 - 2\alpha)\%$ confidence interval is

$$\overline{Y}_T - \overline{Y}_R \pm t_{1-\alpha:\upsilon}\sqrt{S_T^2/n_T + S_R^2/n_R} \tag{17.9}$$

where

$$\nu = \frac{\left(S_T^2/n_T + S_R^2/n_R\right)^2}{\frac{S_T^4}{(n_T-1)n_T^2} + \frac{S_R^4}{(n_R-1)n_R^2}} \tag{17.10}$$

is the degrees of freedom based on the Satterthwaite's approximation (Satterthwaite 1946).

Regardless of the sample size, the test guarantees that when $|\mu_T - \mu_R| = \delta$, the probability of rejecting H_0 is α. This represents the patient's risk. As n_T and n_R increase, the confidence intervals in (17.8) and (17.9) will shorten, providing a greater chance of passing the test of equivalence. Therefore, the equivalence test ensures that a claim of equivalence (resulting from rejecting the null hypothesis) is not the consequence of small sample sizes or large variability.

An Example of a Tier 1 Test

The Tier 1 equivalence test is demonstrated using the CQA TNF-α RGA binding method data shown in Table 17.4 (Table 17.8 (FDA 2016)).

For this demonstration, the equivalence margin δ is chosen to be $1.5 \times S_R = 1.5 \times 7.47 = 11.2$.Unequal variances were assumed, and so from (17.10) the degrees of freedom are calculated to be

$$\begin{aligned} \nu &= \frac{\left(S_T^2/n_T+S_R^2/n_R\right)^2}{\frac{S_T^4}{(n_T-1)n_T^2}+\frac{S_R^4}{(n_R-1)n_R^2}} \\ \nu &= \frac{\left(2.07^2/9+7.47^2/13\right)^2}{\frac{2.07^4}{(9-1)9^2}+\frac{7.47^4}{(13-1)13^2}} = 14.5 \text{ (rounded up to 15)} \end{aligned} \tag{17.11}$$

Using (17.9) the 90% confidence interval of the difference is computed

$$\begin{aligned} &\overline{Y}_T - \overline{Y}_R \pm t_{1-\alpha:\upsilon}\sqrt{S_T^2/n_T + S_R^2/n_R} \\ &92.44 - 93.69 \pm 1.757\sqrt{2.07^2/9 + 7.47^2/13} \\ &-1.25 \pm 3.82 \\ &(-5.1 \text{ to } 2.6) \end{aligned} \tag{17.12}$$

Table 17.4 Summary statistics of TNF-α RGA binding method

Product	No. of lots	Mean	SD
GP2015 (TP)	$n_T = 9$	$\overline{Y}_T = 92.44$	$S_T = 2.07$
US Enbrel (RP)	$n_R = 13$	$\overline{Y}_R = 93.69$	$S_R = 7.47$

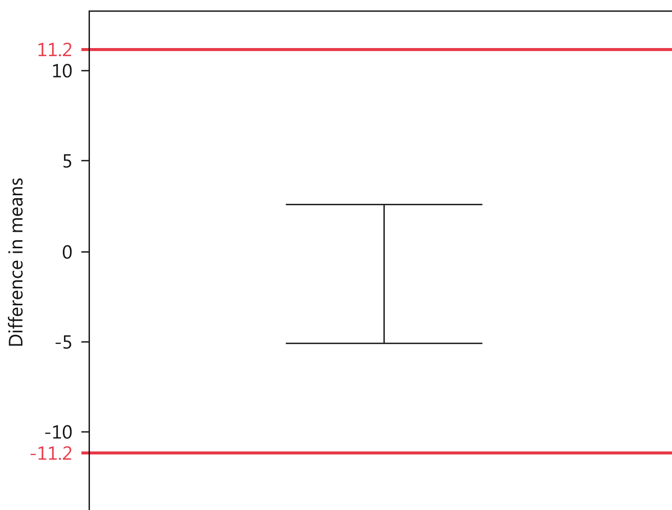


Fig. 17.3 90% confidence interval on difference in means versus equivalence margin

Since the interval falls entirely within the range from –11.2 to +11.2, equivalence has been demonstrated. Figure 17.3 displays this result visually. The vertical line represents the computed confidence interval on $\mu_T - \mu_R$ and the red horizontal lines at –11.2 and +11.2 represent the equivalence margin. The entire confidence interval is contained within this range, demonstrating equivalence of means.

Tier 2 Testing

For the comparison of Tier 2 QAs, FDA recommends an approach based on a quality range defined as

$$\mu_R \pm K \times \sigma_R \tag{17.13}$$

where K is a scale factor that needs to be appropriately justified. Similarity for a Tier 2 quality attribute is established if a sufficient percentage of TP lot values (e.g., 90%) fall within the quality range defined by (17.13). As with Tier 1 testing, the parameters μ_R and σ_R are estimated by $\overline{Y}_R$ and S_R, respectively. In the majority of Biologics License Application (BLA) filings to this point, sponsors have selected $K = 3$ and 90% as the required percentage of TP lots falling in this range (see FDA webpage under Advisory Committees > Committees Meeting Materials to find applications with this information, (FDA 2018)). Considerations for selecting K are further discussed in Selection of K for Tier 2 CQAs. Table 17.5 provides summary data for one of the Tier 2 attributes reported in Table 17.10 of FDA (2016).

Table 17.5 Summary statistics of TNF-α neutralization apoptosis method

Product	No. of lots	Mean	SD
GP2015 (TP)	$n_T = 8$	$\overline{Y}_T = 101.0$	$S_T = 5.50$
US Enbrel (RP)	$n_R = 11$	$\overline{Y}_R = 117.7$	$S_R = 10.25$

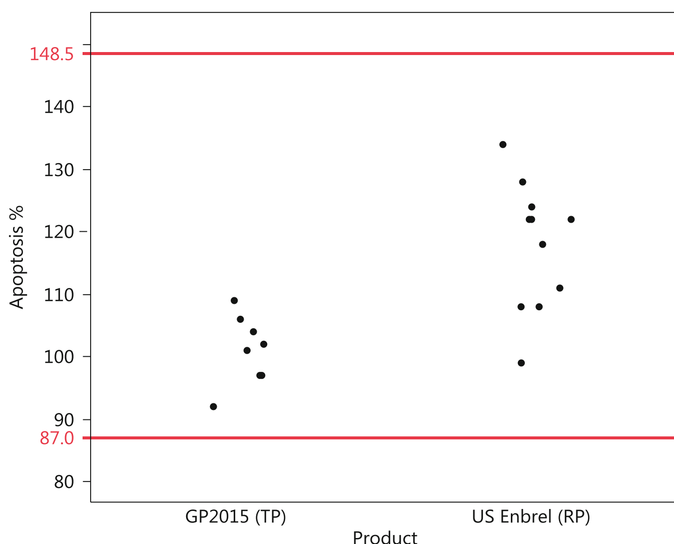


Fig. 17.4 Quality range with individual TP and RP values

The quality range using $K = 3$ with the appropriate estimators is

$$\begin{aligned} &\overline{Y}_R \pm K \times S_R \\ &117.7 \pm 3 \times 10.25 \\ &(87.0, 148.5) \end{aligned} \tag{17.14}$$

All TP lots fit within the Tier 2 quality range as shown in Fig. 17.4.

Adjustments for Age Differences

The guidance notes that age differences between RP and TP lots at the time of testing may result in analytical differences, and that a pre-specified plan is needed to address how changes in attributes over the shelf-life will be incorporated into the determination of the similarity acceptance criterion. One statistical method that can be used for this purpose is regression analysis.

Figure 17.5 presents a small representative data set for the stability indicating method SE-HPLC Main peak measured as percentage of total peak area.

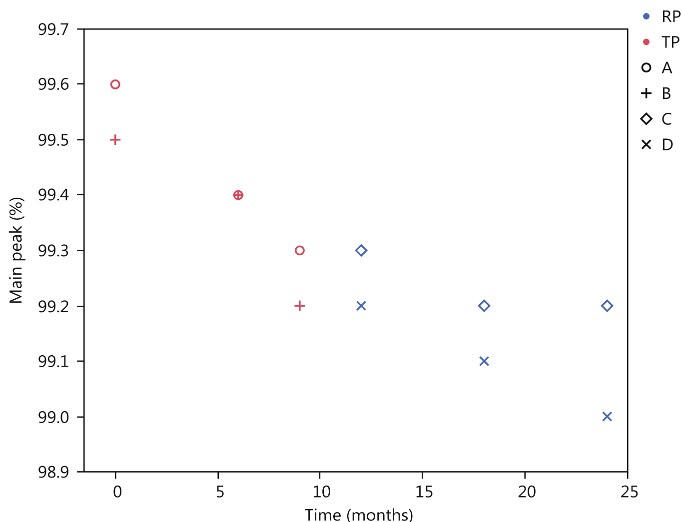


Fig. 17.5 Plot of main peak (%) versus time in months

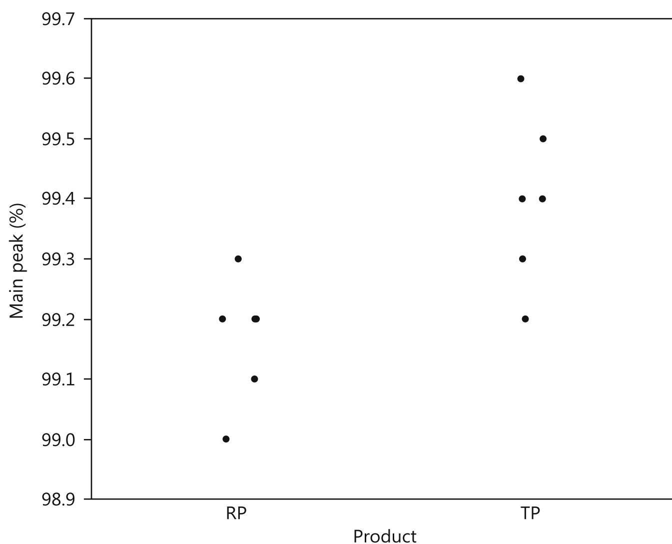


Fig. 17.6 Plot of main peak values ignoring time

There are two TP lots, A and B, and two RP lots, C and D. Note that the RP lots on the market are generally older on average than the newly manufactured biosimilar TP lots. Figure 17.6 shows a side-by-side plot of the two product groups without acknowledgement of the lot age.

Table 17.6 Estimate of the regression line

Term	Estimate	Prob > |*t*|
Intercept	99.5	<0.0001
Time (months)	– 0.02	0.0029
Product [RP]	0.012	0.8155
Random effect	Var component	Pct of total
Lot	0.0040706	47.746
Residual	0.0044549	52.254
Total	0.0085255	100.000

The TP and RP lots look quite different in Fig. 17.6, even though in Fig. 17.5 there is no apparent disconnect in the stability trend. The apparent difference in Fig. 17.6 suggests it would be difficult to demonstrate similarity for either a Tier 1 or Tier 2 analysis. In order to fairly compare the two product groups, it is first necessary to adjust for the age bias.

To do this, one can estimate the slope of the CQA over time and then project each individual TP and RP value to the expected value at the same point in time, t^*. Assuming the RP and TP profiles are parallel, the value of t^* is unimportant. It is typically selected in the middle of the data set, or perhaps at shelf life expiry or at the release since most of the biosimilar lots are tested at release. If one cannot assume equal degradation rates, one must fit a separate slope for each product, and then selection of t^* will impact the results.

Table 17.6 reports the results of a regression model of SE main peak on time, product, and lots (as a random effect) where it is assumed the slopes of RP and TP are equal. This is the "separate intercepts, common slopes" model described in ICH Q1E (ICH 2003) for the error structure, with an allowance for unequal intercepts for product (RP or TP).

The estimated slope shown in Table 17.6 is –0.02% per month. That is, the main peak is decreasing on average by 0.02% per month. Each individual value in the data set is now projected to its value at t^* using the equation

$$\text{Value at } t^* = \text{Original} + \left(t^* - t_O\right) \times Slope \tag{17.15}$$

where t_O is the time point associated with the original observation. For example, TP Lot A has value 99.6 at time $t_O = 0$. Thus, the value at $t^* = 12$ is

$$\text{Value at } t^* = 99.6 + (12 - 0) \times (-0.02) = 99.36\%. \tag{17.16}$$

Likewise, RP lot C has value 99.2 in month $t_O = 24$. The value at $t^* = 12$ for this original value is

$$\text{Value at } t^* = 99.2 + (12 - 24) \times (-0.02) = 99.44\%. \tag{17.17}$$

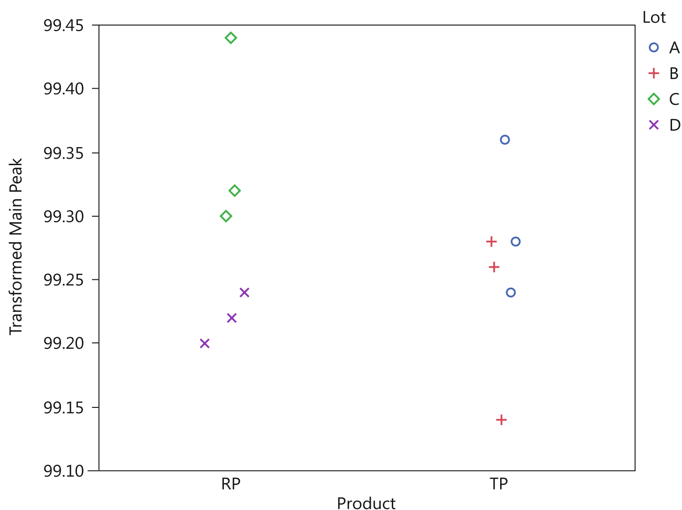


Fig. 17.7 Data values at $t^* = 12$

Figure 17.7 shows a side-by-side plot of the projected data at $t^* = 12$. Once the age bias has been removed, the two groups compare quite favorably. The values at $t^* = 12$ can now be used for either a Tier 1 or Tier 2 analysis.

Tier 3 Testing

For Tier 3 QAs, the FDA recommends carrying out similarity assessment using raw data and graphical comparisons. The method is appropriate because these QAs have the lowest impact on product safety, efficacy, and quality.

FDA Approach to Define the Equivalence Margin for Tier 1 CQAs

One of the key components of an equivalence test is the equivalence margin, δ. The equivalence margin should represent a difference of practical significance. Ideally, δ should be based on biological and scientific understanding of the CQA and its impact on product safety and efficacy. However, it is difficult to establish a causal relationship between the analytical measurement of a CQA and the product. As stated in the guidance, often a scientific argument for selecting δ cannot be made. For this reason, the FDA defined δ as a function of the variability of the RP. In

Table 17.7 Test size and power at $|\mu_T - \mu_R| = \sigma_R/8$ assuming $\delta = 1.5\sigma_R$

Number of lots per product	Power	Type I error rate (confidence level, %[a])
6	0.76	0.10 (80)
8	0.81	0.065 (87)
10	0.87	0.05 (90)
15	0.90	0.05 (90)

[a]Confidence level of two-sided interval on mean difference used in equivalence test

particular, the guidance specifies that $\delta = 1.5\sigma_R$. Selection of this value was made to satisfy three objectives:

1. Ensure that values of the attribute being tested for the proposed biosimilar generally fall within the reference product distribution,
2. Provide a unified representation of δ for all Tier 1 quality attributes despite different levels of product variability, and
3. Ensure sufficient power for sample sizes that are practical for the sponsor to produce.

Given these objectives, the FDA determined a target power for numbers of RP lots that could be sourced and TP lots that could be manufactured within a reasonable time frame. After performing many simulations, the FDA selected $\delta = 1.5\sigma_R$ to ensure reasonable powers at the true difference of means, $|\mu_T - \mu_R| = \sigma_R/8$ with several type I error rates.

As a visual representation of this definition, Fig. 17.8 depicts two standard normal distributions with equal variances where the means differ by $1.5\sigma_R$. The percentage of overlap for these two curves is 45%.

Table 17.7 reports computed powers when $\delta = 1.5\sigma_R$ and the true mean difference is $|\mu_T - \mu_R| = \sigma_R/8$. These calculations assume that σ_R is known. Note that originally, a greater type I error rate was allowed when the number of lots per product is less than 10. However, the 2017 draft guidance recommends at least 10 lots for both TP and RP.

Burdick (2015) noted that there are several advantages of the FDA approach for determining δ:

1. It provides a criterion in a situation where no scientifically based value is available.
2. The criterion can be applied across all products.
3. The criterion appropriately considers practical sample sizes for RP and TP lots.
4. Sponsors are awarded for increasing TP lots.
5. Although necessarily subjective, criterion can be described as “reasonable” from a visual perspective (see Fig. 17.8).

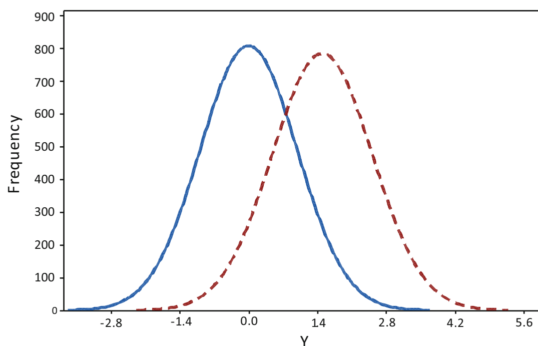


Fig. 17.8 Two populations of equal variance separated by $1.5\sigma_R$

Selection of K for Tier 2 CQAs

The FDA guidance states that K in Eq. (17.13) must be appropriately justified, but provides no guidance. Earlier publications (Chow et al. 2016; Tsong et al. 2016) recommended multiples of σ_R around the mean of the RP as a basis to select K. For example, multipliers 1.645, 1.96, and 3 cover 90%, 95% and 99.7%, respectively, of a standard normal population. The limitation of this approach is that the multiplier can correctly bound the area under a normal curve only if the population mean and standard deviation are known or if the sample sizes are sufficiently large. In practice, the mean and standard deviation are estimated from samples that are typically small. If this rationale is followed, the multiplier for the requisite 90% coverage for Tier 2 assessment is 1.645 which will result in an extremely narrow quality range. Further, the consideration of the required percentage of TP lots that must fall within the quality range under this paradigm implicitly assumes that the TP and RP have equal means and variances. One might argue that this renders the Tier 2 assessment more stringent than the Tier 1 assessment because Tier 1 criteria allows a non-zero shift in means of $\sigma_R/8$. This is counter to the FDA desire that a Tier 2 assessment should be less rigorous than the Tier 1 assessment.

One way to scientifically justify K is to assess the width of the calculated quality range relative to some scientifically based threshold value that is known to have no scientific impact on safety or efficacy. As an example, consider an impurity attribute that has no known impact on potency if it is under the threshold value. If the threshold value exceeds the width of the quality range, then K is deemed scientifically justified.

Although a scientific justification of K is ideal, estimation of the quality range still relies on estimates from typically small sample sizes. Sampling variability has to be taken into consideration in the scientific justification. Sample size-dependent probabilistic inferences can be made using simulations when the respective true population means and population variances of the RP and TP are assumed known. Montes (2016) developed an algorithm that for a given set of parameter values $[\mu_R, \sigma_R, \mu_T, \sigma_T]$, allows the value of K to be chosen such that at least 90% of the

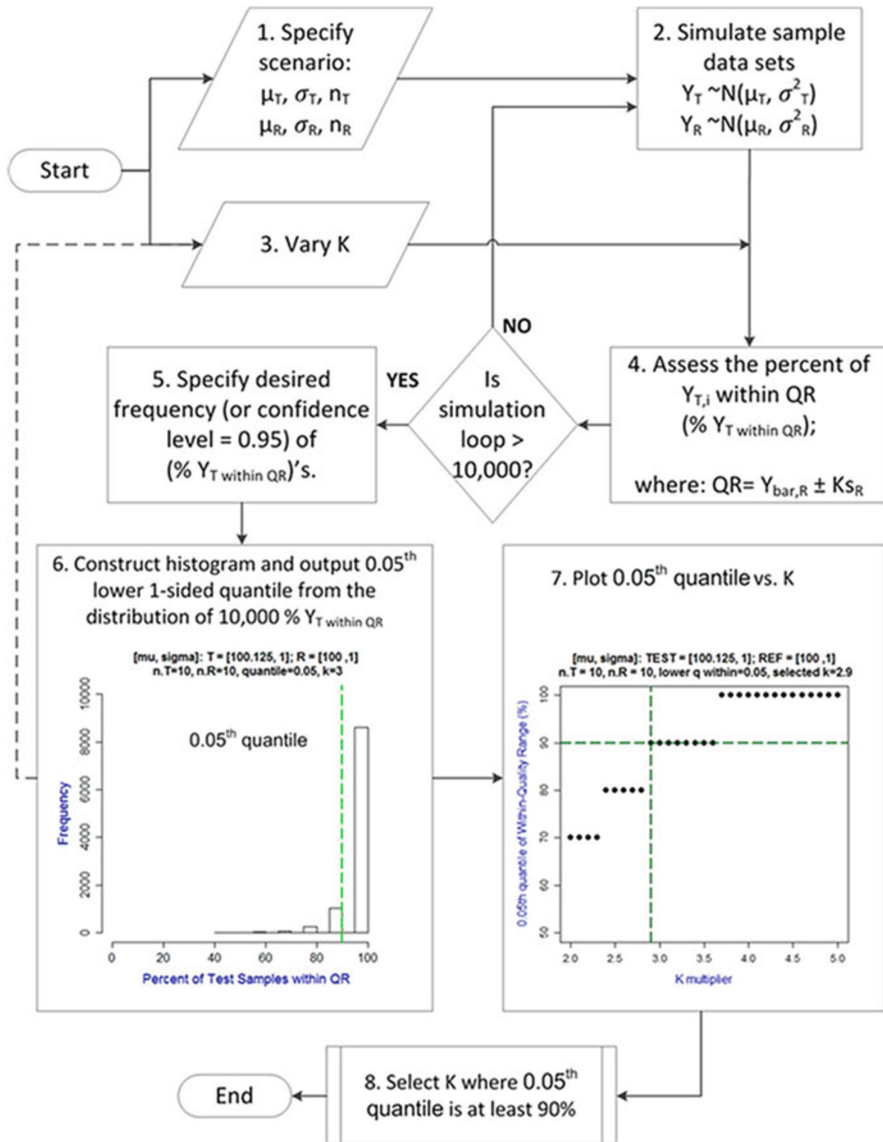


Fig. 17.9 Flow diagram of algorithm to derive a *K* multiplier customized to the allowable mean shift and sample size

sampled TP values are within the quality range at a high frequency (i.e., ≥95%) when TP and RP are similar. The details of the simulation-based algorithm are diagrammed in Fig. 17.9.

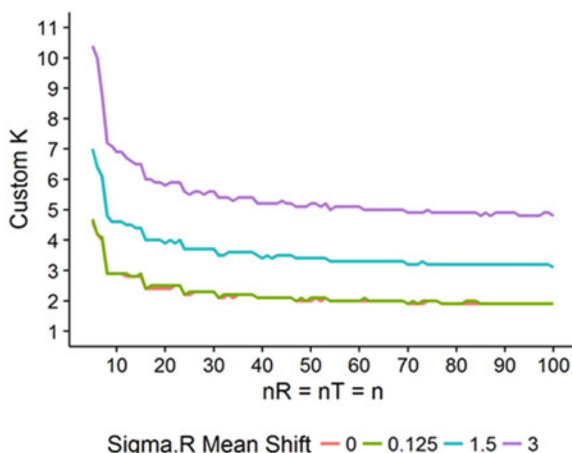


Fig. 17.10 Custom K vs. sample size by mean shift

Briefly, 10,000 random samples are simulated from two independent normal populations with means and standard deviations $[\mu_R, \sigma_R]$ and $[\mu_T, \sigma_T]$. The percent of sampled TP lots falling within the computed quality range based on the RP sample and a fixed value of K is computed and denoted as $\%Y_{T,withinQR}$. The lower 0.05th quantile of $\%Y_{T,withinQR}$ (i.e., the value less than 95% of the simulated values) is output for that K value. These steps are repeated for varying values of K. The smallest K value where the 0.05th quantile is at least equal to the targeted coverage of 90% is chosen as the custom K for the given set of $[\mu_R, \sigma_R, \mu_T, \sigma_T]$. In the example scenario where $n_R = n_T = 10$ with $|\mu_T - \mu_R| = 0.125 \times \sigma_R$, the selected custom K is 2.9. Figure 17.10 reports custom K values as a function of sample size (varying from 5 to 100) and mean shifts of $|\mu_T - \mu_R|$equal to 0, 0.125, 1.5, and 3 times σ_R when TP and RP lots have equal sample sizes and variances. Note the lines for 0 and 0.125 are effectively on top of each other.

Figure 17.10 demonstrates that a selection of $K = 3$ seems quite reasonable with the typical samples sizes of 10 for both TP and RP as recommended in the FDA guidance. With smaller sample sizes, there is more uncertainty in the true values of the means and variances, and hence a greater K multiplier is required to assure that at least 90% of sampled biosimilar lots fall within the quality range with a frequency of at least 95%. The custom K for mean shift $0.125 \times \sigma_R$ starts at 4.7 for $n = 5$ but decreases to 2 as sample size increases. The custom K for mean shifts greater than $0.125 \times \sigma_R$ has to be greater to achieve the same frequency of passing Tier 2. The custom K for mean shift of $1.5 \times \sigma_R$, which is at the boundary of what is considered equivalent in a Tier 1 assessment, starts at 7 for $n = 5$ but decreases to 3 as sample size increases. The custom K for mean shift of $3.0 \times \sigma_R$ is as great as 10 but decreases to 5 as sample size increases. The algorithm is predicated on the premise that the claimed mean shift can be scientifically justified for the attribute. Using a large K multiplier for the sake of passing Tier 2 assessment should not take precedence over scientific judgment that when biosimilar significantly differ from reference (e.g., mean shift of $3.0 \times \sigma_R$), it should fail the Tier 2 assessment.

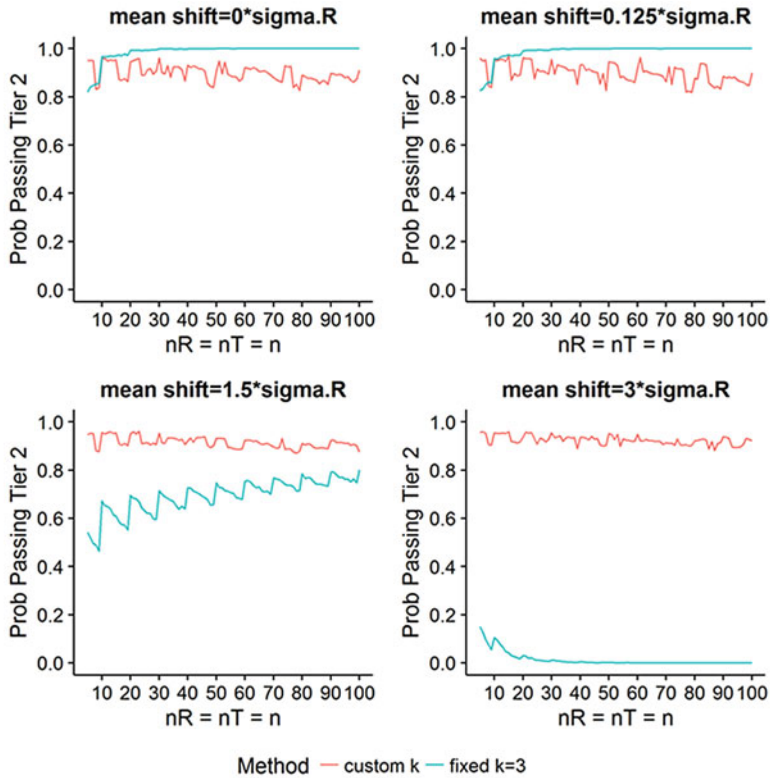


Fig. 17.11 Probability of passing Tier 2 vs. sample size and method to select K

As referenced in Tier 2 Testing, the majority of sponsor BLA submissions to date have used $K = 3$ for Tier 2 assessment. Whether the choice of $K = 3$ is truly scientifically justified for each attribute is not clear. The probability of passing Tier 2 assessment under varying mean shifts for equal sample sizes, equal variances scenario using $K = 3$ and the custom K approach described in Fig. 17.9 are compared in Fig. 17.11.

Note that the probability profiles for both K methods are non-smooth and non-monotonic. As more fully discussed in Tier 2 criticisms, this occurs due to the binary pass-fail outcome for each individual value. For a mean shift less than or equal to $0.125 \times \sigma_R$, $K = 3$ has at least 80% probability of passing Tier 2 at $n = 5$, and this increases to 100% when n is at least 30. At mean shift $1.5 \times \sigma_R$ with sample sizes $n = 10$, the probability of passing Tier 2 is less than 50% with $K = 3$. So if a mean shift as large as $1.5 \times \sigma_R$ can be scientifically justified, the probability of passing the Tier 2 assessment with $K = 3$ will be no better than flipping a coin. In contrast, the probability of passing the Tier 2 assessment for an acceptable mean shift is maintained around 90% using the custom K approach regardless of the claimed mean shift scenario because the custom K approach adapts to the mean shift scenario

and sample size as shown in Fig. 17.10. Therefore the custom K approach performs better than $K = 3$ in correctly passing Tier 2 assessment for any given mean shift scenario that can be scientifically justified.

The proposed custom K algorithm aims to detect departure from the three key assumptions made in the simulation.

1. Data are normally distributed,
2. The actual value of $|\mu_T - \mu_R|$ does not exceed the hypothesized true mean shift, and
3. Variances of the two products are equal.

If Tier 2 assessment fails to conclude similarity, sponsors should investigate which of these assumptions are not met. If a biosimilar has much greater variability than the RP, future manufacturing control strategies to reduce biosimilar variability can be instituted. If procured manufactured RP lots are correlated, this may explain the relatively smaller variance for the RP lots. If an attribute is not normally distributed, it will be incongruent with the quality range framework where normality is assumed, hence impacting the probability of passing a Tier 2 assessment. Investigation findings such as these can be incorporated in the totality of evidence to support conclusion of overall biosimilarity.

Criticisms of the FDA Approach

The requirement for developing an analytical similarity plan that contains a statistical analysis is new. Although it is a good practice to promote, an analytical similarity plan is challenging to implement. One problem is that there is a lack of control over the RP process and it may change during the development of the proposed biosimilar product.

On the statistical side, the recommendations in the draft guidance are quite consistent with the previous presentations and publications by FDA statisticians (Tsong et al. 2016). However, the draft guidance does not mention three specific details that were previously communicated to sponsors either via public conferences or private sponsor-FDA meetings:

1. Lot independence: In early communications, independent lots were required for the statistical similarity assessment due to the potential impact of correlated lots on the equivalence test (Shen et al. 2016; Yang et al. 2016). However, drug product (DP) lots from the same drug substance lot are not independent, and this causes difficulties with the proposed statistical methods.
2. Equal vs. unequal variance: The unequal variance assumption seems to be the preferred approach based on FDA analyses to this point, but this is not specifically stated in the guidance.
3. Unequal sample size: Early in the process, the FDA was concerned about a sponsor's attempt to increase power of the Tier 1 equivalence tests by merely

increasing the number of RP lots. To mitigate this concern, they recommended splitting the RP lots into two groups. One group was to be used to estimate σ_R, and the other group to perform the equivalence test (Tsong et al. 2016). However, this approach can lead to different conclusions based on the manner in which the data are split (Burdick et al. 2016). For these reasons, an alternative approach was developed based on an adjustment to the degrees of freedom in the equivalence test (Burdick et al. 2017; Dong et al. 2017). However, this requirement is not mentioned in the guidance, and it is not clear if this concern remains within the Agency.

Tier 1 Criticisms

The draft guidance notes that the Tier 1 equivalence test suffers from inflation of the Type I error (i.e. the patient risk). This is because σ_R is unknown and must be estimated from the RP lots sourced by the sponsor from the market. In addition, the RP lots may be correlated if sourced from the same drug substance. This introduces an additional source of variation, which potentially may further inflate the Type I error (Burdick et al. 2017; Yang et al. 2016). Methods are proposed in Alternative Approaches to Address Criticisms of FDA Approach to mitigate these problems.

Some other practical questions not addressed in the guidance include

1. What approach should be taken if the data distribution is not well-represented with the normal distribution, or not even symmetric?
2. Can the equivalence margin be adjusted to ensure adequate power if there are fewer than 10 lots?
3. If TP and/or RP lots are found to be correlated, what is the desired statistical approach to handle the analysis?
4. Is equivalence testing the best approach for Tier 1 attributes?

The last question is particularly important as the equivalence tests focuses only on the mean difference between the RP and the TP. It is probably more clinically relevant to ensure that the TP distribution is similar to, or falls completely within the RP distribution. Failure to assess similarity over the entire distribution potentially increases both patient and manufacturer risk. Giacoletti and Heyse (2011) provide one such approach for considering the entire distribution as opposed to only the mean.

Inflation of Type I Error Rate in Tier 1 Test

Burdick et al. (2017) performed computer simulations to determine the realized type I error rate and power for the Tier 1 equivalence test recommended in the FDA guidance. Sample means and variances were simulated for TP and RP lots assuming equal variances for given values of n_T and n_R where n_T ranges from 6 to 20 and n_R/n_T ranges from 1 to 5. The true mean difference is either $|\mu_T - \mu_R| = 0.125$ or

Table 17.8 Simulated type I error (nominal level of 0.05) and power at $|\mu_T - \mu_R| = 0.125$

n_T	n_R	Simulated type I error rate at $\lvert\mu_T - \mu_R\rvert = 1.5$	Simulated power at $\lvert\mu_T - \mu_R\rvert = 0.125$
6	6	0.051	0.479
6	12	0.049	0.729
10	10	0.063	0.803
10	20	0.057	0.939
15	15	0.067	0.944
15	30	0.063	0.990

$|\mu_T - \mu_R| = 1.5$ with $\sigma_R = 1$. The simulated sample statistics were used to compute 90% two-sided confidence intervals on $\mu_T - \mu_R$. The simulated type I error rate is calculated as the proportion of times that the confidence interval falls within the appropriate equivalence margin when the true mean difference is $|\mu_T - \mu_R| = 1.5$. The specified type I error rate is 0.05. The power of the test is calculated as the proportion of times that the confidence interval falls within the appropriate equivalence margin when the true mean difference is $|\mu_T - \mu_R| = 0.125$. A portion of the results is shown in Table 17.8.

Examination of Table 17.8 demonstrates the following results:

1. The test recommended by the FDA yields an inflated type I error rate. That is, the true error rate exceeds the desired 0.05 in most every row of Table 17.8. This is an increased risk to patient that is caused by the need to estimate σ_R.
2. The inflated type I error rate increases as n_T increases for a fixed value of n_R/n_T. Thus, increasing both n_R and n_T at the same rate makes the problem worse.
3. The results in Table 17.8 differ from those in Table 17.7 because FDA assumed σ_R was known and did not account for estimation error. Note that for the design where there are 10 product lots from each group, the type I error rate increases from 0.05 (Table 17.7) to 0.063 (Table 17.8) and that the power decreases from 0.87 (Table 7) to 0.803 (Table 8). This drop in power represents an increased risk to the sponsor.

In summary, as noted in the FDA guidance and demonstrated in Table 17.8, the proposed FDA approach does not control type I error rate. Additionally, this approach suffers a loss in power at $|\mu_T - \mu_R| = 0.125$, thereby increasing the risk of not declaring two products to be analytically similar, when this is indeed the case. Burdick et al. (2017) propose a simple change to the FDA formulation that remedies this problem and is described in Alternative Approaches to Address Criticisms of FDA Approach of this chapter.

Impact of Correlation on Tier 1 Test

Yang et al. (2016) conducted a simulation study to evaluate the effect of correlation on both type I error rate and power. Ten biosimilar lots and either 10 or 20 RP lots were used in the simulation. The correlation of RP lots sourced with the same drug

Table 17.9 Impact of correlation on Tier 1 equivalence test

n_T	n_R	Correlation of RP lots sourced with same DS	Simulated type 1 error rate (desired is 0.05)	Simulated power at $0.125 \times \sigma_R$
10	10	0.8	0.081	0.674
10	10	0	0.062	0.802
10	20	0.8	0.075	0.885
10	20	0	0.057	0.937

substance lot was assumed to be either 0 or 0.8 with pairs of DP lots sourced from a single DS lot, and the RP and TP variances were assumed equal. The results are presented in Table 17.9.

As seen from the table, the type I error rate increases and the power decreases as the correlation increases. Mitigations to this problem are described in Alternative Approaches to Address Criticisms of FDA Approach.

Tier 2 Criticisms

The major decision with the Tier 2 analysis is selection of K in Eq. (17.13). This value is not clearly prescribed in the guidance, but $K = 3$ has been the most selected value in submissions to date.

There is a problem with the rule that 90% of the TP items must fall in the quality range, because it creates a disincentive for sponsors to increase the number of TP lots beyond the required number of 10. Assuming that every TP lot has an equal chance, p, to fall within the quality range, the probability that at least 90% of the TP lots fall within the quality range is defined by the binomial probability distribution as

$$\sum_{x=c}^{n_T} \frac{n_T!}{x!\,(n_T - x)!} p^x (1 - p)^{n_T - x} \tag{17.18}$$

where c is the smallest integer greater than or equal to $90\,\% \times n_T$. The disincentive of acquiring more lots is created because the probability of passing the test decreases as n_T increases when $n_T - c$ is constant. This is demonstrated in Table 17.10. The probability of passing assumes p = 0.997 and that the TP and RP populations have the same mean and variance.

Table 17.10 demonstrates that when n_T increases and $n_T - c$ remains constant, the probability of passing the Tier 2 test decreases. Thus, unless a sponsor wants to increase TP lots from 10 to 20, the chances of passing the test will decrease with increasing TP lots. This penalty for increasing TP lots is counter to the Tier 1 test that is designed to reward taking additional lots beyond $n_T = 10$.

Table 17.10 Probability of passing Tier 2 test with p = 0.997

n_T	c	$n_T - c$	Probability of passing
10	9	1	0.970
11	10	1	0.967
12	11	1	0.965
13	12	1	0.962
14	13	1	0.959
15	14	1	0.956
16	15	1	0.953
17	16	1	0.950
18	17	1	0.947
19	18	1	0.945
20	19	1	0.942
21	19	2	0.998

Table 17.11 Simulated probabilities of passing Tier 2 test

K	n_R	n_T	Number of TP lots in quality range to pass	RP lots uncorrelated	RP lots correlation of 0.80
2	16	8	8	0.594	0.510
3	16	8	8	0.923	0.870
2	16	10	9	0.825	0.733
3	16	10	9	0.987	0.959

Impact of Correlation on Tier 2 Test

Yang et al. (2016) carried out a simulation study to assess the effect of correlation on demonstration of Tier 2 analytical similarity. The simulation consists of various scenarios with various combinations of n_T and n_R with a correlation of 0.80 and $K = 2$ and 3, when RP and TP lots have equal means and variances. Analytical similarity is claimed if over 90% of TP lots fall within the quality range computed from the RP lot data. The probability of successful demonstration of analytical similarity is presented in Table 17.11.

It is evident that the correlation has a significant effect on probability of passing the Tier 2 test. As expected, the probability of passing Tier 2 is greater for $K = 3$ than for $K = 2$, and $K = 2$ is deemed to be too small given the RP and TP lots in the simulation were identical.

Alternative Approaches to Address Criticisms of FDA Approach

Several authors have recommended improvements within the FDA paradigm of equivalence testing for Tier 1 and quality ranges for Tier 2. Some of these approaches are described in this section.

Reformulation of Hypotheses to Consider Effect Size in the Tier 1 Test

Burdick et al. (2017) provide a simple adjustment to the hypotheses shown in (17.1) that provides an equivalence test that maintains the desired type I error rate of 0.05. In particular, they suggest reformulating the hypotheses as

$$\begin{aligned} &H_0 : \tfrac{\mu_T-\mu_R}{\sigma_R} \leq -1.5 \text{ or } \tfrac{\mu_T-\mu_R}{\sigma_R} \geq 1.5 \\ &H_A : \ -1.5 < \tfrac{\mu_T-\mu_R}{\sigma_R} < 1.5. \end{aligned} \tag{17.19}$$

That is, each side of the equations in (17.1) are divided by σ_R. By making this change, the right-hand sides of the equations are free of unknown parameters, and σ_R is now contained in the parameter of interest, $\lambda = \frac{\mu_T-\mu_R}{\sigma_R}$. This ratio λ is often referred to as the effect size relative to the RP. The TOST procedure can again be used to test the hypotheses by constructing a 90% confidence interval on λ. If this computed confidence interval falls within the range from –1.5 to +1.5, the null hypothesis is rejected and equivalence is demonstrated.

Confidence Interval on Effect Size Assuming Equal Variances

The 90% confidence interval on the effect size is based on the inversion confidence interval principle as described by Kelley (2007). If it is assumed that the variances for the TP and RP lots are equal, the maximum likelihood estimator (with degree of freedom correction) for λ is

$$\begin{aligned} &\widehat{\lambda} = \tfrac{\overline{Y}_T-\overline{Y}_R}{S_P} = \left[\frac{\overline{Y}_T-\overline{Y}_R}{S_P\sqrt{\frac{1}{n_T}+\frac{1}{n_R}}}\right] \times \sqrt{\tfrac{1}{n_T}+\tfrac{1}{n_R}} = t_{calc} \times \sqrt{\tfrac{1}{n_T}+\tfrac{1}{n_R}} \\ &S_P = \sqrt{\tfrac{(n_T-1)S_T^2+(n_R-1)S_R^2}{n_T+n_R-2}}. \end{aligned} \tag{17.20}$$

The statistic t_{calc} is the familiar test statistic used to test equality of means for two independent groups. To form a confidence interval for λ based on $\widehat{\lambda}$, first form a confidence interval on the non-centrality parameter

$$ncp = \frac{\lambda}{\sqrt{\frac{1}{n_T}+\frac{1}{n_R}}} \tag{17.21}$$

associated with t_{calc}. Once a two-sided confidence interval is obtained for *ncp*, each bound is multiplied by $\sqrt{\frac{1}{n_T}+\frac{1}{n_R}}$ to obtain a confidence interval on λ.

To demonstrate, the following SAS code is used to compute a 90% confidence interval on λ. The function "tnonct" returns the value of *ncp* that yields the specified probability (e.g., 0.95 or 0.05) for the observed value of t_{calc}.

```
tcalc = meandiff/sqrt(pooledvar*(1/nt + 1/nr));
lbeffect = tnonct(tcalc,nt + nr − 2,.95)*sqrt(1/nt + 1/nr);
ubeffect = tnonct(tcalc,nt + nr − 2,.05)*sqrt(1/nt + 1/nr);
```

Note that this procedure provides an exact confidence interval on λ.

Confidence Interval on Effect Size With Unequal Variances

When it cannot be assumed that $\sigma_R = \sigma_T$, estimation of σ_R must be based only on the RP lots with no pooling of the variances for the two products. A useful approach for constructing a confidence interval on the effect size in this situation is to employ a generalized confidence interval (GCI). Tsui and Weerahandi (1989) introduced the concept of generalized inference for testing hypotheses when exact methods do not exist. Weerahandi (1993) extended this concept to construct GCIs. Hannig et al. (2006) have shown that under most practical conditions, these intervals provide correct frequentist coverage.

For the present application, a GCI can be computed using the following steps:

1. Compute $\overline{Y}_R$, $\overline{Y}_T$, and S_R^2 and S_T^2 for the sample data sets of size n_R and n_T, respectively.
2. Simulate N values of the effect size:

$$\lambda_{sim} = \frac{\overline{Y}_T - \overline{Y}_R - Z \times \sqrt{\frac{(n_T-1)\times S_T^2}{n_T \times W_T} + \frac{(n_R-1)\times S_R^2}{n_R \times W_R}}}{\sqrt{\frac{(n_R-1)\times S_R^2}{W_R}}} \tag{17.22}$$

where W_R is a chi-squared random variable with $n_R - 1$ degrees of freedom, W_T is a chi-squared random variable with $n_T - 1$ degrees of freedom, and Z is a standard normal random variable with mean 0 and variance 1. A value of $N \geq 100{,}000$ simulations is recommended, although a value as low as 10,000 generally works satisfactorily.

3. Order the N simulated λ_{sim} values obtained in Step 2 from least to greatest.
4. Define the lower bound for a two-sided $100(1 - 2\alpha)\%$ confidence interval as the value in position $N \times \alpha$ of the ordered data set in Step 3. Define the upper bound as the value in position $N \times (1 - \alpha)$ of this same ordered set. For example, if $N = 100{,}000$ the lower bound of a 90% two-sided confidence interval is the value in position $100{,}000 \times 0.05 = 5{,}000$ and the upper bound is the value in position $100{,}000 \times 0.95 = 95{,}000$.

Table 17.12 Summary statistics of TNF-α RGA binding method

Product	No. of lots	Mean	SD
GP2015 (TP)	$n_T = 9$	$\overline{Y}_T = 92.44$	$S_T = 2.07$
US Enbrel (RP)	$n_R = 13$	$\overline{Y}_R = 93.69$	$S_R = 7.47$

Note that these steps can be computed with any software package that contains sorting and simulation functions including Excel. An example calculation is provided in the next section. Simulation results using effect size provides simulation results that demonstrate this procedure provides desired type I error rates and powers.

Numerical Example

A 90% confidence interval is computed for the effect size using the Tier 1 data from Table 17.4 and reproduced in Table 17.12.

Although the standard deviations suggest an unequal variance model is more appropriate than the equal variance model, for illustration the confidence interval is first computed assuming equal variances. For this calculation

$$\begin{aligned} S_P &= \sqrt{\frac{(n_T-1)S_T^2+(n_R-1)S_R^2}{n_T+n_R-2}} = \sqrt{\frac{(9-1)2.07^2+(13-1)7.47^2}{9+13-2}} = 5.932 \\ \widehat{\lambda} &= \frac{\overline{Y}_T-\overline{Y}_R}{S_P} = \frac{92.44-93.69}{5.932} = -0.211 \\ t_{calc} &= \frac{\widehat{\lambda}}{\sqrt{\frac{1}{n_T}+\frac{1}{n_R}}} = \frac{-0.211}{\sqrt{\frac{1}{9}+\frac{1}{13}}} = -0.486 \end{aligned} \tag{17.23}$$

and the resulting 90% two-sided confidence interval on *ncp* is from −2.130 to 1.170 (using the SAS code provided earlier). Converting this interval to the 90% confidence interval on the effect size

$$\begin{aligned} L &= \sqrt{\frac{1}{n_T}+\frac{1}{n_R}} \times -2.130 = \sqrt{\frac{1}{9}+\frac{1}{13}} \times -2.130 = -0.92 \\ U &= \sqrt{\frac{1}{n_T}+\frac{1}{n_R}} \times 1.170 = \sqrt{\frac{1}{9}+\frac{1}{13}} \times 1.170 = 0.51. \end{aligned} \tag{17.24}$$

Note the effect size has no units of measure. Since this interval falls entirely within the range from –1.5 to +1.5, equivalence has been demonstrated.

The algorithm described when assuming unequal variances is now used to compute the 90% confidence interval on λ. Figure 17.12 presents 10 rows of an Excel sheet that demonstrates the required calculation by simulating the random chi-squared values using the uniform distribution and the Excel function CHIINV to obtain W_1 and W_2.

Table 17.13 presents a summary for 10,000 iterations that produces the 90% confidence interval from –0.66 to 0.33.

W1 uniform	W1	W2 uniform	W2	Z	Mean1	Mean2	SD1	SD2	n1	n2	λ sim
0.361	13.118	0.425	8.084	-0.133	93.690	92.440	7.470	2.070	13	9	0.214
0.300	14.014	0.474	7.593	0.620	93.690	92.440	7.470	2.070	13	9	-0.003
0.677	9.304	0.566	6.729	-0.045	93.690	92.440	7.470	2.070	13	9	0.160
0.379	12.861	0.419	8.145	-1.203	93.690	92.440	7.470	2.070	13	9	0.526
0.020	24.056	0.436	7.970	1.985	93.690	92.440	7.470	2.070	13	9	-0.372
0.610	10.063	0.262	10.046	-0.247	93.690	92.440	7.470	2.070	13	9	0.224
0.529	11.004	0.836	4.227	0.592	93.690	92.440	7.470	2.070	13	9	-0.019
0.522	11.084	0.424	8.095	0.156	93.690	92.440	7.470	2.070	13	9	0.115
0.352	13.241	0.839	4.192	-0.379	93.690	92.440	7.470	2.070	13	9	0.293
0.777	8.108	0.217	10.741	-0.164	93.690	92.440	7.470	2.070	13	9	0.184

Fig. 17.12 Example Excel worksheet

Table 17.13 Summary of 10,000 Values of λ_{sim}

Mean	–0.16
Standard deviation	0.301
Minimum	–1.24
Maximum	1.00
Count	10,000
Largest(500)-95th percentile	0.33
Smallest(500)-5th percentile	–0.66

Again, since this confidence interval falls within –1.5 and +1.5, equivalence has been demonstrated. Note that the mean in Table 17.13 is very close to the estimated effect size,

$$\frac{\overline{Y}_T - \overline{Y}_R}{S_R} = \frac{92.44 - 93.69}{7.47} = -0.17. \tag{17.25}$$

Simulation Results Using Effect Size

Burdick et al. (2017) used computer simulation to compare the effect size approach with the FDA approach. Figures 17.13 and 17.14 present type I error rates and powers, respectively, when $\sigma_R = \sigma_T$. Figures 17.15 and 17.16 provide results when it cannot be assumed that variances are equal, and the effect size confidence interval is computed using generalized confidence intervals. These simulations demonstrate that although the FDA procedure has an inflated type I error rate, the methods based on the effect size successfully maintain the type I error rate. This is true for both the equal and unequal variance assumptions. Additionally, as shown in Fig. 17.14, when variances are assumed equal, the effect size interval has greater power than the FDA interval at $|\mu_R - \mu_T| = 0.125 \times \sigma_R$. For the unequal variance case, the effect size interval is somewhat less powerful.

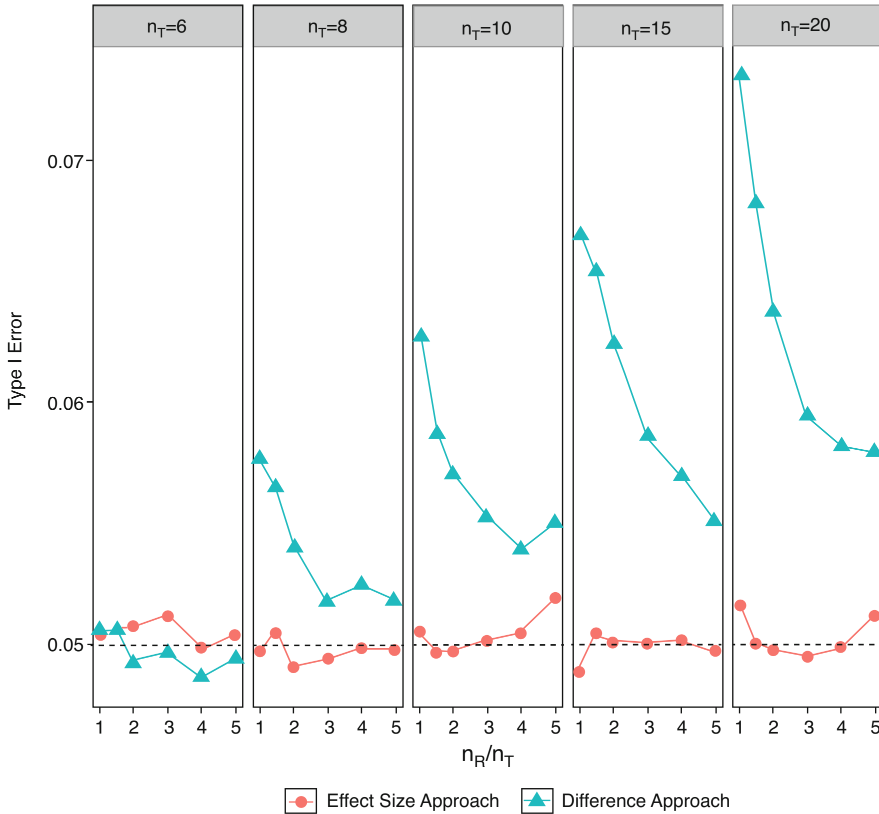


Fig. 17.13 Simulated type I error rates assuming equal variances and desired value of 0.05

Mitigation of Correlation Effects

The second major deficiency with the FDA approach concerns the problems associated with correlated RP lots as described in Criticisms of the FDA approach. Yang et al. (2016) provide two strategies for dealing with this issue:

1. Purchase RP lots over an extended time frame to decrease the likelihood of obtaining drug product (DP) lots sourced with the same drug substance (DS). This strategy will also provide an opportunity for the lot-to-lot variation to fully manifest and provide a better estimate of σ_R.
2. Identify DP lots that have been sourced by the same DS, and use this information to fit a statistical model that properly adjusts for the correlation structure. Two possible approaches for identifying the DS source are

   (a) Analytical determination using a stable isotope profile. (See, e.g. Apostol et al. (2001)), and

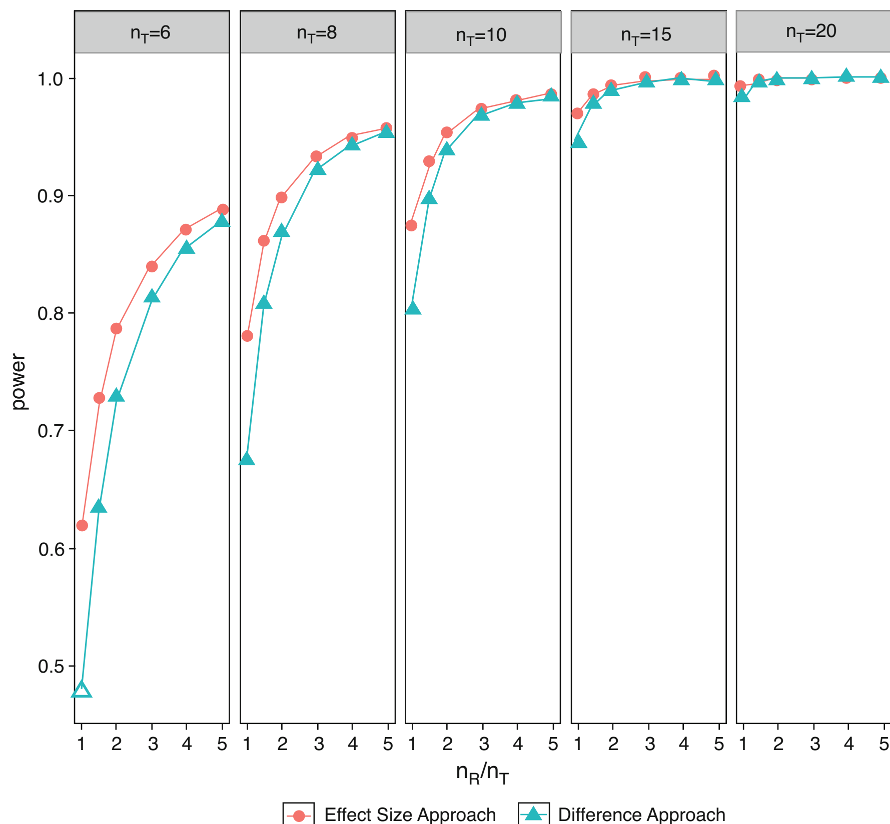


Fig. 17.14 Power assuming equal variances at mean difference of $|\mu_R - \mu_T| = 0.125 \times \sigma_R$

(b) Empirical assessment based on repeated measures of the RP drug product lots.

Strategy 1 provides the recommendation offered in the FDA guidance. Strategy 2 requires the ability to identify the relationship of the RP lots. Yang et al. (2016) provide an example of an empirical assessment to identify such a relationship.

As previously discussed, the correlation among RP lots may have a negative effect on both the type I error rate and the statistical power. However, the impact may be mitigated through proper modelling if sourcing of DP lots can be identified. Such an effort is consistent with the FDA guidance that recommends that the RP variability should encompass both the within-lot and between-lot variance components.

Yang et al. (2016) derived a GCI for the effect size under the assumption that reportable values of the RP lots are correlated. To demonstrate, assume that it is discovered that each drug substance DS lot sources $b = 2$ DP lots. Further, assume

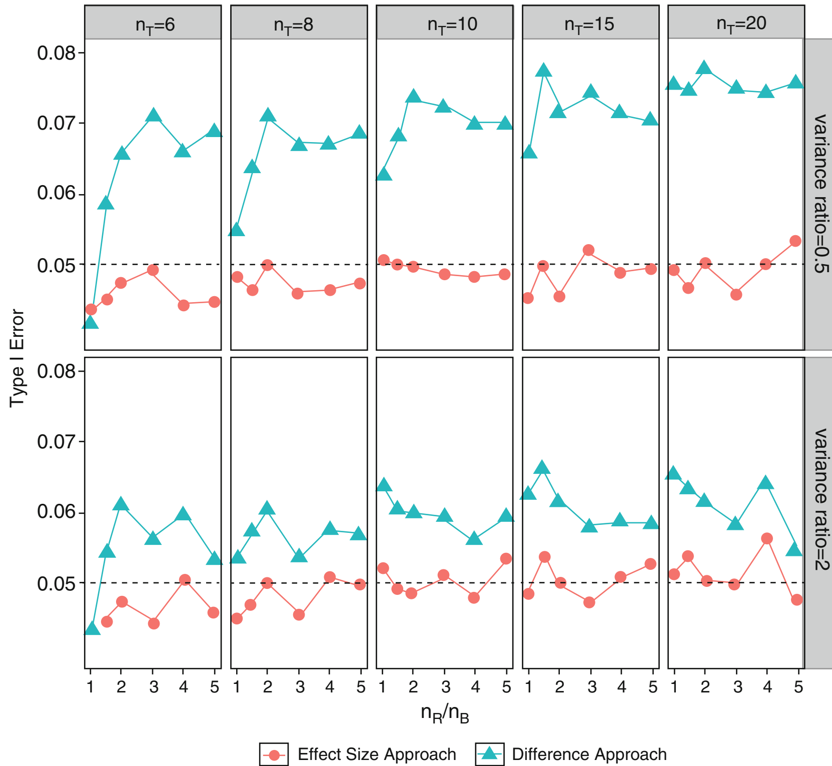


Fig. 17.15 Type I error assuming unequal variances and desired value of 0.05 where variance ratio $= \sigma_T^2/\sigma_R^2$

10 DP lots are collected consisting of pairs of DP lots sourced from $a = 5$ DS lots. A statistical model to describe a CQA from the RP lots is

$$\begin{aligned} Y_{Rij} = \mu + L_{Ri} + E_{Rij} \quad i = 1, \ldots, a; \quad j = 1, \ldots, b; \\ L_{Ri} \sim N\left(0, \sigma_{DS}^2\right) \\ E_{Rij} \sim N\left(0, \sigma_E^2\right) \end{aligned} \tag{17.26}$$

where Y_{Rij} is the CQA value of the jth DP lot filled with the ith DS lot, L_{Ri} is a random effect that represents differences among the $a = 5$ DS lots with mean zero and variance σ_{DS}^2, and E_{Rij} is a random effect that represents variability between the $b = 2$ DP lots filled with the same DS lot with mean zero and variance σ_E^2. Estimators for the variances σ_{DS}^2 and σ_E^2 are based on the statistics

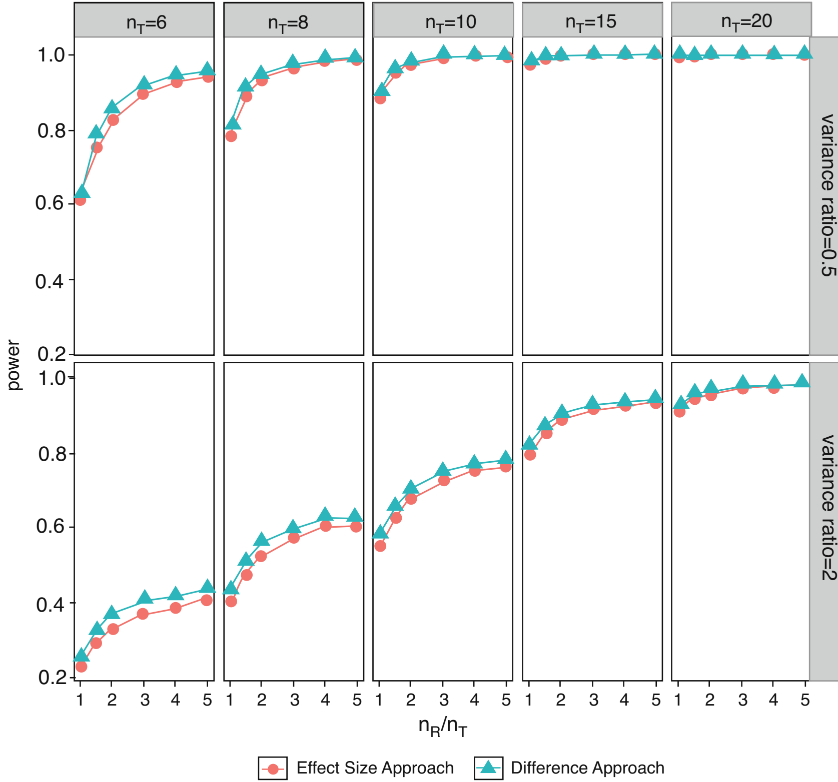


Fig. 17.16 Power assuming unequal variances at mean difference of $|\mu_R - \mu_T| = 0.125 \times \sigma_R$ where variance ratio $= \sigma_T^2/\sigma_R^2$

$$
\begin{aligned}
SSL_R &= b \sum_{i=1}^{a} \left(\overline{Y}_{Ri} - \overline{Y}_R\right)^2 \\
SSW_R &= \sum_{i=1}^{a} \sum_{j=1}^{b} \left(Y_{Rij} - \overline{Y}_{Ri}\right)^2. \\
\overline{Y}_{Ri} &= \frac{\sum_{j=1}^{b} Y_{Rij}}{b} \\
\overline{Y}_R &= \frac{\sum_{i=1}^{a} \overline{Y}_{Ri}}{a}
\end{aligned}
\tag{17.27}
$$

which provide the analysis of variance partitioning of the total variance. The correlation between DP lots sourced with the same DS lot is

$$
\rho = \frac{\sigma_{DS}^2}{\sigma_{DS}^2 + \sigma_E^2}. \tag{17.28}
$$

It is assumed that all TP lots are independent, but this model can also be extended to allow TP lots to be correlated.

A 90% confidence interval on the effect size assuming unequal variances between the RP and TP lots can be computed using a generalized confidence interval approach defined by the following steps.

1. Compute $\overline{Y}_R$ and $\overline{Y}_T$, S_T^2 for the n_T TP lots, and the two sums of squares defined in (17.27) for the $n_R = a \times b$ RP lots, respectively.
2. Simulate N values of the effect size:

$$\lambda_{sim} = \frac{\overline{Y}_T - \overline{Y}_R - Z \times \sqrt{\frac{\frac{SSL_R}{b \times WL_R} + \left(\frac{b-1}{b}\right) \times \left(\frac{SSE_R}{WE_R}\right)}{n_R} + \frac{\frac{(n_T-1)S_T^2}{W_T}}{n_T}}}{\sqrt{\frac{SSL_R}{b \times WL_R} + \left(\frac{b-1}{b}\right) \times \frac{SSE_R}{WE_R}}} \tag{17.29}$$

where W_R is a chi-squared random variable with $n_R - 1$ degrees of freedom, W_T is a chi-squared random variable with $n_T - 1$ degrees of freedom, and Z is a standard normal random variable with mean 0 and variance 1. A value of $N \geq 100{,}000$ simulations is recommended, although a value as low as 10,000 generally works satisfactorily.

3. Order the N simulated λ_{sim} values obtained in Step 2 from least to greatest.
4. Define the lower bound for a two-sided $100(1 - 2\alpha)\%$ confidence interval as the value in position $N \times \alpha$ of the ordered data set in Step 3. Define the upper bound as the value in position $N \times (1 - \alpha)$ of this same ordered set.

Yang et al. (2016) provide simulation results that demonstrate this approach will control the type I error rate in the Tier 1 equivalence test. Figure 17.17 displays simulated type I error rates of the GCI based on the effect size using interval (17.29), GCI interval on the effect size in (17.22) in which independence is assumed, and the FDA approach that also assumes independence. The simulation is performed for $n_R = n_T = 10$, $\rho = 0.8$, and $\sigma_R^2/\sigma_T^2 = 0.5$, 1.0, and 2.0. The horizontal reference line depicts the desired error rate, 0.057, adjusted for simulation error. It can be observed that (1) the FDA approach and the GCI interval based on (17.22) have an inflated type I error rate, and (2) the GCI interval based on (17.29) maintains the type I error rate for all conditions. Yang et al. provide simulations for many more designs that provide similar results.

Power assessment was only carried out for the GCI interval based on (17.29) because it is the only method that maintains type I error rate. The results are shown in Fig. 17.18.

Note that $b = 1$ corresponds to the situation where all DP lots are sourced from different lots, and there is no correlation among DP lots. From the plot, it is evident that even when correlation is properly modeled, its presence has a negative impact on power. Thus, sponsors are encouraged to purchase RP lots over an extended

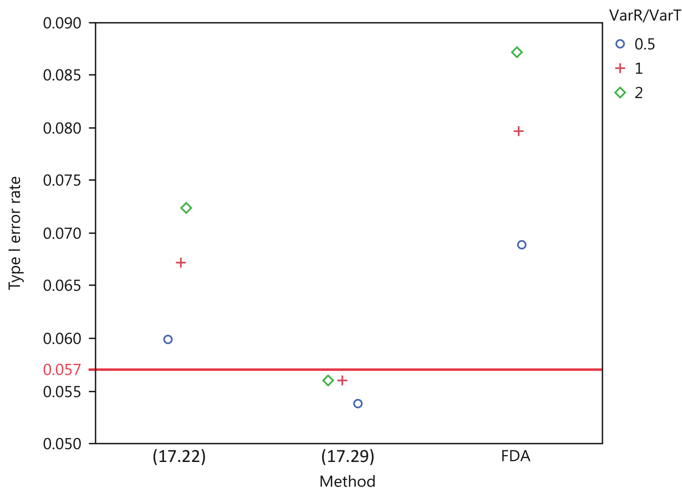


Fig. 17.17 Type I error rate with correlation of 0.8

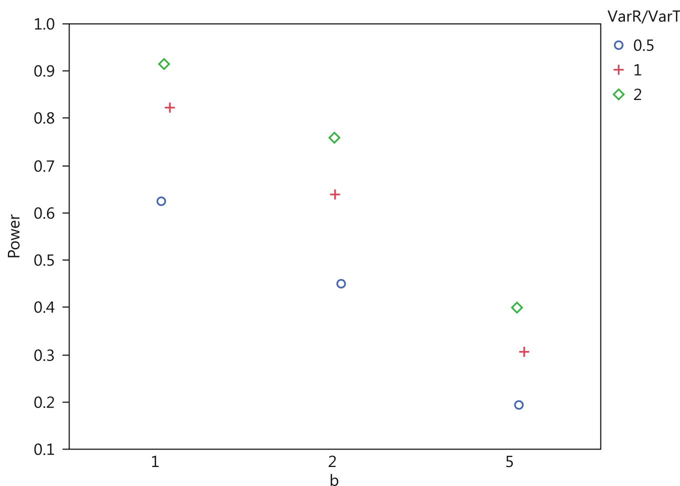


Fig. 17.18 Power of CGI based on (17.28)

time frame to decrease the likelihood of obtaining DP lots sourced with same DS. Likewise, TP lots should be collected so that they are independent.

For Tier 2 testing, Yang et al. (2016) recommend estimating σ_R in the quality range with

$$\sqrt{\frac{SS_L}{(n-1)\,r} + \left(\frac{1}{nr}\right) SS_E}. \tag{17.30}$$

Concluding Remarks

Analytical characterizations are the most sensitive measures for assessing similarity between a biosimilar and reference product. The FDA draft guidance issued in September 2017 but withdrawn June 2018 recommends a tiered approach and use of statistical methods for assessing analytical similarity. When appropriately applied, statistical methods can provide a high degree of confidence for demonstration of analytical similarity. However, development of such statistical methods can be challenging due to the limited number of reference lots and lack of knowledge concerning their sourcing information. Further compounding the issue is the fact that Tier 1 and Tier 2 acceptance criteria must be established using RP lots. The uncertainties in these acceptance criteria have been shown to inflate the type I error rate and reduce statistical power.

Due to the withdrawal of the FDA guidance, several alternative approaches and refinements of the present approaches are likely to appear in the future. The following goals would seem to be useful for comparing these approaches.

1. Protect patients from consequences of concluding similarity when products are not similar.
2. Protect sponsors from consequences of concluding lack of comparability when products are in fact comparable (the consequences include a lack of patient access to lower cost treatments).
3. Incentivize sponsors to acquire process knowledge concerning the biosimilar product.
4. Enable decision making with practical sample sizes.
5. Examine entirety of the process distribution of product.
6. Statistical rigor should consider criticality and measurement scale of the attribute.
7. Demonstrate robustness to violations of assumptions.
8. The approach must be transparent, easy to explain, and easy to compute by scientists with no formal statistical training.

It is unlikely that any one approach will be uniformly better than all competitors across all goals, but it is possible that one or two approaches will work well enough to satisfy most goals.

References

Apostol I, Brooks PD, Mathews AJ. Application of highprecision isotope ratio monitoring mass spectrometry to identify the biosynthetic origins of proteins. Protein Sci. 2001;10:1466–9.

Berger R, Hsu J. Bioequivalence trials, intersection-union tests and equivalence confidence sets. Stat Sci. 1996;11(4):283–319.

Berkowitz SA. Analytical characterization: structural assessment of biosimilar. In: Endrenyi L, Declerck P, Chow S-C, editors. Biosimilar drug product development. Boca Raton: CRC Press; 2017.

Burdick RK. Comments on analytical similarity session. FDA-Industry Statistics Workshop, Washington DC; 2015.
Burdick RK, Coffey T, Gutka H, Gratzl G, Conlon HD, Huang C-T, Boyne M, Kuehne H. Statistical approaches to assess biosimilarity from analytical data. AAPS J. 2016;19(1):4–14. https://doi.org/10.1208/s12248-016-9968-0.
Burdick RK, Thomas N, Cheng A. Statistical considerations in demonstrating CMC analytical similarity for a biosimilar product. Stat Biopharm Res. 2017;9(3):249–57. https://doi.org/10.1080/19466315.2017.1280412.
Chow SC, Song F, Bai H. Analytical similarity assessment in biosimilar studies. AAPS J. 2016;18(3):670–7.
Dong X, Weng Y-T, Tsong Y. Adjustment for unbalanced sample size for analytical biosimilar equivalence assessment. J Biopharm Stat. 2017;27(2):220–32. https://doi.org/10.1080/10543406.2016.1265544.
EMA. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf. Accessed Dec 19, 2017.
EMA. Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development. 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/03/WC500224995.pdf. Accessed Dec 25, 2017.
FDA. Pharmaceutical cGMPs for the 21st century: a risk-based approach: Final report. 2004.
FDA. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. Silver Spring: The United States Food and Drug Administration. 2015. https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf. Accessed Dec 19, 2017.
FDA. FDA briefing document. Arthritis Advisory Committee Meeting; 2016 Jul 13, BLA 761042, GP2015, a proposed biosimilar to Enbrel(etanercept). 2016. https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM510493.pdf. Accessed Dec 25, 2017.
FDA. Draft Guidance to industry: statistical approach to evaluate analytical similarity. 2017. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM576786.pdf (Link is no longer available).
FDA. FDA homepage. 2018. https://www.fda.gov/ (Search for biosimilar advisory committee). Accessed Feb 12, 2018.
Giacoletti K, Heyse J. Using proportion of similar response to evaluate correlates of protection for vaccine efficacy. Stat Methods Med Res. 2011; https://doi.org/10.1177/0962280211416299.
Hannig J, Iyer H, Patterson P. Fiducial generalized confidence intervals. J Am Stat Assoc. 2006;101:254–69.
ICH. Q1E Evaluation for stability data. 2003. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1E/Step4/Q1E_Guideline.pdf.
ICH. Q9 Quality risk management. 2005. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q9/Step4/Q9_Guideline.pdf. Accessed Dec 19, 2017.
ICH. Q8(R2) Pharmaceutical Development. 2009a. http://www.fda.gov/downloads/Drugs/.../Guidances/ucm073507.pdf. Accessed Dec 19, 2017.
ICH. Q10 Pharmaceutical quality systems. 2009b. http://www.fda.gov/downloads/Drugs/.../Guidances/ucm073517.pdf. Accessed Dec 19, 2017.
ICH. Q11 Development and manufacture of drug substances (Chemical entities and biotechnological/biological entities). 2012. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q11/Q11_Step_4.pdf. Accessed Dec 19, 2017.
Kelley K. Confidence intervals for standardized effect sizes: theory, application, and implementation. J Stat Softw. 2007;20(8):1–22.
Montes RO. Analytical similarity assessment: practical challenges and statistical perspectives. Oral presentation at 2016 Midwest Biopharmaceutical Statistics Workshop. 2016. http://www.mbswonline.com/presentationyear.php?year=2016. Accessed Jul 26, 2018.

Satterthwaite FE. An approximate distribution of estimates of variance components. Biom Bull. 1946;2:110–4.

Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm. 1987;15:657–80.

Shen M, Wang T, Tsong Y.Statistical considerations regarding correlated lots in analytical biosimilar equivalence test. J Biopharm Stat. 2016. https://doi.org/10.1080/10543406.2016.1265541.

Tsong Y Dong X, Shen M. Development of statistical methods for analytical similarity assessment. J Biopharm Stat. 2016. https://doi.org/10.1080/10543406.2016.1272606.

Tsui K, Weerahandi S. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. J Am Stat Assoc. 1989;84:602–7. (Corrections: 86 (1991), p. 256).

Weerahandi S. Generalized confidence intervals. J Am Stat Assoc. 1993;88:899–905. (Corrections: 89 (1994), p. 726).

Yang H, Novick S, Burdick R. On statistical approaches for demonstrating analytical similarity in the presence of correlation. PDA J Pharm Sci Technol. 2016;70(6):547–59.

Part V
Clinical Aspects of Biosimilar Development (Clinical Trials, Immunogenicity, Extrapolation and Interchangeability)

Chapter 18
Comparative Clinical Studies for Biosimilars as Part of a Stepwise Approach

Thomas E. Gwise

Abstract The U.S. Food and Drug Administration (FDA) has implemented an abbreviated pathway to market for products that are demonstrated to be biosimilar to or interchangeable with a licensed biological reference product. We discuss comparative clinical studies supporting biosimilarity in the context of the stepwise approach suggested in FDA guidance documents. First we give a general overview of the Biosimilar pathway followed by a discussion of studies intended to support Biosimilar licensure and several examples with a primary focus being given to equivalence margin selection.

Keywords Biosimilar · TOST

Background

This chapter is a discussion of the ways in which various study types have been implemented through the US Food and Drug Administration's approach to evaluating biosimilar products with a special focus on comparative clinical studies designed to support Biosimilarity. The chapter begins with the origin of the biosimilar pathway and a general overview of the review pathway itself. Also for context, we include a short overview of analytic studies used to support claims that products are highly similar to reference products before presenting the main focus of the chapter, comparative clinical studies designed to support claims that products have no clinically meaningful differences. Three examples discussing similarity margin selection follow. A brief description of studies supporting product interchangeability is then followed by a discussion.

The Biologics Price Competition and Innovation Act (BPCI Act) amended the Public Health Service Act (PHS Act) to create an abbreviated licensure pathway

T. E. Gwise (✉)
FDA, Silver Spring, MD, USA
e-mail: Thomas.gwise@fda.hhs.gov

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_18

for biological products shown to be biosimilar to or interchangeable with an FDA-licensed biological reference product (U.S. Food and Drug Administration 2015b). This pathway is conceptually similar to, although more complicated than the abbreviated pathway for generic drugs. The Hatch-Waxman amendments to the Food Drug and Cosmetics Act were enacted in 1984 and established an abbreviated pathway to market for small molecule drugs having the same active ingredients as legally marketed drugs. Mossinghoff (1999) The requirements for submitting an abbreviated application for a generic drug can be found in the Code of Federal Regulations beginning at 21CFR§314.92 (Code of Federal Regulations 2017a). A major underlying condition for the approval of a generic drug is that the active ingredients are the same. Once a developer demonstrates the active ingredients are the same, the next major hurdle is to show bioequivalence between the generic and the reference product. Requirements for such demonstrations are discussed in the Code of Federal Regulations (21CFR§320) (Code of Federal Regulations 2017b). The generic pathway is inappropriate for biological products because of their complexity and expectations that exactly duplicating biological products' active ingredients would be difficult or impossible. See Woodcock et al. (2007) and Christl et al. (2017) for further background.

Biosimilar Evaluation Process

General

To license a product as biosimilar to a reference product, one must demonstrate that the product is "highly similar" to the reference product, not withstanding any minor differences in clinically inactive components and there are "no clinically meaningful differences" between the proposed biosimilar product and the reference product. These two major criteria are in the PHS Act (42 USC 262 2010) and interpreted in the FDA guidance document Scientific Considerations in Demonstrating Biosimilarity to a Reference Product; Guidance for Industry (Scientific Considerations Guidance) (U.S. Food and Drug Administration 2015b). Satisfying the first criterion (highly similar) of the above set is addressed through analytical studies. The second criterion, showing no clinically meaningful differences, is demonstrated through clinical and sometimes animal studies. Among the clinical studies are pharmacokinetic (PK), pharmacodynamic (PD), immunogenicity, and comparative clinical studies.

Additional information is required before a biosimilar product may be substituted at the pharmacy level without the intervention of the prescriber (42 USC 262 2010). Such biosimilar products are known as interchangeable products. For a product to be interchangeable with a reference product it must be shown to be (1) biosimilar to the reference product and (2) expected to produce the same clinical result as the reference product in any given patient, and also (3) not to pose excessive risks to patients if they switch between the reference product and interchangeable product (42 USC 262 2010; U.S. Food and Drug Administration 2015b),

Stepwise Approach and Totality of the Evidence

The Scientific Considerations Guidance (U.S. Food and Drug Administration 2015b) suggests that a stepwise approach to evaluating biosimilar and interchangeable products be followed. Using a stepwise approach, a biosimilar developer with FDA consultation would be able to interpret information as it is developed and accrues from the various studies mentioned above. Given accrued information, residual uncertainties with respect to the biosimilarity of the proposed product can be considered to help plan further studies to resolve those uncertainties. Using a stepwise approach tends to follow logically because structural and functional evaluations of proposed biosimilars need to be done for safety purposes prior to clinical studies and a demonstration that the biosimilar is"highly similar" to the reference product is necessary for the proposed product to be biosimilar to the reference product. Also, the Scientific Considerations Guidance states that under some circumstances, certain studies may not be necessary in a specific product's evaluation. Finally, biosimilarity being a prerequisite to interchangeability also suggests using a stepwise process. While the stepwise approach may be efficient from certain perspectives, it is feasible that some biosimilar and interchangeable developers may consider conducting the various studies in parallel more appropriate with respect to their objectives. Such an approach is not precluded in the Scientific Considerations Guidance.

The FDA uses evidence gained from various studies in its totality when evaluating biosimilar products (U.S. Food and Drug Administration 2015b). That is, all data and information are considered when making the biosimilarity determination. This "totality of the evidence approach" is different from what is considered the typical drug development process. For example, although the vocabulary associated with efficacy trials is often still used, the idea of there being one "pivotal" clinical trial on which a biosimilar licensing determination is made is not accurate. In fact, the need for a comparative clinical study utilizing a clinical outcome endpoint may be obviated in the biosimilar setting if the analytical, PK and PD studies adequately show biosimilarity without any unresolved residual uncertainties. While some of these ideas may seem revolutionary, the biosimilar pathway is conceptually parallel to the generic drug abbreviated pathway and the concept of "totality of the evidence" is foreshadowed in FDA's 1998 guidance document on clinical evidence requirements (U.S. Food and Drug Administration 1998b).

Analytical Similarity

Introduction

A key requirement for establishing biosimilarity is to show the proposed product is highly similar to the reference product notwithstanding minor differences in clinically inactive components. To put the comparative clinical studies in context,

we introduce the analytical studies designed to support the “highly similar” claim. The FDA has accepted a risk based approach to evaluating proposed biosimilars in order to address the complex challenges of demonstrating analytical similarity while simultaneously creating an abbreviated pathway to licensure. Analytic similarity has been divided into three steps:

1. determining the quality attributes that characterize the reference product in terms of its structural, physiochemical and functional properties;
2. ranking these quality attributes according to their risk of potential clinical impact;
3. evaluating the similarity of these attributes according to one of three tiers of statistical approaches based on a consideration of risk ranking.

The three tiers of statistical analysis used in the analytic similarity assessment for biosimilars has been: equivalence testing, the quality range approach, and visual displays. The details of the statistical methods used can be observed in the examples below.

Comparative Clinical Studies

Design Considerations

We now consider the role of comparative clinical studies from the perspective of the Scientific Considerations Guidance (U.S. Food and Drug Administration 2015b). The object of clinical studies supporting biosimilarity claims is markedly different from that of clinical trials supporting efficacy claims. Drawing this distinction is necessary, since the latter is specifically mentioned as not being goal of the program. In biosimilar evaluation, comparative clinical studies are necessary to resolve residual uncertainties that remain after PK, PD studies and the extensive structural and functional testing needed to support a finding that the products are highly similar. In fact, the guidance reminds the reader that the agency can at its discretion decide that some element(s) may not be necessary in a specific biosimilar application. The clinical study methods deemed appropriate and the number of analyses needed to resolve residual uncertainties and demonstrate biosimilarity are determined on a product-specific basis.

The stepwise approach combined with the totality of the evidence approach suggested in FDA’s Scientific Considerations Guidance (U.S. Food and Drug Administration 2015b) tend to imply that there is a degree of flexibility in design of studies supporting biosimilar development. In the context of a stepwise demonstration of biosimilarity and its implied flexibility, a developer and FDA may weigh the relative importance of various factors when sizing a comparative clinical study. Such factors to be considered may include, but are not limited to:

1. residual uncertainties after analytical testing;
2. the ability to detect clinically meaningful differences in effectiveness between the two products;

3. the ability to detect clinically meaningful differences in safety between the two products;
4. sufficient exposure to the proposed biosimilar product to detect safety signals, including immunogenic responses;
5. study benefits versus risks from a study subject's perspective.

Detection of Clinically Meaningful Differences

The clinical study to support a claim that no clinically meaningful differences exist between a proposed biosimilar and the reference product is a natural focus for developers. This may be in part due to the importance of clinical studies in demonstrating the effectiveness of new biological products. But this is in fact a point of contrast between the biosimilar pathway and that of a product being newly developed. As with generic drugs, the goal of the biosimilar exercise is to show the proposed product is very close to being the same as the reference product. Therefore, as with generic product evaluations, a different focus is placed on supporting clinical studies. Efficacy of biosimilars is inferred via their close similarity to the reference product. Once that is established, efficacy of the biosimilar is inferred from the reference product and its supporting clinical trials. The role of the clinical study in the biosimilar exercise is to resolve residual uncertainties about possible differences between the reference and the proposed biosimilar product given the information obtained in the analytical studies. Since the goal of the biosimilar comparative clinical study is not to demonstrate efficacy, but to show no difference in activity, the endpoint of the study and thus the parameter to be the subject of statistical testing could plausibly be different from that which was used to demonstrate efficacy of the reference product.

Given an appropriate endpoint, demonstrating that there are no clinically meaningful differences between the products can be broken into two pieces, showing the product has neither lesser nor greater activity compared to the reference product. Non-inferiority (NI) tests are commonly used to show that a product is no less effective than another. To support both parts of the goal, intersection-union tests, sometimes referred to as equivalence tests can be used. The basic idea of such a test is to show the difference between the two parameters being tested lies between pre-specified margins, customarily denoted by plus or minus δ. While it is common to use symmetric margins for this part of the biosimilarity exercise, the Scientific Considerations Guidance (U.S. Food and Drug Administration 2015b) suggests that given scientific justification, asymmetric margins could be an alternative. The remainder of our discussion will assume symmetric margins are used. The choice of δ is commonly based on a combination of historical data and clinical opinion. We let Δ denote the difference between the parameters and formally write the statistical null and alternative hypothesis as:

- Null Hypothesis: $\Delta \in \left\{(-\infty, -\delta] \bigcup [\delta, \infty)\right\}$
- Alternative Hypothesis: $\Delta \in (-\delta, \delta) \;=\; \left\{(-\infty, -\delta]^c \bigcap [\delta, \infty)^c\right\}$

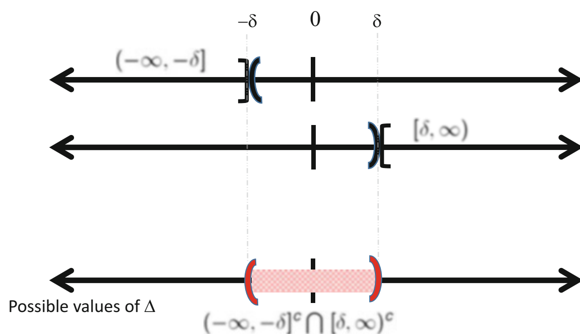


Fig. 18.1 The union of the sets labeled on the top two number lines comprise the set of outcomes under the null hypothesis. The intersection of their complements (bottom number line) is the set of outcomes under the alternative hypothesis

Figure 18.1 illustrates the hypotheses of an intersection-union test.

It is convenient to view intersection-union tests as two one sided tests (TOST) (Schuirmann 1987; Berger and Hsu 1996). For simplicity, we have chosen to discuss comparison in terms of differences. Comparisons in the context of TOST may also use other means of comparison, such as risk ratios or odd ratios. The third example presented later in this chapter is an example comparing outcomes via ratios of proportions. To execute TOST, two tests are performed, one having null hypothesis $\Delta \leq -\delta$, and the other having null hypothesis $\Delta \geq \delta$. The corresponding one-sided confidence intervals for Δ can be used to perform these tests. The TOST procedure is often executed by comparing the bounds of a two-sided confidence interval about the parameter Δ to the margins. For example the null hypothesis is rejected at the 5% level if the entire 90% confidence interval lies on the interval $(-\delta, \delta)$. The probability of a Type I error using this procedure is 5% which is intuitive and follows from the theory around closed testing procedures since both one-sided tests of the TOST must be rejected at the 5% level. A detailed discussion can be found in Berger and Hsu (1996).

Margins for clinical similarity studies are chosen such that rejecting the null hypotheses of the TOST procedure rules out clinically meaningful differences with a type I error probability of the chosen alpha level. Biosimilar developers generally select margins in consultation with the applicable regulatory agency because determining what is a clinically meaningful difference is subjective and opinions of regulatory stakeholders can play a role. Margin selection usually considers the effect size of the reference product because a difference which includes the possibility of the proposed biosimilar being no more effective than a placebo should be considered clinically meaningful. Therefore, a frequent starting point for margin selection is to estimate the effect size of the reference product through a meta-analysis of historical trials and use some function of that quantity as the margin. This approach parallels the concepts presented in FDA's guidance on non-inferiority (U.S. Food and Drug Administration 2016c) which gives examples of using 50% of the reference product effect size as a margin in NI studies for cardiovascular studies and margins of 10–15% of the reference product effect size for margins in antibiotic NI studies. The NI guidance states that a popular approach to NI testing is the so called 95–95%

or fixed margin approach. In this approach, a fraction—traditionally one half—of the more conservative bound of the 95% confidence interval of the reference product effect estimate is used for the NI margin. The appropriate bound of the 95% confidence interval of the difference parameter excluding the margin then rejects the NI null hypothesis.

While the margin selection procedures for biosimilar comparative clinical studies borrow from NI procedures, because the goals of the two designs are different, the biosimilar margin selection process has been more adaptable. If for example, the endpoint for a biosimilar study is different from that used in originally evaluating the reference product, the effect size could be difficult to estimate due to unavailability of historical data. Such a case might require more reliance on expert clinical opinion in margin selection. Situations in which historical trials are rare or the results vary also present motivation to explore different margin selection procedures.

As in any modeling procedure, the value of model output depends on the underlying assumptions. The validity of inference based on a TOST procedure indirectly depends on the assumptions surrounding clinically meaningfulness. The assumption that the medical practice and other pertinent conditions in place during the historical studies used to derive the margin are comparable to those conditions during the comparative clinical study is known as the constancy assumption (U.S. Food and Drug Administration 2016c).

Detection of Safety Signals

A clinical study may be necessary to resolve residual uncertainty with respect to possible safety signals associated with the proposed biosimilar product. A sample size can be determined based on an assumed event rate. For example, unobserved events have a rate of 0/n in a sample of n patients. We can determine sample size (n) by back calculating the upper bound of a 95% confidence interval. Assume we wish to rule out an event rate of 1/100 assuming we see no events. We calculated n such that the upper bound of the 95% confidence interval for 0/n is 0.01. Doing so provides a sample size estimate of approximately 300 (Jovanovic and Levy 1997) The duration study subjects are exposed to a proposed biosimilar product will depend on the nature of the product and the indication under study.

Examples

Following a stepwise approach, the type and depth of clinical studies needed to evaluate a specific proposed biosimilar product depend on how well the product is characterized at the time the comparative clinical study is designed. The following three examples illustrate slightly different approaches to selecting margins representing clinically meaningful differences between proposed biosimilar products and

licensed reference products. The first example illustrates a proposal that compares the margin to the historical effect size. As described in FDA's documentation (U.S. Food and Drug Administration 2016a), it appears that the developer used a combination of clinical opinion and historical data evaluation to select the margin. The second example shows the use of a surrogate endpoint and details considered in selecting the clinically meaningful margin. The final example uses a ratio comparison rather than a difference comparison and describes the margin in an unconventional way.

Etanercept

Etanercept is a tumor necrosis factor (TNF) blocker indicated for the treatment of: Rheumatoid Arthritis (RA), Polyarticular Juvenile Idiopathic Arthritis (JIA) in patients aged 2 years or older, Psoriatic Arthritis (PsA), Ankylosing Spondylitis (AS), and Plaque Psoriasis (PsO). The FDA's Arthritis Advisory Committee (AAC) met on July 13, 2016 to consider an application for a biosimilar to etanercept. The FDA briefing document provided to the AAC (U.S. Food and Drug Administration 2016a) gives a brief summary of the studies used to support the biosimilarity claim. What follows is a recap of a comparative clinical study used to support the assertion that there are no clinically meaningful differences between etanercept and the proposed biosimilar.

Our focus in these examples will highlight the approach to determining margins representing clinically meaningful differences. The etanercept comparative clinical study margin of plus or minus 18% was pre-specified by the developer and described as retaining 60% of the reference product effect size. This margin appears to have been chosen by considering clinical opinion and reference product effect size (U.S. Food and Drug Administration 2016a).

The study under discussion was a randomized, double blind comparative clinical study of the proposed biosimilar and the reference product in subjects age 18 years and older with chronic plaque psoriasis. The treatment group receiving the proposed biosimilar included 264 patients, and 267 were assigned to receive the reference product. The primary endpoint was the fraction of subjects in the respective treatment arms obtaining at least a 75% reduction from baseline in the Psoriasis Area Severity Index (PASI 75) after 12 weeks on the study. The margin representing a clinically meaningful difference was set at 18% based on historical data. Leonardi et al. (2003), reported PASI 75 of 49% for etanercept and 4% for placebo. Papp et al. (2005), reported PASI 75 of 49% for etanercept compared to 3% for placebo. Combining information from these studies resulted in an estimated effect size of approximately 45%. We note that the margin preserves approximately 60% of the estimated etanercept effect size. On study conclusion, the PASI 75 for the proposed biosimilar and the reference product were 70.5% and 71.5%, respectively; differing by −1.1% with a 90% confidence interval of (−8.3%, 6.0%) for the difference. The 90% confidence interval was wholely within the plus or minus 18% interval allowing rejection of the null hypothesis using the TOST procedure

described above. Of interest is the large discrepancy between the study results and the historical trials. The FDA noted in their briefing document (U.S. Food and Drug Administration 2016a) that patients in the historical trials were from The US, Canada and Western Europe, while those in the biosimilar comparative trial were mostly from Eastern Europe and South Africa. An observation such as this should signal to study designers that the constancy assumption may need re-evaluation. In this case the higher than expected responses did not show a loss of efficacy and were not considered to be a sign that the study's assay sensitivity was in jepardy (U.S. Food and Drug Administration 2016a).

Filgrastim

Filgrastim is a leukocyte growth factor most commonly used to decrease the likelihood of infections in patients receiving certain myelosuppressive or myeloablative chemotherapy treatments and to mobilize hematopoietic progenitor cells into the peripheral blood in preparation for leukapheresis. A precise description of the approved indications are available in the product label (U.S. Food and Drug Administration 2016b).

On January 7, 2015, the FDA Oncologic Drug Advisory Committee (ODAC) was asked to consider whether data submitted to the Agency supported the licensure of a product proposed to be biosimilar to filgrastim. We consider some details of one clinical study supporting the claim that are no clinically meaningful differences between the two products. A summary of this biosimilar exercise can be found in the FDA briefing document presented to the ODAC (U.S. Food and Drug Administration 2015a).

The study under discussion was a double-blind parallel group comparative study in which women with breast cancer who were eligible for neoadjuvant or adjuvant treatment were treated with six cycles of chemotherapy that included docetaxel at 75 mg/m^2, doxorubicine at 50 mg/m^2 and cyclophosphamide at 500 mg/m^2. The chemotherapy regimen, also called TAC and known to cause neutropenia, was given intravenously on the first day of each 21-day cycle. Twenty-four hours after chemotherapy each woman also received according to random assignment either the proposed biosimilar or the reference product. The primary endpoint was the duration of severe neutropenia (DSN), defined as an absolute neutrophil count less than 500 per micro-liter. We note that DSN was used as a surrogate for the clinical endpoint febrile neutropenia (FN) used in the original evaluation of filgrastim (U.S. Food and Drug Administration 1998a). Studying the difference between the two products with respect to FN would have required a much larger sample size, as it has been demonstrated that use of filgrastim lowers the incidence of FN and a 1 day change in DSN correlated to a 10% change in the risk of FN (Blackwell and Crawford 1994). This relationship between the DSN and FN was a factor in determining the margin representing a clinically meaningful difference. A 10% difference in FN may be considered to be a clinically significant as discussed in FDA's medical review of peg-filgrastim (U.S. Food and Drug Administration 2002a), thus the choice of 1 day

DSN as the margin representing a clinically meaningful difference. The margin should be compared to the reference product's effect size to ensure the study's ability to detect a treatment no better than placebo. Nabholtz et al. (2001) report median DSN for patients receiving TAC but not receiving G-CSF prophylaxis to be approximately 7 days. In FDA's review of peg-filgrastim (U.S. Food and Drug Administration 2002b), the mean DSN for patients receiving TAC and filgrastim was reported to be approximately 1.6 days. We observe that a roughly estimated effect size of filgrastim for patients with breast cancer and receiving TAC is approximately 4–5 days difference in DSN. The margin of 1 day being less than one quarter of the filgrastim effect size is additional support for its adoption. Nie et al. (2018) give a detailed discussion of the 1 day margin and its relationship to the effect size estimated from historical data.

The outcome of the 204 person study showed the difference in mean DSN between the reference product group and the test product group to be 0.04 days DSN with 90% confidence interval (−0.21, 0.28). The confidence interval falling within the zero plus or minus 1 day margins allows us to use TOST to reject the null hypothesis (section "Detection of Clinically Meaningful Differences") that the difference in mean DSN lies outside the interval defined by the clinically meaningful margin with a type I error probability of 5%.

Bevacizumab

The third example reviews a study presented at the Oncologic Drug Advisory Committee meeting of July 13, 2017 on a proposed biosimilar to bevacizumab. The reference product, bevacizumab is a humanized monoclonal antibody designed to inhibit tumor angiogenesis by targeting vascular endothelial growth factor (VEGF) (Semenza 2008). As with the other examples presented, this example focuses on one comparative clinical study designed to support the claim that there are no clinically meaningful differences between the proposed biosimilar and the reference products.

The comparative clinical study, conducted in a population of patients having non-small cell lung cancer (NSCLC) receiving first-line therapy with carboplatin and paclitaxel, was randomized and double-blinded. A total of 642 patients were randomly assigned in a 1:1 ratio to either the reference product (314) or the proposed biosimilar (328). The study was designed to enroll 620 patients, equally allocated to the two treatment arms such that it would have more than 95% power to demonstrate no clinically meaningful difference existed between treatments with respect to overall response rate (ORR). The general approach used to compare the treatments is the one discussed in section "Detection of Clinically Meaningful Differences" above. An interesting difference between this example and the previous two is the comparison of treatment outcomes in this example is via a ratio. The statistical test proposed by the biosimilar developer was such that the ratio of ORRs would need to be between 0.67 and 1.5 to support the claim that the two products were not clinically meaningfully different. The FDA did not agree with the proposed margins (U.S. Food and Drug Administration 2017b) and used alternatives in the evaluation of the data.

As mentioned before, calculations of margins used in biosimilar studies parallel those used in non-inferiority (NI) studies. The latter study designs are often used to show that one treatment is not worse than another with some allowance for variability set by the choice of the NI margin. In drug trials, a major focus of NI study design is to ensure the study treatment is adequately better than a placebo through comparison to a drug known to be active in the setting under study. This consideration is often a starting point for selecting margins in the biosimilar setting with the FDA's guidance on non-inferiority (U.S. Food and Drug Administration 2016c) used as an aid. The most common implementation of the approach can be summarized in two steps: (1) estimate the effect size of the reference product and (2) choose some fraction of that effect size that ensures rejecting the null hypothesis does not permit concluding an ineffective product is effective. The fixed margin approach accounts for variability in the historical effect size estimate by using a function of the confidence bound as a margin. This is only partially satisfying because the fraction of the effect size, based on a confidence interval or not, is still a subjective choice. We therefore understand that although the approach considers variability in the algorithm, the margin selection process is a subjective benefit-risk evaluation that considers variability. This is especially important if one is executing a hypothesis test using a p-value because that value will be a function of the margin.

He et al. (2016) performed the two margin development steps just described for bevacizumab biosimilar studies in patients having NSCLC. The effect size estimation included four published clinical trials in that population: Sandler et al. (2006), Nishio et al. (2009), Johnson et al. (2004), and Reck et al. (2010). The ratios of ORR (control ORR)/(control + bevacizumab ORR) from the four studies were estimated to be, respectively: 0.43, 0.60, 0.58, and 0.63. He et al. (2016), using a fixed effect meta analysis, the reproduced results of which are partially presented in Fig. 18.2, estimated the combined ORR ratio to be 0.53, 95%CI (0.45, 0.63).

Each trial in the meta-analysis was a comparison of a chemotherapy regimen plus bevacizumab to the chemotherapy regimen alone. Except for the study reported by Reck et al. in which the chemotherapy regimen was cisplatin and gemcitabine, the chemotherapy regimens were paclitaxel and carboplatin. We notice from Fig. 18.2 that the point estimate from this study (Reck et al. 2010) is consistent with two of the three other studies and it is that of the largest study that deviates slightly.

As in this case, ratios are commonly used in oncology settings to compare treatments. This is natural in a time-to-event setting, but ratios also tend to be preferred when comparing event rates in oncology studies. It is important to consider how relative comparisons impact equivalence margins. For given margins on the ratio scale, the interval width of acceptable response rates for a proposed biosimilar product will vary with the assumed response rate of the reference product. Consider for example, the commonly used bioequivalence margins of 0.80–1.25. When designing the study, assuming the reference product response rate is 10% permits response rates of 8–12.5% for the proposed biosimilar. If one were to assume a different reference product response rate, say 50%, for example, the proposed biosimilar product could vary by more (40–72.5%). If a ratio comparison is used and the constancy assumption is invalid such that the observed event rate is smaller than expected, the study is in danger of being underpowered.

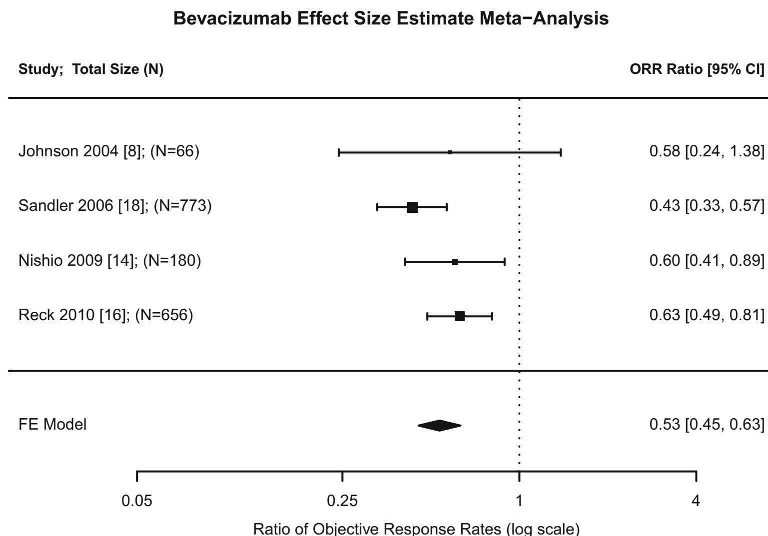


Fig. 18.2 The bevacizumab effect size presented here was estimated using a fixed effect meta-analysis, as in He et al. (2016)

Feasibility of a study depends on its size. He et al. (2016) and the FDA briefing document (U.S. Food and Drug Administration 2017b) present possible study sizes in the context of different similarity margins and confidence levels of the bevacizumab effect size. To help illustrate the approach we consider a simplified sample size formula assuming normality (18.1).

$$n = \frac{C * \left(Z_{\alpha} + Z_{1-\beta}\right)^2 \sigma^2}{\psi^2} \tag{18.1}$$

In this expression, C represents constant terms used in specific models. The influence of type I error probability and power, represented by α and $(1 - \beta)$ respectively, are well known factors impacting sample size. Sample size increases as lower type I error and/or higher power are demanded. Large study data variation, represented by (σ) requires a relatively larger sample size. In the bevacizumab example, as with other examples having binary outcomes, the assumed variance of the ORR ratio depends on the assumed effect size. The final term in the sample size calculation is the desired difference to be detected, represented here by ψ. Requiring the two products to differ very little requires a small value of ψ and in turn a large sample size. In biosimilar examples, ψ includes the margin representing clinically meaningful differences which in this case is described as being half of the more conservative effect size confidence bound. Recall that while 1/2 is commonly used, it is nonetheless determined subjectively. This approach is a variation of the fixed margin approach in which fractions of the effect size 95% confidence bound are

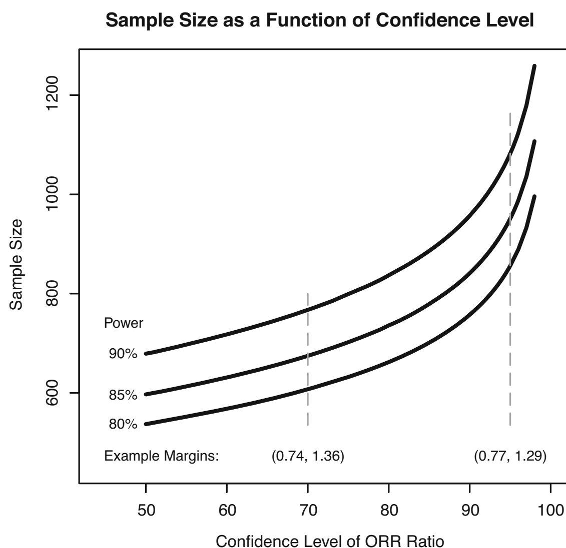


Fig. 18.3 Study size is presented as a function of confidence level. Sample sizes depend on the clinically meaningful margins which are based on 1/2 of the effect size lower confidence bound

considered for the margin. Here the fraction 1/2 is held constant and the effect size confidence level is varied. Figure 18.3 shows sample size as a function of effect size confidence level at three power levels. Type I error probability is held constant at 0.05. The margins FDA considered to be clinically meaningful and those depending on 1/2 of the 95% confidence bound are included on the plot for easy reference. These sample size calculations are based on equation 11.19 of Rothmann et al. (2012) and include an allowance for 10% attrition due to dropouts.

The FDA determined that the interval of ORR ratios (0.74, 1.36) represented no clinically meaningful difference in this particular comparison of bevacizumab and the proposed biosimilar. The margin selection was based on a benefit-risk analysis considering power, type I error, sample size, knowledge of historical data, preliminary analytical studies of the proposed product, and confidence in the effect size estimate. The ORR results were 39% and 41.7% for the proposed biosimilar and the reference product, respectively. The ratio of ORRs was 0.93 with 90% CI (0.80, 1.09). The 90% CI falling completely within the margins 0.74 and 1.36 allowed rejecting the two null hypotheses in TOST and support the claim that there are no clinically meaningful differences between the products.

Interchangeability

The final piece of the stepwise approach we consider is interchangeability. A biological product may be determined to be interchangeable with a reference product if it is (1) biosimilar to the reference product and (2) expected to produce

the same clinical result as the reference product in any given patient, and also (3) not to pose excessive risks to patients if they switch between the reference product and interchangeable product (42 USC 262 2010; U.S. Food and Drug Administration 2015b, 2017a). A practical distinction between biosimilar products and interchangeable products is that an interchangeable product may be substituted for the reference product without the intervention of the product's prescriber. That is, the interchangeable product may be given in place of the reference product at the pharmacy level (42 USC 262 2010).

The draft guidance document, Considerations in Demonstrating Interchangeability With a Reference Product (Interchangeability Draft Guidance) (U.S. Food and Drug Administration 2017a) again promotes the use of a stepwise approach in developing interchangeable products by suggesting that postmarketing information on a biosimilar product could be used as partial support for claims the product is also interchangeable. However, the guidance goes on to state that postmarketing data alone would likely not suffice to support interchangeability claims and that a dedicated study to support interchangeability would usually be needed.

The FDA expects immunogenicity related outcomes to be the primary risk incurred by patients switching back and forth between the reference product and an interchangeable product (U.S. Food and Drug Administration 2017a). Therefore because pharmakometric (PK) and pharmacodynamic (PD) endpoints are believed to be sensitive to immunogenicity changes they should be the focus of the switching study. The Interchangeability Draft Guidance (U.S. Food and Drug Administration 2017a) suggests that a switching study be a randomized comparison of a switching arm vs a non-switching arm with three changes from one product to the other; the first product administered being the reference product. Equivalence tests with PK and PD endpoints are discussed elsewhere in this volume.

Discussion

We discussed the origins and some important aspects of FDA's suggested stepwise approach to evaluating biosimilar products in sections one and two. An important feature of the approach is the idea of considering the evidence supporting biosimilarity assertions in its totality, rather than relying on pivotal trials for decision making purposes. Such an approach affords stakeholders, the product developers and FDA flexibility in using risk based design of experiments and evaluation criteria so that an adequate evaluation may be accomplished while not departing from the goal of having an abbreviated pathway. Sections three through five discuss component studies of the evaluation process, but focus on design elements of comparative clinical studies. The examples illustrate design flexibility. In the first example, the margins representing a clinically meaningful difference were obtained by considering the effect size of the reference product, as well as clinical opinion. The filgrastim example is an example in which a surrogate endpoint was used in comparing the proposed biosimilar to the reference product rather than using that

of the original approval of the reference product. The use of the surrogate allowed designing a study that adequately compared the two products without needing a large sample size. The final example, in which a biosimilar to bevacizumab was compared to the reference product, illustrates the advantages of comparing products through a ratio when uncertainty exists about the reference product effect size.

Acknowledgements I express my gratitude to Dr. Huanyu Chen for programming advice and assistance.

References

42 USC §262. Regulation of biological products (2010). http://uscode.house.gov/view.xhtml?req=(title:42%20section:262%20edition:prelim)#sourcecredit. Accessed 1 Oct 2017.

Berger RL, Hsu JC. Bioequivalence trials, intersection-union tests and equivalence confidence sets. Stat. Sci. 1996;11:283–319.

Blackwell S, Crawford J. Filgrastim (r-methug-csf) in the chemotherapy setting. In: Morstyn G, Dexter TM, editors. Filgrastim (r-metHuG-CSF) in clinical practice. New York: Marcel Dekker; 1994. pp. 103–116.

Christl LA, Woodcock J, Kozlowski S. Biosimilars: the us regulatory framework. Annu. Rev. Med. 2017;68:243–254.

Code of Federal Regulations. 21cfr314.92, current as of April 1 2017. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=314; 2017a. Accessed 29 Dec 2017.

Code of Federal Regulations. 21cfr320, current as of April 1 2017. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=320; 2017b. Accessed 29 Dec 2017.

He K, Chen H, Gwise T, Casak S, Lemery S, Keegan P, Pazdur R, Sridhara R. Statistical considerations in evaluating a biosimilar product in an oncology clinical study. Clin. Cancer Res. 2016;22:5167–5170.

Johnson DH, Fehrenbacher L, Novotny WF, Herbst RS, Nemunaitis JJ, Jablons DM, Langer CJ, DeVore RF III, Gaudreault J, Damico LA, Holmgren E, Kabbinavar F. Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer. J. Clin. Oncol. 2004;22 11:2184–2191. PMID: 15169807.

Jovanovic BD, Levy PS. A look at the rule of three. Am. Stat. 1997;51 2:137–139.

Leonardi CL, Powers JL, Matheson RT, Goffe BS, Zitnik R, Wang A, Gottlieb AB. Etanercept as monotherapy in patients with psoriasis. N. Engl. J. Med. 2003;349 21:2014–2022.

Mossinghoff GJ. Overview of the hatch-waxman act and its impact on the drug development process. Food Drug Law J. 1999;54:187–194.

Nabholtz JM, Mackey JR, Smylie M, Paterson A, Noël DR, Al-Tweigeri T, Tonkin K, North S, Azli N, Riva A. Phase II study of docetaxel, doxorubicin, and cyclophosphamide as first-line chemotherapy for metastatic breast cancer. J. Clin. Oncol. 2001;19:314–321.

Nie L, Przepiorka D, Deisseroth A, Sridhara R, Gwise TE. Determination of similarity margin in comparative clinical studies to support the development of biosimilar products of neupogen. Bio Drugs 2018;32:325–330.

Nishio M, Horai T, Kunitoh H, Ichinose Y, Nishiwaki Y, Hida T, Yamamoto N, Kawahara M, Saijo N, Fukuoka M, JO19907 Study Group. Randomized, open-label, multicenter phase ii study of bevacizumab in combination with carboplatin and paclitaxel in chemotherapy-naive Japanese patients with advanced or recurrent nonsquamous non-small cell lung cancer (nsclc): Jo19907. J. Clin. Oncol. 2009;27. 15s abstract 8036.

Papp KA, Tyring S, Lahfa M, Prinz J, Griffiths CEM, Nakanishi AM, Zitnik R, Van De Kerkhof PCM, the Etanercept Psoriasis Study Group. A global phase III randomized controlled trial of etanercept in psoriasis: safety, efficacy, and effect of dose reduction. Br. J. Dermatol. 2005;152 6:1304–1312.

Reck M, von Pawel J, Zatloukal P, Ramlau R, Gorbounova V, Hirsh V, Leighl N, Mezger J, Archer V, Moore N, Manegold C. Overall survival with cisplatin-gemcitabine and bevacizumab or placebo as first-line therapy for nonsquamous non-small-cell lung cancer: results from a randomised phase III trial (avail). Ann. Oncol. 2010;21 9:1804–1809.

Rothmann MD, Wiens BL, Chan ISF. Design and analysis of non-inferiority trials. Boca Raton: Chapman & Hall/CRC; 2012.

Sandler A, Gray R, Perry MC, Brahmer J, Schiller JH, Dowlati A, Lilenbaum R, Johnson DH. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N. Engl. J. Med. 2006;355 24:2542–2550. PMID: 17167137.

Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J. Pharmacokinet. Biopharm. 1987;15:657–680.

Semenza GL. A new weapon for attacking tumor blood vessels. N. Engl. J. Med. 2008;358 19:2066–2067. PMID: 18463385.

U.S. Food and Drug Administration. Filgrastim. https://www.accessdata.fda.gov/drugsatfda_docs/label/1998/filgamg040298lb.pdf; 1998a. Accessed 25 Feb 2018.

U.S. Food and Drug Administration. Providing clinical evidence of effectiveness for human drug and biological products; guidance for industry. https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm072008.pdf; 1998b. Accessed 29 Dec 2017.

U.S. Food and Drug Administration. Application number 125031 medical review. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2002/125031_000_Neulasta_medr_P3.pdf; 2002a. Accessed 07 Jan 2018.

U.S. Food and Drug Administration. Application number 125031 statistical review. http://www.accessdata.fda.gov/drugsatfda_docs/nda/2002/125031_000_Neulasta_statr.pdf; 2002b. Accessed 07 Jan 2018.

U.S. Food and Drug Administration. FDA briefing document oncologic drugs advisory committee meeting January 7, 2015 bla 125553 ep2006, a proposed biosimilar to filgrastim. https://wayback.archive-it.org/7993/20170405222922/https:/www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM428780.pdf; 2015a. Accessed 06 Jan 2018.

U.S. Food and Drug Administration. Scientific considerations in demonstrating biosimilarity to a reference product. www.fda.gov/downloads/drugs/guidances/ucm291128.pdf; 2015b. Accessed 07 Jan 2018.

U.S. Food and Drug Administration. FDA briefing document arthritis advisory committee meeting July 13, 2016, bla 761042 gp2015, a proposed biosimilar to enbrel (etanercept). https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM510493.pdf; 2016a. Accessed 07 Jan 2018.

U.S. Food and Drug Administration. Filgrastim. https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/103353s5188.pdf; 2016b. Accessed 06 Jan 2018.

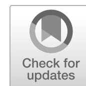
U.S. Food and Drug Administration. Non-inferiority clinical trials to establish effectiveness, guidance for industry. https://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf; 2016c. Accessed 07 Jan 2018.

U.S. Food and Drug Administration. Considerations in demonstrating interchangeability with a reference product; guidance for industry (draft guidance). https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf; 2017a. Accessed 18 Jan 2018.

U.S. Food and Drug Administration. FDA briefing document oncologic drugs advisory committee July 13, 2017 bla 761028 abp215, a proposed biosimilar to avastin (bevacizumab). https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM566365.pdf; 2017b. Accessed 07 Jan 2018.

Woodcock J, Griffin J, Behrman R, Cherney B, Crescenzi T, Fraser B, Hixon D, Joneckis C, Kozlowski S, Rosenberg A, Schrager L, Shacter E, Temple R, Webber K, Winkle H. The FDA's assessment of follow-on protein products: a historical perspective. Nat. Rev. Drug Discov. 2007;6:437–442.

Chapter 19
Immunogenicity Assessment of Biosimilars: A Multidisciplinary Perspective



Paul Chamberlain and Pekka Kurki

Abstract Evaluation of the relative immunogenicity of a biosimilar candidate in direct comparison to the reference product is a general regulatory requirement, with the main weight of evidence deriving from head-to-head clinical studies in populations that are adequately sensitive to reveal clinically meaningful differences across the proposed conditions of use.

This chapter provides case examples to illustrate the multi-disciplinary nature of the exercise, which involves interpretation of bioanalytical measures of the immune response in relation to differences detected in the product quality profile, as well as potential biases in bioanalytical methodology and confounding patient-related factors. Most importantly, the design of the clinical immunogenicity evaluation needs to reflect the risk profile established for the reference product, allied to uncertainty about possible impact of minor heterogeneity in product-related variants and process-derived impurities, and limitations of the methodology for detection of clinically meaningful consequences. Thus, a difference in the measured anti-drug antibody incidence may not necessarily preclude a conclusion of biosimilarity if this does not translate into a negative impact on efficacy or safety: examples of authorized biosimilar products are discussed to show how regulators have applied a "totality of evidence" approach to deal with apparent numerical differences in immune response parameters.

Validity of extrapolating conclusions about relative immunogenicity in one therapeutic setting to other indications is reviewed based on actual clinical results obtained for products associated with clinically impactful immunogenicity. Then, experience gained from transitioning between product versions is discussed in the context of interchangeability considerations. Finally, reflection is given to

---

P. Chamberlain (✉)
NDA Advisory Services Ltd, Leatherhead, Surrey, UK
e-mail: paul.chamberlain@ndareg.com

P. Kurki
Clinical Immunology, University of Helsinki, Helsinki, Finland
e-mail: pekka.kurki@fimea.fi

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_19

longer-term management of potential immunogenicity-related risks associated with manufacturing changes during the independent post-authorization life-cycles of different product versions.

Keywords Immunogenicity · Biosimilars · Aggregates · Extrapolation · Switching · Interchangeability · Substitution

Introduction

Immunogenicity is the host immune response to administration of a therapeutic agent. The immune response includes cell-mediated (innate and adaptive effector/regulatory cells) and humoral elements (antibodies), pre-existing and treatment-emergent effects. Immunogenicity of therapeutic proteins may result in harmful local and systemic manifestations, and can be associated with diminished drug exposure and efficacy. However, for the majority of biological medicinal products—whether they be innovator or biosimilar versions—immunogenicity has no impact on overall clinical benefit and risk.

Assessment of the relative immunogenicity of a biosimilar candidate to the reference product is an essential element of the biosimilarity exercise, whose objective is to demonstrate the absence of a clinically meaningful difference in immunogenicity between the comparator products. This is a multi-disciplinary exercise that involves consideration of:

- Immunogenicity-related risks identified for the reference product and potential risks associated with the biosimilar candidate product
- Qualitative and quantitative differences in product-related variants and process-derived impurities
- Stability of the drug product in the formulation-primary container to be commercialized
- Potential for reactivity with pre-existing antibodies, and dynamics of treatment-emergent immune response, in suitably sensitive clinical populations
- Interpretation of bioanalytical measures of the immune response in relation to PK/PD, efficacy and safety parameters
- Impact on overall clinical benefit and risk for all populations to be treated.

Regulatory authorities apply a risk-based approach to the assessment of relative immunogenicity (FDA 2014; EMA 2006), reflecting a hierarchy of concerns driven by severity of consequences (Rosenberg 2003). While the same principles apply to authorization of innovator and biosimilar candidates, the scale of evaluation for biosimilars can be guided by the clinical experience gained for the reference product, taking into account uncertainty about product-specific variables that could result in incremental immunogenicity. Since it not possible to predict how these variables may interact to modify the balance between innate and adaptive immune responses and immune tolerance in individual human subjects, the current regulatory standard in EU and US requires directly comparative clinical evaluation of immunogenicity

of biosimilar and reference products prior to marketing authorization. The purpose of this chapter is to illustrate the extent of the clinical evaluation of immunogenicity for different product types, based on experience of the EU and US regulatory approval processes.

Immune Response to Therapeutic Proteins

Endogenous and therapeutic proteins contain structural motifs (T- and B-cell epitopes) that drive the adaptive immune response by stimulating antigen-specific effector and regulatory T-lymphocytes, and B-lymphocytes (de Groot and Scott 2007; Weber et al. 2009; Baker et al. 2017). The T-dependent response is mediated by structurally-constrained binding of the peptide sequences, which have been taken up and processed by antigen-presenting cells, to intracellular MHC Class II proteins; the peptide-MHC Class II complexes can then migrate to the cell surface, which enable the T-cell epitope to be recognized by the T-cell receptor on CD4+ T-helper cells.

Stimulation of an effective immune response depends on co-stimulatory factors provided by antigen-presenting cells, including B-lymphocytes, following recognition by B-cell receptors of antigen-specific structural motifs, referred to as B-cell epitopes, in the endogenous or therapeutic protein. B-cell epitopes can be overlapping with, or distinct from, the amino acid sequence corresponding to the T-cell epitopes in the protein.

Endogenous proteins such as erythropoietin are intrinsically immunogenic by virtue of peptide sequences that can bind to MHC Class II (Tangri et al. 2005), as evidenced by detection of erythropoietin-specific CD4+ T-cells in the human circulation (Delluc et al. 2010). Central and peripheral immune tolerance mechanisms then act to suppress immune responsiveness to endogenous ("self") proteins—although the level of tolerance differs between the B- and T-cell populations, and may be bypassed via excessive stimulation by therapeutic versions of the protein, particularly if these contain altered conformers or aggregates of the therapeutic protein (Sauerborn et al. 2010).

Given the highly polymorphic nature of human MHC Class II proteins, allied to restriction of the T-dependent response to productive presentation of T-cell epitopes by MHC Class II, the immune response can be related to the particular MHC Class II allelic variants expressed by individual subjects (Stickler et al. 2004). In addition to the impact of genetic polymorphism on T-dependent responses, post-translational modifications of proteins can alter recognition by B-cell receptors (Baker et al. 2017).

Activation of innate immune cells (e.g. macrophages and dendritic cells) by process-derived impurities can provide an important source of co-stimulatory signals to enhance the adaptive immune response to a therapeutic protein (Verthelyi and Wang 2010). These cells have also been shown to be stimulated by the presence of aggregates and sub-visible particles induced by mechanical or thermal stress of

therapeutic monoclonal antibodies (Rombach-Riegraf et al. 2014; Ahmadi et al. 2015; Joubert et al. 2016).

Finally, truncation of protein molecules could expose neoepitopes in a manner that alters recognition by pre-existing antibodies (Brezski et al. 2011; van Schie et al. 2015a; Kim et al. 2016) and/or induction of a treatment-emergent response (Li et al. 2001). Therefore, high quality, purity, and stability are all key elements in the development of biosimilars.

Differential Factors for Biosimilar Candidates

Product Quality Variables with Potential to Influence Immunogenicity

The analytical similarity exercise places high emphasis on demonstration that the biosimilar candidate has an identical primary amino acid sequence to the reference product, and that any minor differences in product related variants are not expected to have a clinical impact. State-of-art physicochemical and biological characterization of multiples batches of the biosimilar candidate are compared directly to multiple batches of the reference product to establish that the respective analytical profiles are comparable; non-clinical and clinical comparison may then be performed to confirm absence of impact of detectable (and undetectable) differences in pharmacological and therapeutic properties, including immunogenicity.

As discussed above, therapeutic proteins could contain different levels of product-related variants and/or process-derived impurities that have potential to modify recognition by the different elements of the innate and adaptive immune systems. Thus, even if the primary amino acid sequence of different versions of the same therapeutic protein are identical—implying equivalence of intrinsic immunogenic potential—there could be differences in the immune response to drug products associated with different levels of:

- Molecular size variants, including truncated or fragmented protein, oligomers, aggregates or sub-visible particles;
- Host-cell derived impurities including lipopolysaccharides and non-human proteins;
- Post-translational glycosylation, particularly the qualitative and quantitative content of non-human glycans.

While any these factors might influence the immune response to an exogenous protein, as well as interacting with different factors to enhance risk, these factors are detectable using state-of-art analytical methods and their levels are controlled within defined limits. *In addition*, comparative (biosimilar *vs*. reference product) stability testing, designed to model “worst-case” storage and handling conditions, is performed as part of the similarity exercise to evaluate a potential influence of

differences in the formulation-primary container combination of the drug product to be commercialized (Pisupati et al. 2017).

Although it is evident that deamidation of T-dependent immunogens can modify T-cell help to reduce immunogenicity of a vaccine (Verma et al. 2016), many proteins undergo deamidation in vivo as a naturally occurring process (Liu et al. 2009), making it difficult to evaluate impact of small differences of deamidated variants that may be present in the drug product prior to administration. Oxidation of amino acid residues in therapeutic proteins is also commonly observed, and can be associated with conformational changes that can lead to aggregation of the protein (Torosantucci et al. 2014). The methods applied during the analytical similarity exercise are expected to define the relative levels of these product-related variants in the biosimilar version and reference product, as well as their potential impact on biological potency. In the case that significant analytical differences were detected in these parameters, results from comparative clinical studies would then assist the interpretation of clinical impact.

Both FDA (https://www.fda.gov) and EMA (http://www.ema.europa.eu/ema/) publish detailed information describing how the analytical results were assessed in relation to immunogenicity for approved biosimilar products, and multiple publications present comparative analytical profiles for the biosimilar versus reference versions (Visser et al. 2013; Jung et al. 2014; Cho et al. 2016; Liu et al. 2016; Lee et al. 2018; Seo et al. 2018). Case examples using published information for adalimumab and infliximab are presented later in this chapter to illustrate the interdependencies between the clinical evaluation and the product quality review for the assessment of relative immunogenicity of the biosimilar candidate and reference products.

Mitigation of Risk due HMW Variants and Sub-visible Particles

Clinical data to define the relationship between levels of high molecular weight variants, aggregates or sub-visible particles to the immunogenicity of biological medicinal products are not available. This reflects the limitations of analytical and bioanalytical methods applied historically (Carpenter et al. 2009; Wang et al. 2012). Improved methodology (Hawe et al. 2012; Weinbuch et al. 2013) is now available to characterize the levels of HMW variants and sub-visible particles, and more sensitive, drug-tolerant assays, can be applied to monitor the immune response in subjects treated in clinical trials (Bourdage et al. 2007; Smith et al. 2007; Bloem et al. 2015).

In the case of therapeutic monoclonal antibodies, it is known that dimerization or oligomerization can enhance binding to Fcγ receptors by some 100-fold (Luo et al. 2009), and that Fcγ receptors appear to play a role in the activation of innate immune effector cells by aggregated IgG (Joubert et al. 2016). Theoretically, both dimerization and aggregation could contribute co-stimulatory signals that enhance an adaptive immune response to a therapeutic monoclonal antibody—the scale of

such enhancement is likely to depend on the intrinsic immunogenicity associated with T- and B-cell epitopes in the monomeric molecule. Therefore, application of orthogonal analytical techniques is required by regulatory authorities to control levels of product-related size variants in the size ranges (100 nm to 10 μm diameter) that were not effectively monitored by traditional analytical methods such as SE-HPLC or light obscuration (Carpenter et al. 2009).

The multi-disciplinary team approach applied by regulatory authorities for the analytical similarity exercise includes a rigorous review of product quality variables that have a potential impact on immunogenicity, with particular emphasis on aggregates, sub-visible particles and process-derived impurities.

Inferences

- Demonstration of analytical similarity is the basis for minimizing risk of incremental immunogenicity of biosimilar versions relative to the reference product.
- Although multiple extrinsic factors *might* influence immunogenicity of different therapeutic proteins, levels of the individual factors are monitored by state-of-art analytical methods. By definition and in practice, designation of a product as a biosimilar is based on sound evidence that only minor differences in product quality attributes are detected, and that such differences do not have a clinically meaningful impact.
- Potential for an influence of differences that may not detected by analytical methods, or which has an uncertain impact immunogenicity, is evaluated in randomized, controlled, comparative clinical studies prior to authorization of biosimilar candidates; with regulatory provision for post-marketing monitoring adapted to the risk profile of the product and level of residual uncertainty associated with detectability of the risk.

Role of Non-clinical Evaluation

The regulatory assessment of the first biosimilar products to be approved in EU included a review of the relative ADA response detected in comparative non-clinical toxicology studies (Chamberlain 2014). However, the results were not instructive for the biosimilarity assessment due to the doubtful relevance of the immune response in non-human species for detection of clinically meaningful differences between product versions meeting the high standard of analytical similarity required for approval as a biosimilar. EU regulatory guidance (EMA 2014) was then revised actively to discourage use of animals for the biosimilarity exercise. EU

regulators (Van Aerts et al. 2014) concluded “that animal studies have no role in a comparability exercise to establish the biosimilarity of a product with regard to its propensity to induce ADA in humans”. The authors of the present chapter strongly agree with this conclusion, not least because it would be unethical to perform such studies without a sound scientific justification.

Clinical Study Design Considerations for the Biosimilarity Exercise

Health Authority Recommendations

EU and FDA guidance (EMA 2014; FDA 2015b; EMA 2010) recommend use of a comparative parallel-group design study in treatment-naïve patients as the most sensitive pre-authorization test to evaluate potential differences in clinical impactful immunogenicity between the biosimilar candidate and the reference product. Data from comparative PK/PD and therapeutic studies are relevant, particularly if these use different populations, i.e. healthy volunteers for comparator PK/PD, and patients for the therapeutic study. Although not usually required in EU,[1] FDA guidance states that “depending on the clinical experience of the reference and proposed products (taking into consideration the conditions of use and patient population), a sponsor may need to evaluate a subset of patients to provide a substantive descriptive assessment of whether a single cross-over from the reference product to the proposed biosimilar would result in a major risk in terms of hypersensitivity, immunogenicity, or other reactions.”

If there is uncertainty about the relative immunogenicity risk profiles for the proposed therapeutic indications at the time of marketing authorization, post-authorization studies may be requested; these might be interventional or observational, depending on the risk profile of the product and the scale if uncertainty.

To date, one biosimilar product (teriparatide, MOVYMIA) was approved in the EU without a comparative clinical evaluation of immunogenicity; the decision was driven by the low risk profile for the originator product allied to a commitment from the Applicant to submit forthcoming results from a randomized controlled study that had been planned to support marketing authorization in another region (EMA 2016b). This last case illustrates the flexibility of regulatory authorities to accept alternative approaches for addressing immunogenicity-related risks, based on an objective analysis of the scale of risk for the particular product.

[1]One exception is the case of a biosimilar version of a recombinant erythropoietin, for which a switch from the reference product to the biosimilar is recommended (EMA 2010).

Risk Profile for Reference Product

Wide Range of Clinical Impact of Immunogenicity for Different Products

Understanding the immunogenicity risk profile of the reference product represents the starting point for the evaluation of relative immunogenicity of the biosimilar versus reference products. This reflects the wide range of undesirable clinical consequences observed for the innovator versions during controlled studies performed at the pre- and post-authorisation stages, supported by ongoing pharmacovigilance: for many therapeutic protein-based products, there are no identified risks associated with immunogenicity; while severe consequences have been observed for some products (Table 19.1). Case examples are presented below to exemplify how the scale of risk should be considered on a product-by-product basis for the biosimilarity exercise.

Table 19.1 Identified immunogenicity-related risks for reference products

Product	Clinical impact of immunogenicity/ADA	Published references
Epoetin alfa	Cross-reactive neutralizing ADAs causing amPRCA (rare)	Casadevall et al. (1996)
Darbepoetin	Cross-reactive neutralizing ADAs causing amPRCA (rare)	Macdougall et al. (2015)
Cetuximab	Severe allergic reactions in pre-sensitized subjects	Chung et al. (2008)
Natalizumab	Loss of efficacy and increased incidence of infusion-related reactions	Subramanyam (2008)
Infliximab	Immune complex-related hypersensitivity and loss of efficacy	Beart et al. (2003) and Bendtzen et al. (2006)
Adalimumab	Loss of efficacy and increased incidence of injection site reactions	Bartelds et al. (2007) and Murdaca et al. (2016)
Rituximab	Loss of efficacy in patients with severe pemphigus and rare cases of hypersensitivity reactions	Schmidt et al. (2009) and Ataca et al. (2015)
Somatropin	Possible reduction in PK/PD/efficacy in very rare cases	Pfizer (2016)
Insulin glargine	Possible reduction in PK/PD/efficacy in very rare cases	Fineberg et al. (2007)
Follitropin-alfa	Negative impact not identified	Loumaye et al. (1998)
Bevacizumab	Negative impact not identified	EMA (2005a)
Trastuzumab	Negative impact not identified	EMA (2005b)
Abatacept	Negative impact not identified	BMS (2017)
Omalizumab	Negative impact not identified	Somerville et al. (2014)
Filgrastim	Negative impact not identified	Amgen (2016a)
Pegfilgrastim	Negative impact not identified	Amgen (2017)

ADA anti-drug antibody
amPRCA antibody-mediated pure red cell aplasia

It is important to understand the limitations of historical data describing ADA responses to originator products, because in many cases the ADA responses may have been under-estimated due to limitations of the bioanalytical methods that were applied (Wang et al. 2012). Knowledge of the relationship between the ADA response and clinical efficacy and safety parameters is often lacking for the same reason. Therefore, directly comparative, parallel-design, clinical studies that use "state-of-art" bioanalytical methods are required to provide a reliable index of the ADA response for the purpose of the biosimilarity exercise.

Implications: Case of Pegfilgrastim

Since risk should take into account both severity and rate of occurrence, it could be argued that a biosimilar pegfilgrastim candidate might be considered in the "low to negligible" risk category: Although there could be a theoretical risk of severe consequences if a pegfilgrastim product were to induce anti-drug antibodies (ADA) that resulted in neutralization of the function of the endogenous factor, G-CSF, such an outcome has not yet been reported.

Post-marketing experience for both pegfilgrastim and filgrastim has also demonstrated an absence of clinically impactful immunogenicity associated with the use of either product, even in fully immune competent populations. For example, a 2014 publication by Pulsipher et al. (2014) provided results from a prospective 5-year study of 6768 peripheral blood stem cell donors who were treated with G-CSF and 2726 bone marrow donors who were not treated with G-CSF. The results of that study showed that peripheral blood stem cell donors were not at increased risk for developing an autoimmune disease when compared to bone marrow donors. In addition, FDA has stated that the Agency is unaware of reports of neutralising antibodies to G-CSF products, concluding that the literature indicates that G-CSF products are low risk for causing anti-drug antibody-related severe adverse effects (FDA 2015a).

Nevertheless, as pre-existing PEG-reactive antibodies have been detected in the human population, there could be a potential impact of pre-existing or treatment-boosted PEG-reactive antibodies to bind to and alter the pharmacokinetics and pharmacodynamics of Pegfilgrastim. Although such an effect would be expected to be distributed across the treatment groups in a comparative study, binding of PEG-reactive antibodies might contribute to an increase in inter-subject variability, thereby compromising statistical power for demonstration of bioequivalence. Accordingly, additional bioanalytical testing may be useful to confirm the specificity of ADA-positive samples for the PEG moiety, and to exclude induction of ADA reactive with the filgrastim moiety; and, descriptive analysis of impact on PEG-reactive ADA status on PK and PD parameters should be planned. Careful optimization and validation of the ADA assay is needed to achieve suitable sensitivity and precision for detection of PEG-reactive ADA, because affinity/avidity might

be relatively low if the ADA is primarily of the IgM isotype (Armstrong et al. 2007), and the presence of detergent in assay buffer can substantially reduce detectability (Sherman et al. 2012).

Therapeutic application in the oncology setting would tend to decrease risk since patients are likely to be immunocompromised. Therefore, conducting comparative studies of the relative immunogenicity of a biosimilar pegfilgrastim versus the reference product in healthy volunteers would favour increased sensitivity to detect differences in immunogenicity. As pegfilgrastim is not used for chronic treatment, and there are no identifiable immunogenicity-related risks for the reference product, clinical evaluation in short-term studies should be adequate to assess relative risk for marketing authorization. Administration of two doses may be sufficient to assess impact of (1) pre-existing ADA on the PK and PD response to a first administration, and (2) treatment-induced or treatment-boosted ADA following a second administration. The potential for a confounding effect of a cross-over between treatments would favour a parallel-group design, such that a 3-period study involving cross-over between periods 1 and 2, followed by treatment with the same product version in periods 2 and 3, might represent the most suitable study design for comparative evaluation of PK, PD and immunogenicity of a biosimilar pegfilgrastim candidate versus the reference product (Karsten Roth et al. 2017).

Implications: Case of Adalimumab

Treatment-induced ADA following treatment with Humira has been reported to decrease drug exposure and efficacy in some autoimmune disease patients, thereby negatively influencing overall clinical benefit versus risk (Bartelds et al. 2007). Although adalimumab is described as a “fully human” therapeutic monoclonal antibody, the non-human germ-line sequences in the antigen-binding regions of the variable chains appear to contain immunogenic motifs (Harding et al. 2010) which induce antibodies that block binding to the target antigen, TNFa (van Schouwenburg et al. 2013a; van Schie et al. 2015a, b). As a consequence, there appears to be an inverse relationship between the level of ADA (“ADA titer”) and drug trough concentration—which, for subjects with relatively high ADA titers, appear to reduce active drug levels below a threshold for maximal efficacy (Bartelds et al. 2011). The ADA response appears to increase during the first 6 month of chronic treatment, before reaching a plateau of “persistent” ADA. The induction of a sufficient level of persistent ADA, with consequence reduction in the level of active drug, may explain the loss of response in some treated subjects during chronic treatment (van Schouwenburg et al. 2013b).

The experience gained for Humira led to a cautious approach for approval of biosimilar versions of adalimumab, involving monitoring of transient and persistent ADA responses relative to drug exposure (C_{trough}) and efficacy during a minimum treatment period of 12 months in a directly comparative therapeutic setting. In addition, immunogenicity was also compared following a single-dose administration in healthy volunteers, in which ADA was detected in up to 100% of subjects (Table 19.2).

Table 19.2 Proportion of ADA in healthy volunteers receiving a single dose of adalimumab (adapted from EPAR for SB5, EMA 2017b)

Study design	Treatment group	ADA incidence
Phase 1, comparative PK	SB5	62/63 (98.4%)
• 3-way parallel-group	EU-Humira	60/63 (95.2%)
• SB5 vs. EU-Humira vs. US-Humira	US-Humira	63/63 (100%)
• Single dose s.c. injection of 40 mg • Post-dose monitoring of drug concentration and ADA for 71 days		

Table 19.3 AUC 0-inf by ADA titer category (sourced from EMA 2017b)

	SB5 (N = 52/62)		EU-Humira (N = 58/63)		US-Humira (N = 57/62)	
ADA titer sub-group	n	Mean	n	Mean	n	Mean
Low	19	2973	17	3057	15	3218
Medium	22	2291	30	2236	27	2209
High	11	1621	11	1697	15	2013

In the single-dose pharmacokinetic study of the biosimilar adalimumab, SB5 (IMRALDI, Samsung), in healthy individuals, the number of ADA-negative individuals was too low for comparisons. Therefore, the Applicant divided the subjects to low, medium, and high titer groups. The $AUC_{0\text{-inf}}$ and $AUC_{0\text{-last}}$ had an inverse correlation to the titers of ADAs. High titer ADA-positive patients had higher adalimumab clearance rates. The impact of high titer ADAs was also seen on Cmax. The magnitude of the influence of ADA titer on AUC 0-inf appeared similar for the SB5-, EU-Humira, and US-Humira subgroups (Table 19.3, EMA 2017b). In addition, PK parameters were consistent across the treatment groups when ADA-positive and negative patients were analyzed separately (EMA 2017b).

A comparison of ADA titers *vs*. drug concentration for individual subjects indicated that although drug concentration was reduced with increasing ADA titer, there was no difference in the magnitude of this effect for the biosimilar versus reference product versions, confirming that there was no difference in clinically impactful immunogenicity. These results indicate that, even in the case of a product that induces an ADA response in most treated subjects, comparison of the time-course and magnitude of ADA titers relative to drug concentration provides a suitably sensitive index for the assessment of clinically impactful immunogenicity.

Implications: Case of Darbepoetin

As stated in the prescribing information (Amgen 2018):

> Pure red cell aplasia (PRCA) in association with neutralising antibodies to native erythropoietin has been observed in patients treated with Aranesp. This has been reported predominantly in patients with chronic renal failure and in patients with hepatitis C

treated with interferon and ribavirin. Most cases have been associated with subcutaneous administration of ESAs.

A high severity of consequences allied to low rate of occurrence implies a need for cautious strategy involving comparative evaluation of immunogenicity using subcutaneous administration in chronic renal failure patients during 12-month maintenance phase therapy, in combination with a comparative multiple dose PK/PD study in healthy volunteers using intravenous administration (EMA 2010); evaluation of a single switch from the reference product to the biosimilar version is also recommended. Bioanalytical methods will need to demonstrate adequate sensitivity to detect markers of an early treatment-induced immune response, including low-affinity ADA of the IgM isotype, as well as pre-existing EPO-reactive antibodies (Barger et al. 2012).

Implications: Case of Abatacept

For abatacept, there is a *theoretical* concern that antibodies could be induced by the CTLA4 moiety of the fusion protein that are then reactive with endogenous CTLA4 expressed on T-lymphocytes. If this were to occur, anti-CTLA4 antibodies might neutralize the activity of the endogenous protein and compromise its immune-modulatory role. Hypothetically, this could result in worsening of the autoimmune disease that abatacept was intended to treat, or increase risk of development of other autoimmune disease/events.

Studies with IV abatacept revealed a relatively low incidence (<5%) of CTLA4-reactive non-neutralizing antibodies—actually a lower incidence than detected in the placebo control group (EMA 2017c). However, an increased ADA positive frequency following end of treatment was noted—an effect that would be consistent with the intended mode of action of abatacept to suppress the adaptive immune response. Alternatively, appearance of anti-abatacept antibodies after discontinuation of the treatment may signal a poor drug tolerance of the ADA assay. These considerations imply that it is prudent to monitor the longer-term evolution of the immune response to therapeutic proteins that have an immune-suppressive action, both during treatment and following end of treatment.

Parameters for Evaluation of Relative Immunogenicity

Immune Response Parameters

Priority is given to measuring pre-existing and treatment-emergent ADA and neutralizing antibody (NAb) at suitable time-points to provide a descriptive comparative analysis of the dynamics (frequency, magnitude and persistence) of the immune response. Excellent advice on data presentation is provided in consensus documents

developed jointly by industry and regulatory authorities (Shankar et al. 2014; Rup et al. 2015), in addition to published examples from biosimilar programs.

ADA/NAb titer (reciprocal of dilution) can represent the most useful parameter for comparative purposes because it represents a continuous variable that reflects the magnitude of the response, at least in a quasi-quantitative manner; demonstrating overlap of the distribution of ADA titers for individual subjects provides reassurance of similar capacity of the biosimilar and reference products to induce a humoral immune response. Titer is considered to be a quasi-quantitative parameter because it reflects a combination of both the amount and avidity of the polyclonal immune response, rather than a quantitative measure of the amount of analyte. Nevertheless, for immunogenic therapeutic proteins, ADA titer has been shown to correlate with the scale of clinical impact (Bartelds et al. 2011), and has been used very effectively to compare the relative time-course of the magnitude of the ADA response to biosimilar vs. Reference versions (EMA 2013).

In the case of therapeutic monoclonal antibodies, non-human germline amino acid sequences located in the target antigen-binding regions of the variable chains can represent the immunodominant motifs recognised by the human immune system (van Schie et al. 2015b), even though there appears to be a clonally diverse immune response in adalimumab-treated patients (van Schouwenburg et al. 2014). Because ADA can compete with the target for binding to these regions, the ADA and NAb responses may actually represent the same polyclonal antibody population (van Schie et al. 2017), measured in different assay formats—this was illustrated by high concordance of results from the ADA and NAb assays performed for a biosimilar infliximab (EMA 2016a). Therefore, although requested by regulatory authorities, the added value of the NAb assay for the biosimilarity exercise of antagonist therapeutic monoclonal antibodies is equivocal.

As discussed earlier, in the case of a product that has an identified risk associated with induction of ADA that can bind to, and neutralize the function of, and endogenous counterpart, e.g. epoetin or darbepoetin, application of bioanalytical methods with sensitivity to detect an early immune response (e.g. via ADA of IgM isotype) is required. In addition, potential reactivity of positive signals with the endogenous counterpart should be performed using suitable ADA and NAb assay formats.

Demonstration of the specificity of the bioanalytical methods to detect ADA/NAb induced by either the biosimilar or reference product version is rigorously assessed by regulatory authorities; although published examples indicating that biosimilars meeting the analytical similarity standard can induce differential ADA specificities from those reactive with the reference product are currently lacking. For infliximab and adalimumab, ADA induced by the biosimilar version was shown to have identical reactivity to that induced by the reference product (Ben-Horin et al. 2015; Reinisch et al. 2017; Fiorino et al. 2018; Amgen 2016a, b). Thus, a single-assay approach, using the biosimilar candidate as the antigen, is usually the preferred approach, if supported be demonstration of antigenic equivalence, because this will avoid imprecision associated with different assay formats (Chamberlain 2014); in conjunction with demonstrating inhibition

of positive signals by unlabelled version of both the biosimilar and the reference product to confirm antibody equivalence (Chamberlain 2014; Ryding et al. 2017). In case of an increased incidence of ADAs in the biosimilar arm, regulators may ask the sponsor to develop another assay with the reference product as the antigen. The demonstration that the results of the two assays are highly concordant will suggest that the difference is not qualitative, i.e. does not reflect a difference in specificity of the signals.

If the product is a fusion protein or a conjugate, regulatory authorities expect that the specificity of positive signals for the respective moieties is confirmed (EMA 2017a) as the two types of ADA may have different consequences. In the case of a pegylated therapeutic protein, it may be necessary to report the titer of pre-existing and treatment-boosted PEG-reactive ADA, in addition to reporting of the titer of ADA reactive with the protein moiety. A published example of the presentation of the moiety specificity of ADA signals for a pegfilgrastim biosimilar candidate is shown in Table 19.4. These results indicate that the ADA detected in the screening ADA assay using labeled pegfilgrastim as the antigen were then confirmed to have specificity for the PEG moiety, rather than for the filgrastim moiety; two samples with confirmed PEG-reactive ADA were found to have neutralizing capacity in an in vitro cell-based assay, possibly reflecting steric hindrance of engagement of G-CSF receptors on the cell surface when ADA is associated with the PEG moiety of pegfilgrastim. Interestingly, healthy volunteers appeared to have a higher incidence of detectable PEG-reactive antibodies than breast cancer subjects.

Depending on the manufacturing process for the biosimilar, and on differences in analytical profile compared to the reference product, it may be necessary to monitor for induction of treatment-emergent immune responses to process-derived impurities *and/or* post-translational modifications (e.g. non-human glycan variants). This consideration increases the value of soliciting endorsement for the proposed bioanalytical approach from regulatory authorities, via engagement of scientific advice procedures prior to initiation of clinical studies.

It is important to remember that the assays applied for different products are not standardized in terms of format, operating conditions or control reagents. Surrogate positive controls antibody reagents developed in non-human species have equivocal relevance to the human immune response, and many factors may contribute bias to these assays—particularly interference by the residual drug, which may lead to under-detection of ADA and NAb (Wang et al. 2012; Bloem et al. 2015). Thus, results from these assays should be interpreted only in relation to appropriately sensitive and relevant clinical parameters. Certainly, direct comparison of reported ADA incidence detected by different assay formats can be highly misleading, and should be avoided because there may be no consistent relationship with clinical impact. Also, the apparent time-course of development of ADA can be dramatically influenced by analytical sensitivity of the assay applied for ADA monitoring (van Schouwenburg et al. 2013b)—and this aspect needs to be considered carefully in planning clinical studies as well as in the interpretation of the results. The potential methodological limitations associated with some ADA assays represent another reason for needing comparative, parallel-group, studies to evaluate the relative immunogenicity of biosimilar and reference products.

Table 19.4 Incidence and specificity of ADA signals detected in clinical studies for pegfilgrastim biosimilar candidate, MYL-1401H

	MYL-1401H-1001 Healthy volunteers			MYL-1401H-1002 Healthy volunteers		MYL-1401H-3001 Breast cancer	
Category n (%)	MYL-1401H	EU-Neulasta	US-Neulasta	MYL-1401H	EU-Neulasta	MYL-1401H	EU-Neulasta
ADA positive	14 (22.2)	16 (23.5)	21 (30.4)	6 (26.1)	7 (29.2)	0	1 (1.8)
Specificity							
PEG only	8 (12.7)	13 (19.1)	14 (20.3)	5 (21.7)	6 (25.0)	0	0
G-CSF only	0	0	0	1 (4.3)	0	0	1 (1.8)
PEG and G-CSF	6 (9.5)	4 (5.9)	6 (8.7)	1 (4.3)	2 (8.3)	0	0
Neither	1 (1.6)	2 (2.9)	2 (2.9)	2 (8.7)	2 (8.3)	0	0
NAb positive	2 (3.2)	0	0	0	0	0	0

Adapted from Waller et al. (2017)

PK/PD

Since PK and PD parameters are generally more sensitive than clinical efficacy endpoints in assessing the similarity of two products, comparative PK/PD studies using dose-levels that fall within the linear part of the dose-response range represent an essential element of the biosimilarity exercise (FDA 2016a). For some products, e.g. filgrastim and pegfilgrastim and insulin, comparative PK/PD data can represent pivotal evidence of therapeutic equivalence (EMA 2006).

Often, PK parameters represent the most sensitive index of clinically impactful ADA formation (Wolbink et al. 2005; Bendtzen et al. 2006; Radstake et al. 2009), such that descriptive analyses of the relationship of ADA and PK provides important information for the assessment of relative immunogenicity of the biosimilar candidate and the reference product. If a suitable PD is available, a similar descriptive analysis could be performed to qualify interpretation of clinical impact of any differences observed in PK parameters for the ADA positive *vs.* ADA negative subpopulations within and across the respective treatment groups (see Shankar et al. 2014; Rup et al. 2015 for examples of data analysis and presentation).

ADA sampling time-points are synchronized with sampling for drug concentration to enable direct comparison of the outputs from the respective assays throughout the treatment period and, if necessary to minimize drug interference in the ADA assay, at a suitable time following end-of-treatment: this enables descriptive analysis of the relationship between ADA/NAb status (positive at least one time-point or negative throughout) or ADA/NAb titer (maximum value during treatment period) and PK parameters (Cmax, AUC, T1/2, drug trough concentration).

Since the assay used to measure drug concentration may be influenced by the ADA level, an apparent reduction in the drug level may not necessarily reflect enhanced clearance of the drug from the circulation—rather, the reduced signal could be due to competitive inhibition of binding of the drug to the target antigen used in the assay to measure drug concentration. Conversely, if the assay does not distinguish free from ADA-bound drug, an increase in signal due to circulating inactive drug-ADA complexes may be detected. Thus, careful validation of specificity and ADA interference of the bioanalytical method used for the PK analysis is required to enable reliable interpretation of the results for the immunogenicity assessment. Application of a drug concentration assay that measures the free drug (i.e. not bound to the target or to ADA) tends to provide the most sensitive index of ADA formation. Whether any observed difference between the relative ADA versus PK parameters for the comparator products is truly “clinically meaningful” would still need to be assessed relative to the therapeutic dose-response curve.

Efficacy

If a therapeutic equivalence study is required for confirmation of biosimilarity, the relationship of ADA positive/negative status, or quartiles of ADA titers, to the primary efficacy endpoint in each treatment group should be analysed in a

descriptive manner. This would normally be done at the primary efficacy time-point. If the study includes a treatment transition or switch, the comparison could also be performed at a suitable time following the transition or switch.

The equivalence range will be calculated on the basis of the therapeutic effect in previous clinical studies of the reference product. Since the dose- response curve of several therapeutic proteins, notably monoclonal antibodies, is flat in the upper therapeutic dose range, non-inferiority studies could be considered. This is the case especially in oncology where treatment effect of the reference product may be variable due to methodological challenges. Nevertheless, it is appropriate to consult relevant regulatory authorities if a non-inferiority design is considered.

Depending on the number of data points available, descriptive analyses of ADA titer may be useful, e.g. to illustrate that the relationship of efficacy to different ADA titer tertiles/quartiles and to the ADA negative sub-population is similar for each treatment group (Shankar et al. 2014; Rup et al. 2015).

It is important to have a sufficiently long follow up of immunogenicity. This will allow the characterization of the kinetics of ADA response. A difference between comparator groups in ADA frequency or ADA titer that appears to be widening with time could question similarity of clinical impact of the immune response. Thus, it is important to design the study to enable descriptive analysis of the relationship between ADA and efficacy at a time-point that corresponds to the maximal difference in immune response across the two treatment groups. If the maximal difference in the extent of the immune response is not captured on the pre-authorization phase, it may be necessary to perform a comparative post-authorization study of adequate duration to exclude a possible differential clinical impact.

As discussed earlier, the apparent time-course of the immune response can be substantially altered by the sensitivity of the bioanalytical method used, such that the plateau of ADA frequency may be attained earlier when a more sensitive assay is used (van Schouwenburg et al. 2013b).

Safety

Adverse Events of Special Interest (AESI) should be prospectively defined, using established terminology, e.g. FAAN criteria for anaphylaxis (Manivannan et al. 2009). While possibly unrelated to ADA formation, injection-site or infusion-related reactions are often included in the AESI category. Incidence and severity of AESI's by ADA positive/negative status can then be compared for the treatment groups in a descriptive analysis.

For analysis of individual cases of suspected immune-mediated adverse events, it can be instructive to review the ADA titer value at the sampling time-points prior to and following the observed AE, as this might support classification as drug-related hypersensitivity associated with immune-complex formation.

In general, the number of observed events in randomized controlled clinical studies is too low to enable conclusions about association to the immune response, and association with ADA also compromised by the non-randomized post-hoc classification of ADA positive/negative status.

Inferences

- The term "immunogenicity" has wider meaning than "ADA/NAb frequency/titer profile"
- Bioanalytical measures of the immune response (ADA and NAb) should be interpreted cautiously, taking into account their *quasi*-quantitative nature allied to potential sources of bias such as interference by residual drug
- ADA titer may provide a more useful index than ADA frequency for comparison of the humoral immune response to the biosimilar candidate and reference products; for example, box-and-whisker plots showing the distribution of ADA titer values at each sampling time-point by treatment group to illustrate relative magnitude of the humoral response as a continuous variable, rather than relying on a binary output (positive or negative ADA status)
- PK/PD parameters often provide the most sensitive indices of clinical impact of the treatment-induced immune response, particularly in the case of single-dose administration, parallel-group studies performed in healthy volunteers for demonstration of bioequivalence
- Validation of the selectivity of the assay used to measure drug concentration should include evaluation of interference by ADA, as well as specificity to distinguish bound *vs*. unbound drug
- Evaluation of the relationship of bioanalytical parameters to efficacy endpoints should include descriptive sub-group analyses by ADA/NAb positive/negative status and, if sufficient data points are available, ADA/NAb titer.
- Profiles for adverse events of special interest (AESI) such as drug-related hypersensitivity should also be compared by ADA/NAb positive/negative status and, if sufficient data points are available, ADA/NAb titer; although low numbers of data points and confounding variables may preclude definitive conclusions regarding relationship to an immune response to the product.

Acceptable Margin of Difference

Why Pre-definition is not Feasible

FDA guidance (FDA 2015b) states:

> The design of any study to assess immunogenicity and acceptable differences in the incidence and other parameters of immune response should be discussed with FDA before initiating the study. Differences in immune responses between a proposed product and the reference product in the absence of observed clinical sequelae may be of concern and may warrant further evaluation (e.g., extended period of follow-up evaluation).

In the view of the authors, it would be feasible to pre-define "acceptable differences in the incidence and other parameters of immune response" only if *all* of the following conditions were fulfilled:

1. Bioanalytical methods are standardised and indices of immune parameters are reported relative to common reference standards (van Scouwenberg et al. 2016) that allow quantitative comparison between different studies for different products;
2. There is adequate pre-existing knowledge of the relationship of the bioanalytical outputs to relevant clinical endpoints, taking into account differences in sensitivity of PK, PD, efficacy and safety indices to particular immune response parameters, and to the immune competence of the population(s) to be treated.
3. Frequency of ADA responses associated with clinical impact is sufficient to enable detection of differences in pre-approval studies.

Currently, since conditions (1) and (2) are not fulfilled, attempts to pre-define clinically meaningful acceptance margins for the numerical difference in immune response parameters between the biosimilar and reference product versions would be very difficult. Moreover, the case examples below show how the applied regulatory decision-making process has been strongly influenced primarily by a comparison of clinical pharmacology, efficacy and safety results, and secondarily by similarity in product quality attributes—particularly sub-visible particle levels—that have been associated with incremental immunogenicity of other products. Then, depending on the residual uncertainty allied to the risk profile for that product, supplemental data from post-authorization monitoring (possibly including sampling for measurement of ADA and drug levels in addition to evaluation of sustainability of treatment response and AESI) in "real-world" conditions might provide a tertiary level of evidence for risk mitigation.

Case of Infliximab

Since treatment with infliximab is associated with clinically impactful immunogenicity (Beart et al. 2003; Wolbink et al. 2006; Bendtzen et al. 2006), review of the regulatory assessments made for the marketing authorization of the first two biosimilar versions, CT-P13 (Remsima) and SB2 (Flixabi, Renflexis), is instructive for understanding the acceptability of numerical differences in ADA detected by bioanalytical methods relative to the overall weight of evidence of similarity.

In the EU, CT-P13 was approved on the basis of results from comparative clinical studies in two patient populations, namely ankylosing spondylitis and rheumatoid arthritis. This combination provided data for the relative immunogenicity assessment that reflected populations receiving different dose levels (5 mg/kg in ankylosing spondylitis; 3 mg/kg in rheumatoid arthritis) as well as different levels of concomitant immune-suppressive medications during randomized, double-blind treatment period of 54 weeks. Because these studies compared the biosimilar version to reference product sourced from EU, the Company performed an additional

study to support US marketing authorization: a 3-way, parallel-group, comparative PK study in healthy volunteers, using a single 5 mg/kg dose of CT-P13 vs. US-sourced Remicade vs. EU-sourced Remicade. The ADA results for the respective studies are summarized in Table 19.5.

Although the main therapeutic studies in ankylosing spondylitis and rheumatoid arthritis patients indicated a consistent frequency of ADA positive patients throughout the 54-week treatment period, in addition to comparable ADA titer profiles (FDA 2016b), FDA questioned the numerical difference detected for CT-P13 at the week 8 sampling time-point in the healthy volunteer population (FDA 2016c); this led to re-analysis of the samples using a more drug-tolerant assay, as well as a thorough review of the analytical similarity data for sub-visible particles (FDA 2016d). Based on the totality of evidence, which included a demonstration that PK parameters in ADA positive subjects were within the 90% confidence intervals for the comparison of CT-P13 vs. US-sourced Remicade and CT-P13 vs. EU-sourced Remicade, FDA concluded that that CT-P13 is highly similar to the reference product and that the results of the 3-way comparative PK study confirmed the relevance of clinical immunogenicity data from studies using EU-approved Remicade. The applicant also submitted additional ADA monitoring results from an ongoing randomized, double-blind, controlled, post-marketing study in patients with active Crohn's Disease that were considered to "add to the totality of evidence to support the conclusion that there are no clinically meaningful differences between CT-P13 and the US-licensed Remicade" (FDA 2016e).

Numerical differences in ADA frequency were also detected in rheumatoid arthritis patients during both treatment periods (54-week main study followed by 24-week extension) of a therapeutic equivalence study performed for SB2 (FLIXABI/RENFLEXIS, Samsung). As illustrated in Fig. 19.1, there was a slightly higher frequency of detected ADA positive patients throughout the study in patients treated with SB2 compared to EU Remicade.

As noted by FDA (2017b), although the reported ADA frequency was approximately 5% higher in the SB2 group, the magnitude of the difference did not increase over time. Moreover, the impact of ADA-positive vs. ADA-negative status on the ACR20 response (Table 19.6) and on incidence of infusion-related reactions (Table 19.7) was similar for the SB2 and EU-Remicade treatment groups. ADA frequency did not increase during the extension period for patients who were switched from EU-Remicade to SB2, relative to patients who were maintained on EU-Remicade or SB2.

Finally, the review by the FDA product quality team (FDA 2017c) concluded:

> In the case of high-molecular weight (HMW) species, additional characterization data support that the HMW species observed by size exclusion chromatography are non-covalent and reversible. Stability data demonstrate that the slightly higher levels for SB2 (0.6–0.9% versus ≤0.5% for US-licensed and EU-licensed Remicade) do not impact product stability or lead to excessive sub-visible particle formation.

Taking into account the totality of evidence, FDA authorized SB2 as a biosimilar version of infliximab.

Table 19.5 Comparison of ADA frequency across CT-P13 clinical studies (study 1.4 in healthy subjects, study 1.1 in patients with AS, and study 3.1 in patients with RA)

Assay	Time-point	Study 1.4 in healthy subjects (5 mg/kg single dose)			Study 1.1 in AS (5 mg/kg at week 0, 2, 6, and then q8w to week 54)		Study 3.1 in RA (3 mg/kg at week 0, 2, 6, and then q8w to week 54)		Study 1.3 in AS (5 mg/kg q8w)		Study 3.2 in RA (3 mg/kg q8w)	
		CT-P13 (N = 70)	EU (N = 71)	US (N = 70)	CT-P13 (N = 125)	EU (N = 125)	CT-P13 (N = 302)	EU (N = 304)	CTP-13 to CTP-13 (N = 90)	EU to CTP-13 (N = 84)	CT-P13 to CTP-13 (N = 159)	EU to CT-P13 (N = 143)
ECLA	Pre-dose	2 (2.8%)	1 (1.4%)	1 (1.4%)	2 (2%)	1 (<1%)	9 (3%)	6 (2%)	2 (2%)[a]	1 (<1%)[a]	7 (4%)[a]	4 (3%)[a]
	Week 8	10 (14.3%)	5 (7%)	2 (2.9%)	–	–	–	–	–	–	–	–
	Week 14	–	–	–	11 (9%)	13 (11%)	69 (23%)	70 (23%)	–	–	–	–
	Week 30	–	–	–	32 (25%)	25 (20%)	122 (40%)	122 (40%)	–	–	–	–
	Week 54	–	–	–	25 (20%)	28 (23%)	124 (41%)	108 (36%)	–	–	–	–
	Week 78	–	–	–	–	–	–	–	21 (23%)	25 (30%)	71 (44%)	66 (46%)
	Week 102	–	–	–	–	–	–	–	21 (23%)	23 (27%)	64 (40%)	64 (45%)
ELISA	Pre-dose	4 (5.6%)	0	1 (1.4%)	–	–	–	–	–	–	–	–
	Week 8	19 (26.8%)	18 (25.4%)	8 (11.4%)	–	–	–	–	–	–	–	–

Re-drawn from Table 16, FDA Briefing Document for Arthritis Advisory Committee meeting held on 9 February 2016 (FDA 2016b), based on FDA analysis of data from Celltrion 351(k) BLA submission

[a]Reflects the incidence of ADA positivity at the screening/pre-dose visit in index clinical studies 1.1 and 3.1 of the patients who enrolled in the extension studies 1.3 and 3.2

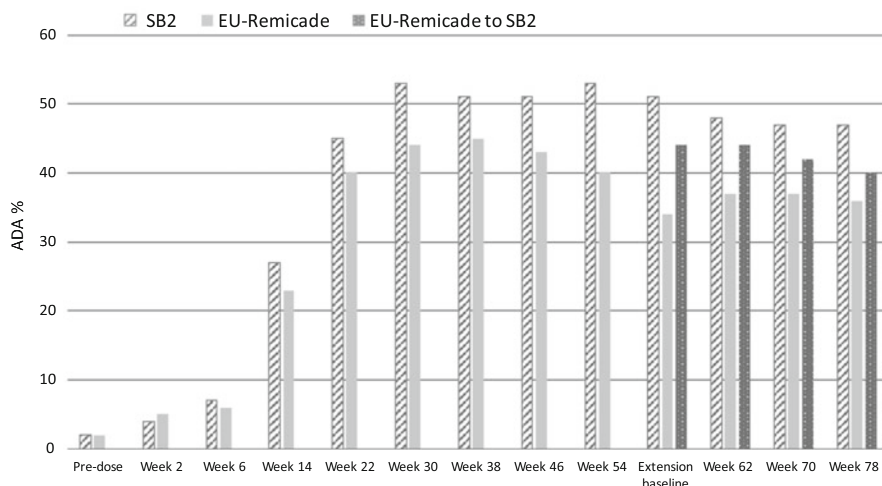


Fig. 19.1 Proportion of ADA positive patients following repeat dosing in study SB2-RA. Source: Redrawn from data shown in FDA clinical review summary (FDA 2017b)

Table 19.6 ACR20 response by ADA status (study SB2-RA, per-protocol set 1)

ADA category	Treatment	Week 30 n/N (%)	Week 54 n/N (%)
ADA positive	SB2	72/127 (57%)	66/117(56%)
	EU-Remicade	74/126 (59%)	69/106(65%)
ADA negative	SB2	76/104 (73%)	73/98 (75%)
	EU-Remicade	89/121 (74%)	81/111 (73%)

Re-drawn from Table 28, FDA clinical review for Samsung 351(k) BLA submission for SB2 (FDA 2017b)

Table 19.7 Incidence of infusion-related reactions by ADA status (Study SB2-RA)

TEAE	ADA category	SB2 (n = 290)	EU-Remicade (N = 293)
Infusion-related reaction	ADA positive	15 (5%)	12 (4%)
	ADA negative	3 (1%)	5 (2%)

Re-drawn from Table 28, FDA clinical review for Samsung 351(k) BLA submission for SB2 (FDA 2017b)

Case of Etanercept

A biosimilar version of etanercept that was reported to contain lower levels of sub-visible particles (detected by micro-flow imaging) and HMW variants (detected by SE-HPLC) than the reference product (Cho et al. 2016; EMA 2015) was associated with a lower incidence of injection-site reactions and injection-site erythema in a randomized controlled clinical study (EMA 2015); incidence of these events did not appear to be associated with ADA positive/negative status. However, because there are some differences in the composition of the formulation of the biosimilar

and reference products, linkage to formulation components rather than differences in sub-visible particle levels cannot be excluded (L-Arginine and latex in the needle shield are absent from SB4).

Assessment of the relative treatment-emergent ADA response was confounded by drug interference in the ADA assay: overall incidence of ADA was reported as 0.7% for the SB4 compared to 13.1% for the reference product; this difference was driven largely by an apparent difference in ADA frequency at earlier (weeks 4 and 8) sampling time-points, i.e. transient antibodies that were not detected at later time-points. Because EMA had concerns about possible interference of residual drug in the ADA assay, the Applicant was asked to re-present the data after exclusion of the results for weeks 4 and 8 (EMA 2015).

Interestingly, the observation of a reduced ADA frequency at earlier sampling time-points for the biosimilar version compared to Enbrel was replicated for a second etanercept biosimilar, namely GP2015 (Erelzi). The Assessment Report states (EMA 2017f)

> In the Healthy Volunteers PK studies, a total of 3 subjects had confirmed binding anti-drug antibodies (ADAs) at the follow-up visit (Day 65) with titers near the detection limit. All 3 subjects were in the treatment sequence of Erelzi/EU-Enbrel (with Enbrel in Period 2), and none of the ADAs were neutralizing. The binding ADA positive results were considered not clinically meaningful due to the very low titers and no other safety issues were identified.
>
> In Study GP15-302, there were no ADA-positive samples detected in the Erelzi group up to Week 52.
>
> In contrast, 5 of the 267 patients in the TP1 Enbrel group (1.9%) had a confirmed positive binding ADA result, all within the first 4 weeks of treatment. Additionally, one subject in the switched Enbrel group, who had undergone the last switch from Enbrel to Erelzi at Week 24, had a confirmed positive binding ADA result at Week 36. For all patients, the obtained titer values were low and transient, and none of the ADAs had neutralizing capacity.

Injection site reactions were more common for Enbrel than for Erelzi (Week 52, Erelzi: 8.5%, Enbrel 15.8%). In general, these were of mild severity, not leading to an increased drop-out rate. It is important to note that there are differences in the drug formulation between the two products: compared to the reference product, phosphate and L-arginine were exchanged for citrate and L-lysine

Information on the comparative levels of product-related aggregates is provided in the FDA backgrounder for the Advisory Committee meeting held on 13 July 2016 (Page 23 in FDA 2016f), which states:

> The average level of aggregates in US-licensed Enbrel quantified by Sandoz's SEC-HPLC assay was 2.1%, while GP2015 was 0.4%. Overall, GP2015 has lower levels of aggregates compared with US-licensed Enbrel. This may be in part due to differences in the ages of the lots at the time they were tested, but some aged GP2015 lots also had lower levels of aggregates compared with US-licensed Enbrel.

Analytical results for the detection of HMW variants in multiple batches of the GP2015 relative to EU- and US-sourced Enbrel were summarized in the Applicants briefing material for the same meeting (Fig. 19.2)

It is not known whether the observed trend for lower transient ADA formation, or for a higher incidence of injection site reactions, for the approved biosimilar versions of etanercept SB4 and GP2015 relative to Enbrel could be related to

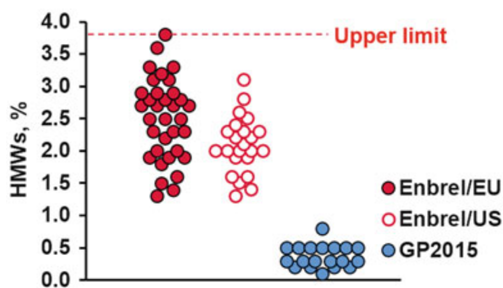


Fig. 19.2 Comparison of HMW variants detected by SE-HPLC. Source Figure 5-13 in Sandoz Backgrounder for FDA Advisory Committee meeting held on 6 June 2016 (Sandoz 2016)

differences in levels of HMW variants. As noted above, there are qualitative differences in excipients between the reference product and the two biosimilar products. Most importantly, the minor and transient difference in ADA frequency did not influence drug exposure or efficacy (EMA 2015; FDA 2016f). From the regulatory perspective, both FDA and EMA accept that a biosimilar version could have a lower immunogenicity relative to the reference product. Also, this case demonstrates that biosimilar manufacturers may achieve more effective control of a product quality attribute (HMW variants/aggregates) associated with increased immunogenicity risk compared to the originator; and that a clear difference in the level of HMW variants did not translate into a clinically meaningful risk for etanercept. Whether this is the case of other products remains to be established.

Case of Rituximab

According to the European Public Assessment Report (EMA 2017d), a twofold lower incidence of ADA was detected in rheumatoid arthritis patients treated with a biosimilar rituximab (RIXATHON, Sandoz), compared to the reference product (MABTHERA, Roche) in a 52-week duration therapeutic equivalence study: 9 ADA positive subjects (11.0% of total) were detected in the biosimilar arm compared to 18 ADA positive subjects (21.4%). This difference in ADA incidence was not associated with a differential impact on efficacy or safety. The Agency concluded that the 10% lower ADA incidence detected in the biosimilar group “could be a chance finding and ultimately it is not of concern since immunogenicity—if anything—is lower than that of the reference product”.

Inferences

- Pre-definition of statistically rigorous acceptance criteria for immune response parameters is of doubtful utility (and scientific validity) due to bioanalytical bias and uncertainty about relationship of bioanalytical outputs to relevant clinical indices.
- Drug interference in the ADA assay is a major source of bias that can compromise the evaluation of relative immunogenicity.

(continued)

- Acceptability of observed differences in immune response parameters should be assessed in relation to PK, PD, efficacy and safety over the full comparative treatment period and, if necessary to address uncertainty about impact on longer-term treatment outcomes, during an appropriate follow-up period.
- A difference in the numerical incidence and/or titer of the immune response does not preclude authorization of a biosimilar candidate where the totality of evidence has demonstrated: (1) acceptable analytical similarity; *and* (2) no clinically meaningful differences in PK/PD, efficacy and safety in randomized controlled clinically studies in adequately sensitive populations; *and*, (3) bioanalytical data enables valid conclusions.
- Levels of relevant product quality attributes, particularly aggregates and sub-visible particles, in the drug product batches used for the clinical immunogenicity evaluation should always be carefully reviewed in relation to any observed differences in the dynamics of the immune response.
- Differences in the drug product formulation-primary container combination appear to represent the most plausible reason for differences observed to date in numerical rates of detected ADA incidence between biosimilar and reference products; whether this is a consequence of differential stability of the monomeric form of the active substance, or to other variables, remains to be established. Nevertheless, these numerical differences in ADA incidence have not translated in a detectable impact on efficacy or safety during chronic administration in controlled clinical studies.
- Interpretation of the significance of observed differences in product quality variables should reflect knowledge of the immunogenicity risk profile of the particular product, which can vary widely between products as well as across different patient populations treated with the same product.

Extrapolation of Conclusions on Relative Immunogenicity to Different Therapeutic Indications

Regulatory Guidance

Regulatory authorities emphasize the importance of evaluation all data, including immunogenicity, in the context of the totality of evidence.

FDA guidance (FDA 2015b) states:

> If a sponsor is seeking to extrapolate immunogenicity findings for one condition of use to other conditions of use, the sponsor should consider using a study population and treatment regimen that are adequately sensitive for predicting a difference in immune responses between the proposed product and the reference product across the conditions of use. Usually, this will be the population and regimen for the reference product for which development of immune responses with adverse outcomes is most likely to occur

> (e.g., patients on background immunosuppressants would be less likely to develop immune responses than patients who are not immunosuppressed).

EU guidance (EMA 2014; Weise et al. 2014) is consistent with these recommendations in stating that:

> the target population of the efficacy, safety and immunogenicity study needs to be sensitive for differences in immunogenicity and its consequences and be representative for the population(s) for whom the product is indicated.

Neither of the above statements imply a need to use the clinical population for which the highest historical numerical incidence of ADA has been reported (Ebbers and Chamberlain 2016)—rather, it is recommended to use the population that provides the most sensitive comparative test of *clinically meaningful differences, including adverse outcomes*, in the immune response. Thus, in practice, the combination of immunogenicity data from the comparative PK/PD and therapeutic equivalence studies has been sufficient to enable extrapolation of conclusions about relative immunogenicity to all licensed indications (Chamberlain 2014; Reinivuori et al. 2018). For products such as pegfilgrastim and therapeutic monoclonal antibodies used in oncological settings, performing a comparative PK/PD study in healthy volunteers is likely to enhance sensitivity to detect differences in ADA response and clinical impact relative to the approved conditions of use.

Case Example: Adalimumab

Adalimumab represents an instructive case example because its use is associated with clinically impactful immune responses during chronic administration (Bartelds et al. 2007; Murdaca et al. 2016), and the scale of the treatment-emergent ADA response is influenced by concomitant immune-suppressive medication used in the rheumatoid arthritis setting (Krieckaert et al. 2012). As discussed above, the population for a therapeutic equivalence study to support biosimilarity should be the one that provides the most sensitive comparative test of clinically meaningful differences in the immune response. In the case of adalimumab, it was not clear if a study performed in rheumatoid arthritis patients (with most patients receiving concomitant methotrexate) would provide the same sensitivity to detect clinically meaningful differences compared to psoriasis patients (not receiving concomitant methotrexate).

The clinical development program for the first biosimilar adalimumab to be approved, ABP501 (AMGEVITA, Amgen), effectively addressed this question by including therapeutic equivalence studies in both settings, i.e. rheumatoid arthritis (ABP501 vs. US-Humira) and plaque psoriasis (ABP501 vs. EU-Humira), in addition to a comparative 3-way (ABP501 vs. US-Humira vs. EU-Humira) PK study in healthy volunteers (Table 19.8). The therapeutic equivalence study in plaque psoriasis also included an extension period to evaluate a single transition from the reference product to ABP501 compared with patients who were maintained on treatment with the reference product of ABP501.

Table 19.8 Summary of design of clinical studies in ABP 501 program

Study No.	Objective	Design	Subjects	Treatments
PK similarity study				
-217	3-way PK similarity, safety, immunogenicity	R, PG, SD, 3-way PK bridging	203 healthy subjects	40 mg SC • ABP 501 • US-Humira • EU-Humira
Therapeutic equivalence studies				
-262	Efficacy, safety, immunogenicity in RA	26 Weeks, R, DB, PG	526 Patients with RA	40 mg SC Q2W+MTX • ABP 501 • US-Humira
-263	Efficacy, safety, immunogenicity in PsO	R, DB, PG; re-randomized at week 16 to either continue EU-Humira or transition to ABP 501	350 Patients with PsO	80 mg SC Day 1, then 40 mg SC Q2W from week 2 • ABP 501 • EU-Humira

Adapted from FDA Briefing Document for Arthritis Advisory Committee meeting dated 12 July 2016 (FDA 2016g)
R randomized, *PG* parallel group, *SD* single dose, *DB* double-blind, *RA* rheumatoid arthritis, *PsO* plaque psoriasis, *SC* subcutaneous, *Q2W* every 2 weeks, *MTX* methotrexate

Table 19.9 Proportion of ADA positive subjects in comparative PK study 217

	Number (%) of ADA positive subjects		
Time-point	ABP 501 (N = 67)	US-Humira (N = 69)	EU-Humira (N = 67)
Day 1, pre-dose	0	0	0
Day 16	12 (18%)	12 (17%)	23 (35%)
Day 29	21 (32%)	27 (42%)	27 (42%)
End-of-study	29 (43%)	34 (50%)	34 (51%)

Table re-drawn from FDA data analysis presented in FDA Briefing Document for Arthritis Advisory Committee meeting held on 12 July 2016 (FDA 2016g)

In the comparative 3-way PK study in healthy volunteers, ADA incidence (Day 1 to End-of-Study) by treatment was: 54% for ABP 501; 55% for Humira-US; and 67% for Humira-EU (Table 19.9).

Comparative ADA frequency detected in the respective therapeutic equivalence studies is summarized in Fig. 19.3.

Both studies demonstrated similar profiles of treatment-emergent ADA in response to treatment: a higher frequency of ADA was evident in the psoriasis study compared to the rheumatoid arthritis study for both treatment groups, with a trend for slightly lower ADA frequency for ABP501 relative to the reference product in the psoriasis study. A single transition to ABP501 had no apparent impact on ADA frequency.

Analysis of the relationship of drug trough concentration for the ADA positive and ADA negative sub-populations in each treatment arm revealed a remarkably consistent impact across the two studies: in both studies, serum drug levels were, as

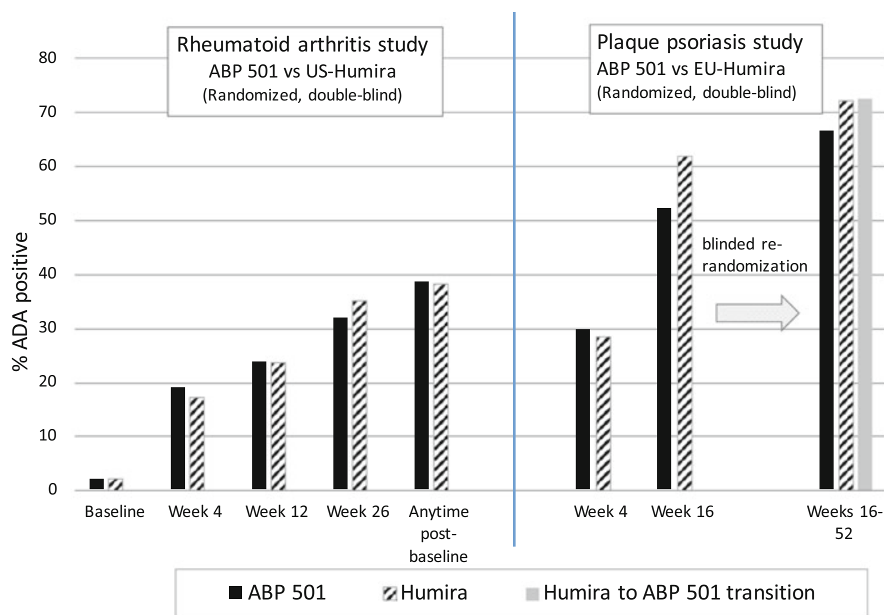


Fig. 19.3 Proportion of ADA positive subjects in therapeutic equivalence studies. Figure drawn using data presented Amgen Briefing Document for Arthritis Advisory Committee meeting held on 12 July 2016 (FDA 2016g)

expected, clearly reduced in the ADA positive sub-population, but the magnitude of the difference between the ADA positive and ADA negative subpopulations was the same for ABP501 and the reference product; and highly comparable profiles were observed across the two clinical studies (Fig. 19.4).

Absence of any difference in clinically impactful immunogenicity associated with ABP501 relative to the reference product was confirmed by similarity of the efficacy endpoints, and no clear difference in safety profile, in the respective therapeutic equivalence studies.

Overall, these results—particularly the drug trough concentration comparisons for the ADA positive vs. ADA negative sub-populations—support extrapolation of conclusions about relative clinically meaningful immunogenicity of different adalimumab versions across different therapeutic populations, despite a lower incidence of ADA in the population receiving concomitant methotrexate. Thus, it should not be necessary to perform therapeutic studies in more than one therapeutic setting to support a demonstration of biosimilarity in the case of adalimumab.

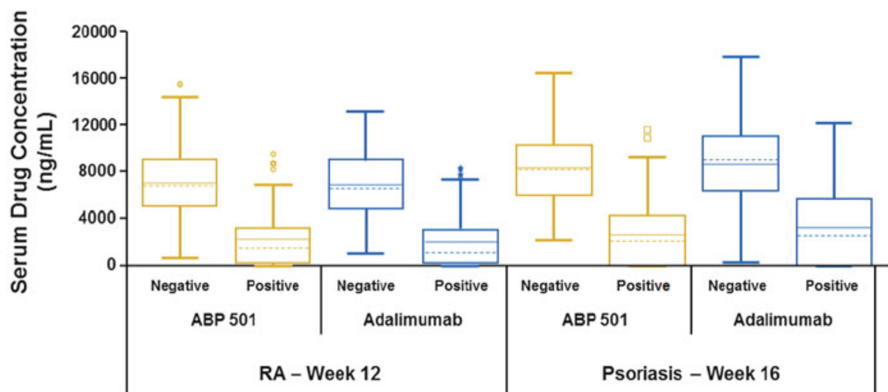


Fig. 19.4 Serum trough concentrations by binding ADA status. Note: Negative and positive denotes ADA status at the displayed time points. Within each box, solid lines represent the median and dashed lines represent the mean. *RA* rheumatoid arthritis. Reproduced without changes from Amgen Briefing Document for Arthritis Advisory Committee meeting held on 12 July 2016 (Amgen 2016b)

Other Products

In the case of the infliximab biosimilar, CT-P13, chronic administration (54-week followed by 48-week extension period) studies were conducted in two populations, namely rheumatoid arthritis and ankylosing spondylitis (Weise et al. 2014). This combination allowed comparison of populations receiving different dose levels of the drug (3 mg/kg in rheumatoid arthritis compared to 5 mg/kg in ankylosing spondylitis), as well as different levels of concomitant immune-suppressive medications. Surprisingly a lower incidence of ADA was detected in the ankylosing spondylitis setting, despite absence of concomitant methotrexate (although use of other immunosuppressive medications was permitted)—possibly, as a consequence of higher drug interference in the ADA assay (Chamberlain 2014). As discussed below, although numerical differences in detected ADA were observed, these did not reach statistical significance and did not translate into a differential impact on efficacy up to and including week 102 in either study.

Two rituximab biosimilars (GP2013, RIXATHON, Sandoz; CT-P10, TRUXIMA, Celltrion) have been approved to date by EMA. Comparative clinical studies were performed in both the rheumatoid arthritis and oncology settings (EMA 2017d, e). In both cases, ADA incidence was higher for the rheumatoid arthritis population, but there was no difference between treatment groups in terms of impact of ADA positive status on PK, efficacy or safety.

Control of Confounding Patient-Related Variables

For most products, subjects enrolled into comparative PK and therapeutic studies to demonstrate biosimilarity will be treatment-naïve. Thus, interference from pre-existing antibodies should not represent a confounding influence. In the case of a pegylated protein such as pegfilgrastim, a significant proportion (>20%) of the treatment-naïve population are likely to have low levels of pre-existing antibodies reactive with the PEG moiety of the conjugated protein. Such pre-existing antibodies can generate confounding signals in the ADA assay and, potentially, could influence PK parameters following a first administration. As discussed earlier, such an effect could increase inter-subject variability of PK parameters in a manner that compromised the statistical power to demonstrate bioequivalence. Since the "real-world" population would also have pre-sensitization to PEG, it would not be valid to pre-screen subjects for exclusion of subjects who are positive for pre-existing PEG-reactive antibodies from the study.

Other patient-related factors can interfere with detectability of treatment-emergent ADA in some therapeutic settings. For example, rheumatoid arthritis patients tend to have elevated levels of Rheumatoid Factors of different isotypes that can generate false positive signals in ADA assay that use a bridging format. A similar problem of false positive ADA signals can occur in samples collected oncology subjects treated with trastuzumab or bevacizumab, which can contain substantially elevated levels of the target antigen bound to the drug.

If differences in the immune response are observed at the treatment-group level, potentially confounding patient-related factors should be thoroughly analysed, particularly imbalances in exposure to concomitant medications with a known immune-modulatory effect, e.g. methotrexate and corticosteroids.

Because the T-dependent immune response is influenced by genetic polymorphism, the distribution of HLA haplotypes within a treatment group could potentially impact the incidence and magnitude of the detected immune response. Although all of these potentially confounding effects effect would be expected to be evenly distributed across the respective randomized treatment arms in a comparative clinical study, they could play a confounding role in the assessment of relative immunogenicity of a product that is associated with a rare but severe outcome.

Inferences

- Performance of clinical studies in multiple therapeutic settings for biosimilar versions of adalimumab, infliximab or rituximab revealed a consistent pattern of clinically meaningful immunogenicity, despite differences in detected ADA incidence between different populations.
- These results indicate that, at least for these products, it is valid to extrapolate conclusions about the relative impact of immunogenicity of different

(continued)

product versions across the different authorized therapeutic indications of the originator.

- If clinical studies are performed in multiple clinical populations in support of biosimilarity, differences in bioanalytical biases (e.g. drug or target interference), dose regimen and confounding patient-related variables (e.g. level of immune-suppression, genotype and co-morbidities) require careful analysis for potential impact on extrapolation of conclusions concerning relative immunogenicity of the biosimilar vs. reference products.

Switching and Interchangeability

Difference in Meaning Between EU and US

In EU, interchangeability means the medical practice based on scientific considerations of changing one medicine for another that is expected to achieve the same clinical effect in a given clinical setting and in any patient on the initiative, or with the agreement of the prescriber. The decision by the treating physician to exchange one medicine with another medicine with the same therapeutic intent in a given patient is referred to as switching. Automatic substitution of biosimilars is a political and administrative decision of dispensing one medicine instead of another equivalent and interchangeable medicine at the pharmacy level without consulting the prescriber (Kurki et al. 2017)

In the US legislation, interchangeability means a new category of biosimilars. Once a product has been licensed as a biosimilar it can be granted a licence as an interchangeable biosimilar provided that it can be expected to produce the same clinical result as the reference product in any given patient when the product 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 product and its reference product should not be greater than the risk of using the reference product without such alternation or switch. An interchangeable product may be substituted for the reference product without the intervention of the health care provider who prescribed the reference product (FDA 2017a). It is obvious that increased immunogenicity is the concern behind the requirement of multiple switches.

Biosimilarity

FDA guidance on the evaluation of immunogenicity for the *biosimilarity* exercise advises that:

> depending on the clinical experience of the reference and proposed products (taking into consideration the conditions of use and patient population), a sponsor may need to evaluate a subset of patients to provide a substantive descriptive assessment of whether a *single cross-over from the reference product to the proposed biosimilar* would result in a major risk in terms of hypersensitivity, immunogenicity, or other reactions.

As a consequence of this recommendation, therapeutic equivalence studies for some biosimilar candidates have included a *single transition* at the end of the double-blind treatment period, in which patients treated with the reference product were transitioned to receive the biosimilar version; this enabled comparison of immunogenicity with patients who continued on treatment with the biosimilar during an open-label extension period (Moots et al. 2017). An example was the Remsima (CT-P13) program, comprising comparative clinical studies in rheumatoid arthritis and in ankylosing spondylitis.

PLANETAS Study: CT-P13 in Ankylosing Spondylitis

In the main PLANETAS study, patients received nine infusions of CT-P13 (CELLTRION, Incheon, Republic of Korea) or the infliximab RP (Janssen Biotech, Horsham, Pennsylvania, USA). After the ninth infusion of PLANETAS study treatment (given at week 54), eligible patients could choose to continue into an open-label extension study during which patients and physicians were blinded to the treatment received in the main study. Patients in the extension study received an additional six infusions of CT-P13 given every 8 weeks from week 62 to week 102. CT-P13 was administered via 2 h intravenous infusion at a dose of 5 mg/kg (Park et al. 2017).

For assessment of immunogenicity, the proportion of patients with ADA was assessed at baseline and weeks 14, 30, 54, 78 and 102 using an electrochemiluminescent immunoassay method (Kim et al. 2015). The proportions of ADA positive patients at each sample time-point are summarized in Fig. 19.5.

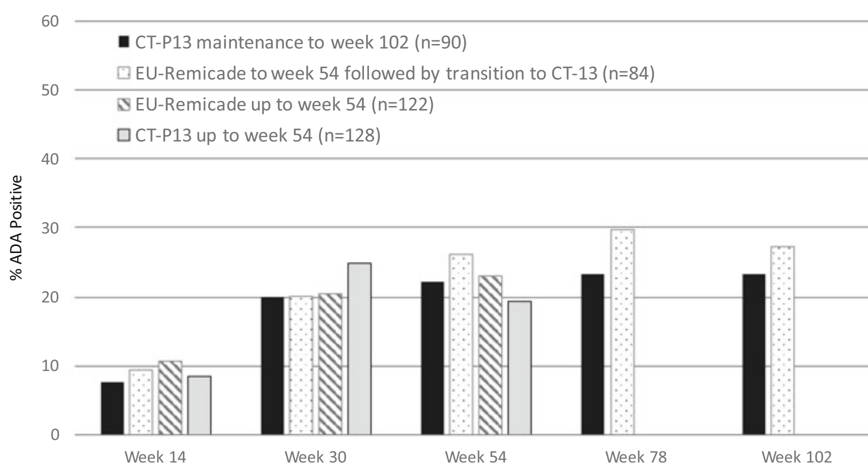


Fig. 19.5 ADA response during 54 weeks of main study and the 48-week extension period of PLANETAS study, with and without single transition from reference product to CT-P13 at week 54. Data extracted from Park et al. (2017) and from Celltrion Briefing Document for FDA Advisory Committee Meeting dated 9 February 2016 (Celltrion 2016)

Table 19.10 Comparison of efficacy for maintenance vs. switch sub-groups at week 102 of PLANETAS study

Efficacy parameter	Group	Responder n/N (%)	Odds Ratio[a]	95% CI of odds ratio	p value[b]
ASAS20	Maintenance[c]	67/83 (80.7)	1.25	0.58–2.70	0.506
	Switch[d]	60/78 (76.9)			
ASAS40	Maintenance	53/83 (63.9)	1.09	0.57–2.07	0.672
	Switch	48/78 (61.5)			
ASAS PR[e]	Maintenance	16/83 (19.3)	0.80	0.37–1.72	0.275
	Switch	18/78 (23.1)			

Data extracted from Park et al. (2017)
ASAS assessment of Spondyloarthritis International Society, *ASAS20* 20% response according to the ASAS International Working Group criteria for improvement, *ASAS40* 40% response according to the ASAS International Working Group criteria for improvement, *n* number of patients with response, *N* number of patients in group, *PR* partial remission, *RP* reference product

[a]The OR was estimated using a logistic regression model with treatment as a fixed effect, and region and baseline BASDAI score as covariates. An OR of >1 indicates increased odds in favour of the maintenance group

[b]The p value was calculated using the Hosmer–Lemeshow test for the goodness-of-fit of the logistic regression model. The test is significant at the 5% level

[c]Patients treated with CT-P13 during the 54 weeks of the main study and the 48-week extension study

[d]Patients treated with RP during the 54 weeks of the main study and then switched to CT-P13 during the 48-week extension study

[e]PR was defined as a value of <20 on a 0–100 scale in each of the following four domains: patient global assessment, pain, function and inflammation

Although there were some numerical differences in ADA frequency between the switch sub-group relative to those maintained on CT-P13, between-group differences did not reach statistical significance at any time point. In particular, ADA incidence did not increase from week 54 to 102 (26.2% vs 27.4%) in the patients treated with Reference product in the main study and then switched to CT-P13 treatment in the 48-week extension period. There was no difference in efficacy at any time-point; treatment response at the end of the extension period (week 102) is compared for the maintenance versus switch groups is summarized in Table 19.10.

PLANETRA Study: CT-P13 in Rheumatoid Arthritis

In the main study, patients received nine infusions of CT-P13 (CELLTRION, Incheon, Republic of Korea) or the infliximab RP (Janssen Biotech, Horsham, Pennsylvania, USA). After study treatment in PLANETRA, eligible patients could choose to continue in the open-label extension study. However, patients and physicians continued to be blinded to the treatment that the patient had received during the main study. All patients participating in and completing this extension study received six infusions of CT-P13 from week 62 to week 102. During the whole study period, CT-P13 was administered via 2 h intravenous infusion at a fixed dose of 3 mg/kg (Yoo et al. 2017). The proportions of ADA positive patients at each sample time-point are summarized in Fig. 19.6.

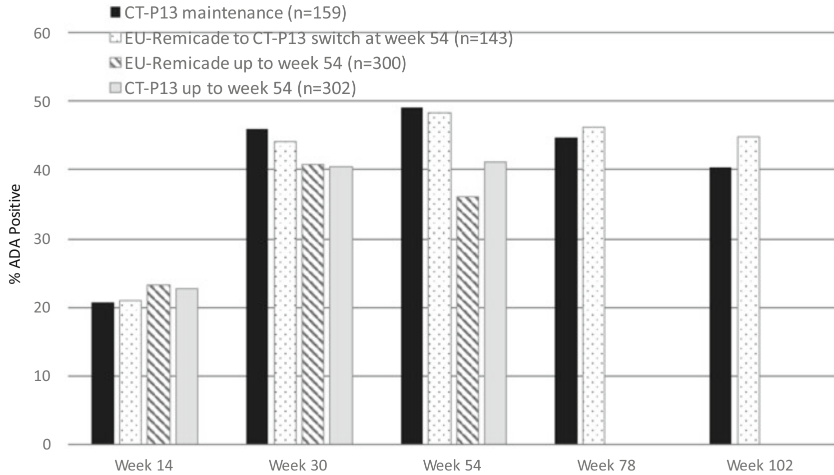


Fig. 19.6 ADA response during 54 weeks of main study and the 48-week extension period of PLANETRA study, with and without single transition from reference product to CT-P13 at week 54

A single switch from the Reference product to CT-P13 was not associated with an increase in ADA frequency at week 78 or week 102 compared to week 54. In addition, efficacy was consistent across the treatment groups throughout the whole study.

Other Studies

Impact of switching on immunogenicity was also evaluated for a second infliximab biosimilar product, SB2, using a different study design that included comparison of the switched group with both the ongoing reference product (US-Remicade) and biosimilar groups (Smolen et al. 2018). This study was fully blinded throughout, and a total of 396 subjects were re-randomized into the 24-week extension period. Among patients who were negative for antidrug antibodies (ADA) up to week 54, newly developed ADAs were reported in 14.6%, 14.9% and 14.1% of the US-Remicade/SB2 (n = 94), US-Remicade/US-Remicade (n = 101) and SB2/SB2 (n = 201) groups, respectively.

The NOR-SWITCH phase 4 trial evaluated immunogenicity associated with switching of Remicade-treated patients to CT-P13, using a randomised, double-blind design to compare non-inferiority (patients switched to CT-P13 vs. patients maintained on Remicade) with respect to worsening of disease during 52-week post-switch follow-up period follow-up (Jørgensen et al. 2017). The per-protocol set for analysis comprised 408 patients: 202 in the Remicade group and 206 in the CT-P13 group; these included patients with Crohn's disease, ulcerative colitis, spondyloarthritis, rheumatoid arthritis, psoriatic arthritis, and chronic plaque

psoriasis. The authors concluded that "there was no suggestion of differences in safety or immunogenicity between the two treatment groups".

The DANBIO registry monitored disease activity and drug trough levels in 802 patients with inflammatory arthritis (rheumatoid arthritis, psoriatic arthritis and axial spondyloarthritis), prior to and following a non-medical switch from Remicade (median treatment >6 years) to Remsima (CT-P13). In switchers with two or more available blood samples, 269/330 (81%) had unaltered sIFX level and 276/330 (84%) had unaltered ADA status during follow-up. ADAs developed in 13/330 (4%) and disappeared in 30/330 (9%). Overall, few changes in sIFX level and ADAs were observed, and they were not statistically significantly associated with CT-P13 withdrawal (Glintborg et al. 2018).

In a prospectively designed, open-label phase 4 study, serum concentrations of infliximab 16 weeks after switching to CT-P13 were non-inferior to those at baseline in patients with stable ulcerative colitis and Crohn's disease (Strik et al. 2018). Overall, 120 consecutive patients with inflammatory bowel disease were recruited: 59 with ulcerative colitis and 61 with Crohn's disease. 46 patients with ulcerative colitis and 42 patients with Crohn's disease comprised the per-protocol population. The geometric mean ratio of serum infliximab concentrations at week 16 (CT-P13) compared with those at baseline (originator) was 110.1% (90% CI: 96.0–126.3) in patients with ulcerative colitis and 107.6% (97.4–118.8) in those with Crohn's disease. In both cases, the lower bound of the 90% CI was higher than the pre-specified non-inferiority margin of 85%.

A retrospective cohort study conducted which compared immunogenicity associated with switching of Remicade to CT-P13 in adult patients with an established diagnosis of Crohn's disease or ulcerative colitis demonstrated a high degree of cross-reactivity of treatment-emergent ADA (i.e. induced by either Remicade or CT-P13) with unlabelled Remicade and CT-P13 and SB2. Both the specificity and magnitude of the ADA responses to the reference products and the biosimilar appeared equivalent (Fiorino et al. 2018).

Rheumatoid arthritis patients competing the randomized, double-blind phase of the therapeutic equivalence study for the biosimilar etanercept, SB4, either continued treatment with SB4 (n = 126) or were switched from the reference product to SB4 n = 119) for an open-label treatment period of 48 weeks duration. One patient in each treatment group developed non-neutralising ADAs after week 52. Both patients had a low titre, and the ADAs did not affect efficacy (Emery et al. 2017). In the psoriasis setting, no ADAs were reported during the multiple switching period for another etanercept biosimilar, GP2015 (Erelzi™), suggesting that multiple switches between GP2015 and the reference product do not adversely impact immunogenicity.

Caution needs to be exercised to avoid misinterpretation of causality of isolated cases of treatment failure or adverse events following a switch between product versions, particularly if there has been no monitoring of either ADA or drug levels in a subject prior to and following a medication switch (Scherlinger et al. 2017). Clinical manifestations suggestive of hypersensitivity reactions may be primarily related to the immune response induced by the prior treatment, and these levels

might not have reached their maximal amplitude at the time of the switch; thus, treatment with *either* version of the product could have induced the negative outcome.

Literature Reviews

The efficacy, safety and immunogenicity of switching between biosimilars and their reference products have been reviewed by two recent reviews dealing with essentially the same switch studies. The review of Cohen et al. (2018), which covered 57 studies and provided a view from generic industry, concluded that "these results provide reassurance to healthcare professionals and the public that the risk of immunogenicity-related safety concerns or diminished efficacy is unchanged after switching from a reference biologic to a biosimilar medicine. Another review written by McKinnon et al. (2018), which was sponsored by pharmaceutical companies under biosimilar threat, also did not reveal any significant efficacy or safety issues.

Experts expressing opinions from the perspective of biosimilars developers appear to support switching, whereas experts receiving funding from by pharmaceutical companies with originator products point to perceived gaps in knowledge, such as immunogenicity of repeated switched and interchangeability of different biosimilar versions of the same reference product.

Inferences

- Despite multiple randomized controlled studies of different biosimilar products—mainly involving infliximab, which is associated with a relatively high incidence of clinically impactful immunogenicity—the transition or switching of patients from treatment with the reference product to a biosimilar version has yet to reveal a negative impact on immunogenicity or treatment outcomes.
- The rigorous regulatory standards applied to assess analytical similarity have proven effective to control risk of incremental immunogenicity associated with minor differences in product quality that are a common feature of both originator products and their biosimilar versions.

Interchangeability in US

In addition to being biosimilar, an *interchangeable* product is expected to produce the same clinical result as the reference product in any given patient. In the case of a product that is administered to a patient more than once, FDA draft guidance indicates that a manufacturer will need to provide data and information to evaluate the risk, in terms of safety and decreased efficacy, of alternating or switching between the biosimilar products (FDA 2017a). While providing examples

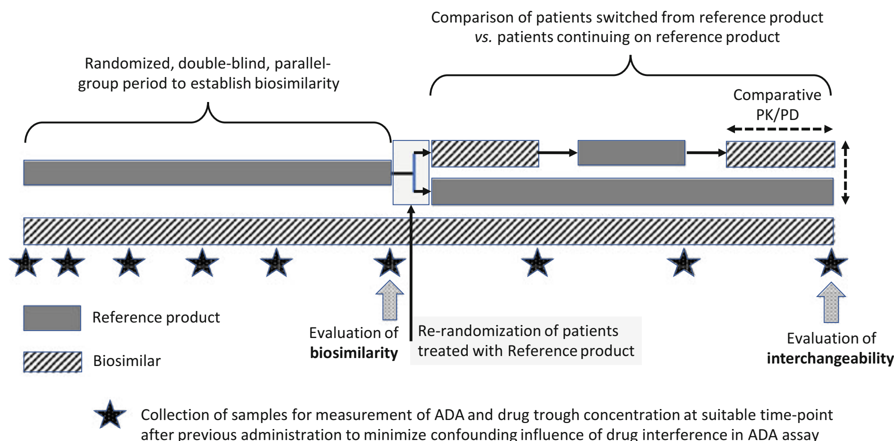


Fig. 19.7 Potential design to establish interchangeability using patients previously treated in therapeutic equivalence study for biosimilarity exercise (adapted from the 'Integrated Study Design' described in Section VI.A.2.b of FDA 2017a)

of recommended study designs for the demonstration of interchangeability, the draft guidance indicates that alternative approaches may be justified.

A possible design for an interchangeability study of a reference product and biosimilar is illustrated in Fig. 19.7.

This design uses the patients treated in the parallel-group comparative phase of the therapeutic equivalence study, and re-randomized the patients in the reference product arm to receive either alternating treatment with the biosimilar and reference product, or to continue treatment with the reference product; patients treated with the biosimilar in the first part of the study continue to receive the biosimilar in the second part. FDA suggests that the primary comparison will be the PK (and, if available, PD) endpoints, derived from intensive sampling following the third switch, for the switch vs. continuing reference product arms. Preference is given to PK parameters, on the basis that these represent the most sensitive test of a potentially clinically relevant impact. FDA draft guidance states:

> An integrated study needs to be adequately powered to evaluate the appropriate endpoint(s) to support the demonstration of no clinically meaningful differences for biosimilarity, where the primary comparison is between the proposed product arm and the reference product arm. In addition, the study needs to be adequately powered to evaluate pharmacokinetics and pharmacodynamics (if available), following the last switch to support a demonstration of interchangeability, where the primary comparison is between the switching arm and the non-switching reference product arm.

From the immunogenicity perspective, samples for ADA (and drug trough concentration) testing could be collected just prior to each switch—with the actual timing depending on the treatment regimen and knowledge of the time-course of the immune response to the particular product. This would enable a comparison of the PK parameters with the ADA response with and without multiple treatment

switches. However, there are a number of important potentially practical limitations associated with this approach, including the large number of subjects that might need to be evaluated to detect low incidence events (Inotai et al. 2017). Since sequential sampling of patients for ADA and drug concentration testing before and after switching would effectively mean that individual subjects could serve as their own controls, it may not be essential to include non-switching control arms, thereby reducing the number of patients required (Strik et al. 2018).

Even if interchangeability studies were feasible, important questions about how to interpret the results would remain, including:

- Is it valid to use the aggregate results for the respective populations to infer an expectation for equivalent therapeutic effect following switching *in any individual subject*?
- What is a clinically relevant margin for an acceptable difference—which could depend *inter alia* on the product risk profile, bioanalytical methodology and the relationship between PK and efficacy)?
- How to deal with the influence of concomitant medication or comorbidities (pro-inflammatory state, disease activity) on the immune response in individual subjects?
- Validity of extrapolation of conclusions about clinical significance of possible differences between the switch and control arms for the "real-world" setting, i.e. multiple confounding factors such as inter- and intra-subject *and* batch-to-batch product variability?

From the scientific perspective, there really is not a compelling case that switching studies are needed—or even able—to support a conclusion of an expectation for equivalent therapeutic effect following switching *in any individual subject*. Aside from feasibility and data interpretation considerations for detecting a change in clinically meaningful immunogenicity associated with switching of medications, there is the fundamental question about utility of interchangeability designation for public health policy and societal benefit relative to resources needed to implement of effective control systems, including traceability of the identity (unique product name and batch number) of the medicine administered (Ebbers and Chamberlain 2014; Minghetti et al. 2015; Inotai et al. 2017). Ultimately, even if interchangeability could be demonstrated to an extent that met the US statutory standard, this would not mean that a product *will* produce the same clinical result as the reference product in every patient (Ebbers and Chamberlain 2014). For this reason, a state-of-the-art demonstration of biosimilarity, allied to an appropriate level of post-marketing surveillance, traceability and physician supervision, could represent the most realistic and effective strategy for mitigation a *potential* risk of a negative impact of switching on immunogenicity of biosimilar medicines (Kurki et al. 2017; Braun and Kudrin 2016).

Inferences

- The EU and US have adopted different strategies in dealing with a potentially altered risk of immunogenicity associated with switching between the biosimilar and its reference product. Several EU Member States encourage switching on the basis of scientific considerations and clinical switching studies that have provided the proof of concept. In the US, the legislation requires systematic switching studies aiming at automatic substitution.
- From the immunogenicity perspective, there are many scientific and feasibility reasons why additional regulatory requirements to demonstrate interchangeability of biosimilars according to US statutory provisions are most unlikely to yield definitive conclusions that justify either the resources required or the benefit for the patient and physician.
- Current regulatory standards applied in EU and US are effective to mitigate a perceived risk of enhanced immunogenicity of different originator and biosimilar products; while this does not exclude the *possibility* of enhancement of immunogenicity that cannot be detected in pre-approval controlled clinical studies, the evidence accumulated since April 2006 for biosimilar products authorized in EU (as of 31 March 2018, EMA has issued a positive opinion for marketing authorization of biosimilar versions of 14 different active substances (http://www.ema.europa.eu/), has shown a highly consistent profile of no incremental risk of clinically meaningful immunogenicity relative to the reference product—evaluated both in directly comparative, randomized controlled clinical studies, including transition or switch designs, and by post-authorization experience in the "real-world" setting.
- Traceability of different product versions administered in clinical practice is critical to assure effectiveness of supervision by the prescribing physician and for ongoing pharmacovigilance.

Post-authorization Risk Mitigation

Control of Risks for Originator Products

Following initial marketing authorization, the manufacturing process of original biopharmaceutical products will invariably be modified to introduce different manufacturing facilities or raw materials, to improve efficiency or introduce different formulations and presentations (Vezér et al. 2016). Some of these changes, e.g. introduction of a new cell line for fermentation or change in the qualitative or

quantitative composition of the drug product formulation, can be associated with changes in the heterogeneity of product quality attributes (Federici et al. 2013). In rare cases, changes have resulted in enhanced immunogenicity of both an originator version (Casadevall et al. 2005) and candidate biosimilar (epoetin-alfa), and follow-up investigations (Bennett et al. 2004; Seidl et al. 2012; Rubic-Schneider et al. 2017) linked the increased risk to a particular combination of variables that were subsequently controlled—both for the originator and the approved biosimilar version—more effectively.

In the case of a "high risk" manufacturing change scenario where there is uncertainty about the impact of the change on immunogenicity, comparative clinical studies have been performed to evaluate clinically impactful immunogenicity prior to implementation of the manufacturing change: examples are a change in cell line for Aranesp (EMA 2008), and introduction of a new formulation of Humira (FDA 2015c); these examples did not reveal an increase in ADA formation or negative clinical impact for the new product versions.

In the vast majority of post-approval manufacturing change scenarios, reliance is placed on analytical and non-clinical comparability data to control risks (ICH 2005). Although some published reports are available that attest to remarkable consistency of many product quality attributes over an extended number of manufacturing campaigns for different products (Tebbey et al. 2015; Hassett et al. 2018), there are some potentially important gaps in the information provided: for example, comparative data on the extent of analysis of sub-visible particles—recognized as a risk factor for enhanced immunogenicity (Carpenter et al. 2009)—is often not presented for the originator products; and it has been the comparative evaluation for biosimilarity that has revealed apparently higher levels of sub-visible particle content in batches of the originator product (Cho et al. 2016). Even for apparently rigorously controlled originator products, notable shifts in glycosylation levels have been reported for different batches placed on the market (Schiestl et al. 2011)—although relationship to immunogenicity risk remains uncertain.

Overall, the regulatory approach (ICH 2005) to assessing comparability (pre- *vs*. post-change) of product quality of biopharmaceutical products *has* proven to be suitably rigorous to control immunogenicity-related risks, despite the changes in the heterogeneity profiles associated with the wide spectrum of post-authorization manufacturing changes. The same regulatory standard (ICH 2005) is applicable to originator and biosimilar products; in the case of biosimilar products, re-establishment of biosimilarity following a change in manufacture is not required.[2]

[2]But, note that FDA expectations for evaluating impact of major manufacturing changes for interchangeable products to ensure continuing compliance with US statutory requirements for interchangeability are yet to be defined.

Uncertainty Associated with Independent Post-Authorization Life-Cycles of Different Biosimilar Versions of the Same Therapeutic Protein

Although each biosimilar version is authorized on the basis of a rigorous comparison with a reference product, manufacturing changes and new formulations/presentations could be introduced during the post-authorization life-cycle, subject to meeting the regulatory standard (ICH 2005) for comparability of the pre- *vs*. post-change versions. The fact that there is no requirement to re-establish biosimilarity with the reference product implies a theoretical risk that there could be a "drift" in the heterogeneity of product-related variants and/or process-derived impurities.

As pointed out earlier, published data describing the consistency of product profile in relation to clinically impactful immunogenicity are lacking, both for originator versions and for biosimilars. The most obvious example of an influence of a post-authorization manufacturing change on immunogenicity is the case of an increased incidence of antibody-mediated Pure Red Cell Aplasia, associated with modifications to the formulation-primary container combination of an originator version of erythropoietin; the risk for the originator product was effectively mitigated by the implementation of improved product quality control provisions, and the experience gained was then taken into account for regulatory approval of biosimilar candidates.

From the immunogenicity risk perspective, rigorous control of molecular size variants of the therapeutic protein as well as process-derived impurities within a suitable range—defined by levels measured in drug product batches that were used in controlled randomized clinical studies—could represent the most effective strategy for avoiding a clinically impactful drift between different versions of the same product. This would necessitate application of a panel of suitably sensitive analytical techniques for testing of multiple batches of drug product, both at the time of produce release and following storage under "worst-case" handling conditions (e.g. comparative stability of mechanically or thermally stressed biosimilar *vs.* reference drug products). In vitro cell-based assays that are highly sensitive to innate immune response modulating impurities (IIRMIs) might also be used to provide supportive risk evaluation data to guide manufacturing process development (Verthelyi and Wang 2010; Haile et al. 2015, 2017).

The authors have noted that the innovator industry has started to emphasize "shifting and drifting" as a problem for interchangeability of biosimilars. The goal is to mandate comparisons to the original product whenever there is a major change in the manufacturing process of either product. However, in our view, comparability requirements upon the change of manufacturing process should be tightened for both innovators and biosimilars if drifting and shifting are regarded as significant risks.

Inferences

- As for the originator products that have been on the market many years, and which have undergone multiple manufacturing changes, neither a risk of change in heterogeneous product quality profile nor a risk of enhanced immunogenicity can be *definitively* excluded for manufacturing or formulation changes to approved biosimilar products.
- Nevertheless, the regulatory assessment of comparability has proven effective for controlling risks of incremental immunogenicity—without requiring directly comparative clinical evidence of similar immunogenicity prior to authorization of the majority changes.
- Current regulatory procedures include provision for comparative (pre- *vs*. post-change) clinical data to assess relative immunogenicity in "high-risk" scenarios, applicable to both originator and biosimilar versions.
- To mitigate a theoretical risk of divergent profiles of different versions of the same therapeutic protein, sensitive analytical techniques can be applied to monitor and control levels of quality attributes associated with risk of incremental immunogenicity, and definition of acceptable control ranges for these attributes linked to historical clinical data.

Future Perspectives

In the light of scientific considerations and experience gained to date, which does not reveal any increased risks associated with the immunogenicity of biosimilar products that have been authorized by EMA (since April 2006) and FDA (since March 2015), the analytical similarity exercise appears sufficiently rigorous to control the potential risk to patients. However, since biosimilarity is defined by demonstration of lack of *clinically* meaningfully differences, and different factors may interact in an unpredictable manner to increase risk, clinical evidence of relative immunogenicity is required by regulatory authorities—as well as being expected by physicians who have responsibility for prescribing these medicines. The extent of this clinical evidence needs to reflect the severity of consequences of immunogenicity, which varies widely between different products, as well as the feasibility of detecting low incidence events. Then, the approval of multiple biosimilar versions of the same originator product raises additional questions about relative risks associated with versions that have not been directly compared in controlled clinical studies, and which are subject to independent manufacturing life-cycles. The authors personal opinions about how some of these questions might be addressed are summarized below.

Would it be feasible to define acceptable levels of individual product variants and process-derived impurities to exclude a risk for enhanced immunogenicity?

Attempts to establish causal relationships in any biological system are prone to almost certain failure. This is particularly true of immunogenicity-related risks, for which a confluence of factors may be involved on provoking adverse events (Bennett et al. 2004). Part of this complexity is related to the inherent heterogeneity of biological medicinal products, and part is related to the diversity of immune response elements that vary across individual subjects; how these different factors impact on the fine balance between the innate and adaptive immune systems and immune tolerance mechanisms cannot be predicted.

That said, individual product quality attributes can be controlled to measurable levels, and their composite impact on the immune response can be evaluated in controlled clinical studies. By standardizing the performance (specificity, sensitivity, precision, relative accuracy etc.) of analytical and bioanalytical methods, substantial progress has been achieved in the scientific evaluation and control of risks associated with individual product variants and process-derived impurities. In the worst-case scenario, i.e. that of antibody-mediated pure red cell aplasia, different factors contributing to enhanced immunogenicity were identified (Bennett et al. 2004; Seidl et al. 2012; Rubic-Schneider et al. 2017) and then subsequently controlled. As discussed earlier in the chapter, clinical evaluation indicated a lower frequency of transient ADA formation for two different biosimilar etanercept products relative to the originator product; however, these apparent differences in transient ADA formation did not translate into clinically meaningful differences in immunogenicity. This underlines the difficulty for regulators to define acceptable levels of individual product quality attributes. Moreover, from an ethical perspective, it would not be acceptable to perform a clinical comparison of the relative immunogenicity of a drug product containing, for example, a range of artificially elevated levels of aggregates.

Could in vitro studies be useful for comparison of relative immunogenicity of biosimilars?

In vitro assays of innate and adaptive immune responses can be useful for comparative purposes, e.g. for application at the lead candidate screening stage, for monitoring process-related impurities or for assessing stability of the active substance within a formulation development exercise (Jawa et al. 2013; Haile et al. 2015; Joubert et al. 2016; Schultz et al. 2017). While sensitivity and precision of these assays may not be adequate to enable *reliable* comparison for the purpose of a biosimilarity exercise (Talotta et al. 2017), in vitro T-cell assays were useful in a root-cause investigation of amPRCA following rhEPO administration (Rubic-Schneider et al. 2017).

In vitro "precision profiling" of the epitopes recognized by ADA in sera obtained from patients treated with infliximab or adalimumab has been reported (Homann et al. 2017), but relevance for detecting differences between different versions of the same active substance is unknown. In the case of infliximab, there is already strong evidence from ligand-binding assays for the equivalent specificity of ADA induced by a biosimilar version compared to the reference product (Ben-Horin et al. 2015; Reinisch et al. 2017; Fiorino et al. 2018). Isolation of B-cell clones from two adalimumab-treated did reveal a clonally diverse response targeting the antigen-binding regions, suggestive of small differences in fine specificity of the humoral

immune response (van Schouwenberg et al. 2016). Overall, although in vitro assays may facilitate the development process of a biosimilar candidate, they are unlikely to replace clinical evidence for the assessment of relative immunogenicity versus the reference product.

Might comparative clinical studies of up to 6 months duration provide adequate pre-approval evidence of similar immunogenicity profile, if supported by post-approval monitoring of drug levels and sustainability of treatment responses?

Arguably, yes for many products that are not associated with severe consequences related to immunogenicity; but, not in the case of a product that has an identified risk of clinically impactful outcomes, e.g. darbepoetin or natalizumab, or for products that could have a delayed effect due to an intended immune-suppressive mode of action, e.g. abatacept.

Accumulated evidence from the chronic administration studies performed to date for immunogenic biosimilar candidates such as infliximab and adalimumab indicates that the peak ADA response (both in terms of frequency and titer) was attained by the 6-month treatment time-point; extension studies, including transition from the reference product to the biosimilar version did not reveal any incremental immunogenicity. In these cases, a 6-month study duration was sufficient to make an informed decision for approval; and, if deemed necessary due to uncertainty about longer-term outcomes, a prospectively designed, "low interventional"-type study (i.e. with periodic monitoring of drug levels and ADA) could be performed in the post-authorization setting.

Could interchangeability studies contribute to controlling risks associated with switching between the reference product and a biosimilar version?

Based on the practical issues in performing studies designed to meet the FDA statutory requirements for interchangeability, allied to difficulties in drawing definitive conclusions to exclude an increased immunogenicity-related risk at the individual subject level based on aggregated results, it seems most unlikely that such studies could be instructive from the scientific perspective—although they might provide re-assurance to prescribing physicians and patients that there is no incremental risk associated with switching.

Are there additional risks in switching between different biosimilar versions containing the same active substance and, if so, how might these be mitigated?

This question often stems from the debate around interchangeability of biosimilars and originator products, and seems to be driven by a perception—rather than by scientific evidence—that different biosimilars of the same active substance could be associated with increased immunogenicity.

If the regulatory review process for marketing authorization of biosimilars has proven effective for controlling the pertinent risks, *and* there is an effective regulatory system for managing risks associated with manufacturing changes to these products (including ICH Q5E provisions for demonstration comparability and post-approval pharmacovigilance), it would be illogical to perceive a higher risk in switching between different biosimilar versions that have been regulated to equivalent standards.

However, effective management of the risk depends on adequate traceability of the product versions administered, allied to effective pharmacovigilance systems for monitoring adverse outcomes that could be associated with increased immunogenicity of particular products, (including the originator as well as multiple biosimilar versions). Thus, if signals are detected in the post-approval phase, regulatory authorities do have the power to request additional clinical data or even suspend or withdraw a marketing authorization.

Certainly, there is always a *potential* risk: but decisions to increase or decrease the rigor of the control systems should be driven by empirical evidence, rather than theoretically-based argumentation.

References

Ahmadi M, et al. Small amounts of sub-visible aggregates enhance the immunogenic potential of monoclonal antibody therapeutics. Pharm Res. 2015;32(4):1383–94.

Amgen. Neupogen prescribing information, revised 06/2016. 2016a. Available at https://pi.amgen.com/~/media/amgen/repositorysites/pi-amgen-com/neupogen/neupogen_pi_hcp_english.pdf. Last Accessed 16 April 2018.

Amgen. Amgen Briefing Document for Arthritis Advisory Committee meeting for ABP 501 held on 12 July 2016. 2016b. Available at https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM510295.pdf. Last Accessed 10 May 2018.

Amgen. Neulasta prescribing information revised 12/2017. 2017. Available at https://pi.amgen.com/~/media/amgen/repositorysites/pi-amgen-com/neulasta/neulasta_pi_hcp_english.pdf. Last Accessed 16 April 2018.

Amgen. Aranesp prescribing information revised 01/2018. 2018. Available at https://pi.amgen.com/~/media/amgen/repositorysites/pi-amgen-com/neupogen/neupogen_pi_hcp_english.pdf. Last Accessed 16 April 2018.

Armstrong JK, Hempel G, Koling S, Chan LS, Fisher T, Meiselman HJ, Garratty G. Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients. Cancer. 2007;110(1):103–11.

Ataca P, Atilla E, Kendir R, Bavbek S, Ozcan M. Successful desensitization of rituximab hypersensitivity. Case Rep Immunol. 2015;2015:524507.

Baker MP, Jones TD, Chamberlain P. Immunogenicity of biologics. In: Vaughan T, Osbourn J, Jallal B, editors. Protein therapeutics. Weinheim: Wiley; 2017. https://doi.org/10.1002/9783527699124.ch12.

Barger TE, Wrona D, Goletz TJ, Mytych DT. A detailed examination of the antibody prevalence and characteristics of anti-ESA antibodies. Nephrol Dial Transplant. 2012;27(10):3892–9. https://doi.org/10.1093/ndt/gfs392.

Bartelds GM, Wijbrandts CA, Nurmohamed MT, Stapel S, Lems WF, Aarden L, Dijkmans BA, Tak PP, Wolbink GJ. Clinical response to adalimumab: relationship to anti-adalimumab antibodies and serum adalimumab concentrations in rheumatoid arthritis. Ann Rheum Dis. 2007;66(7):921–6.

Bartelds GM, Krieckaert CL, Nurmohamed MT, van Schouwenburg PA, Lems WF, Twisk JW, Dijkmans BA, Aarden L, Wolbink GJ. Development of antidrug antibodies against adalimumab and association with disease activity and treatment failure during long-term follow-up. JAMA. 2011;305(14):1460–8. https://doi.org/10.1001/jama.2011.406.

Beart F, Noman M, Vermeire S, Van Assche G, D'Haens G, Carbonez A, Rutgeerts P. Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. N Engl J Med. 2003;348:601–8.

Bendtzen K, Geborek P, Svenson M, Larsson L, Kapetanovic MC, Saxne T. Individualized monitoring of drug bioavailability and immunogenicity in rheumatoid arthritis patients treated with the tumor necrosis factor α inhibitor infliximab. Arthritis Rheum. 2006;54(12):3782–9.

Ben-Horin S, Yavzori M, Benhar I, et al. Cross-immunogenicity: antibodies to infliximab in Remicade-treated patients with IBD similarly recognise the biosimilar Remsima. Gut. 2015;65:1132–8.

Bennett CL, Luminari S, Nissenson AR, Tallman MS, Klinge SA, McWilliams N, McKoy JM, Kim B, Lyons EA, Trifilio SM, Raisch DW, Evens AM, Kuzel TM, Schumock GT, Belknap SM, Locatelli F, Rossert J, Casadevall N. Pure red-cell aplasia and epoetin therapy. N Engl J Med. 2004;351(14):1403–8.

Bloem K, van Leeuwen A, Verbeek G, Nurmohamed MT, Wolbink GJ, van der Kleij D, Rispens T. Systematic comparison of drug-tolerant assays for anti-drug antibodies in a cohort of adalimumab-treated rheumatoid arthritis patients. J Immunol Methods. 2015;418:29–38. https://doi.org/10.1016/j.jim.2015.01.007.

BMS. Orencia® Prescribing Information, June 2017, BMS. 2017. Sourced on 10 May 2018. https://packageinserts.bms.com/pi/pi_orencia.pdf.

Bourdage JS, Cook CA, Farrington DL, Chain JS, Konrad RJ. An affinity capture elution (ACE) assay for detection of anti-drug antibody to monoclonal antibody therapeutics in the presence of high levels of drug. J Immunol Methods. 2007;327(1-2):10–7.

Braun J, Kudrin A. Switching to biosimilar infliximab (CT-P13): evidence of clinical safety, effectiveness and impact on public health. Biologicals. 2016;44(4):257–66. https://doi.org/10.1016/j.biologicals.2016.03.006.

Brezski RJ, Knight DM, Jordan RE. The origins, specificity, and potential biological relevance of human anti-IgG hinge autoantibodies. Sci World J. 2011;11:1153–67. https://doi.org/10.1100/tsw.2011.107.

Carpenter JF, Randolph TW, Jiskoot W, Crommelin DJ, Middaugh CR, Winter G, Fan YX, Kirshner S, Verthelyi D, Kozlowski S, Clouse KA, Swann PG, Rosenberg A, Cherney B. Overlooking subvisible particles in therapeutic protein products: gaps that may compromise product quality. J Pharm Sci. 2009;98(4):1201–5. https://doi.org/10.1002/jps.21530.

Casadevall N, Dupuy E, Molho-Sabatier P, Tobelem G, Varet B, Mayeux P. Brief report: autoantibodies against erythropoietin in a patient with pure red-cell aplasia. N Engl J Med. 1996;334(10):630–3.

Casadevall N, Eckhardt K-U, Rossert J. Epoetin-induced autoimmune pure red cell aplasia. J Am Soc Nephrol. 2005;16:S67–9.

Celltrion. Briefing for FDA AC for CT-P13. 2016. Available at https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/ucm484858.htm. Last Accessed 16 April 2018.

Chamberlain PD. Multidisciplinary approach to evaluating immunogenicity of biosimilars: lessons learnt and open questions based on 10 years' experience of the European Union regulatory pathway. Biosimilars. 2014;4:23–43.

Cho IH, Lee N, Song D, Jung SY, Bou-Assaf G, Sosic Z, Zhang W, Lyubarskaya Y. Evaluation of the structural, physicochemical, and biological characteristics of SB4, a biosimilar of etanercept. MAbs. 2016;8(6):1136–55. https://doi.org/10.1080/19420862.2016.1193659.

Chung CH, Mirakhur B, Chan E, Le QT, Berlin J, Morse M, Murphy BA, Satinover SM, Hosen J, Mauro D, Slebos RJ, Zhou Q, Gold D, Hatley T, Hicklin DJ, Platts-Mills TA. Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. N Engl J Med. 2008;358(11):1109–17. https://doi.org/10.1056/NEJMoa074943.

Cohen HP, Blauvelt A, Rifkin RM, Danese S, Gokhale SB, Woollett G. Switching reference medicines to biosimilars: a systematic literature review of clinical outcomes. Drugs. 2018;78(4):463–78. https://doi.org/10.1007/s40265-018-0881-y.

De Groot AS, Scott DW. Immunogenicity of protein therapeutics. Trends Immunol. 2007;28(11):482–90.

Dellanna F, Goldsmith D, Krendyukov A, Seidl A, Höbel N, Combe C. HX575: established biosimilarity in the treatment of renal anemia and 10 years of clinical experience. Drug Des Devel Ther. 2017;12:9–14. https://doi.org/10.2147/DDDT.S146147.

Delluc S, Ravot G, Maillere B. Quantification of the preexisting CD4 T-cell repertoire specific for human erythropoietin reveals its immunogenicity potential. Blood. 2010;116(22):4542–5.

Ebbers HC, Chamberlain P. Interchangeability: an insurmountable 5th hurdle? GaBI J. 2014;3(2):88–92.

Ebbers HC, Chamberlain P. Controversies in establishing biosimilarity: extrapolation of indications and global labelling practices. BioDrugs. 2016;30(1):1–8.

EMA. European Public Assessment Report (EPAR) for Avastin®. 2005a. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000582/WC500029262.pdf. Last Accessed 16 April 2018.

EMA. EPAR for Herceptin®. 2005b. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000278/WC500049816.pdf. Last Accessed 16 April 2018.

EMA. Guidance on similar medicinal products containing recombinant granulocyte-colony stimulating factor, EMEA/CHMP/BMWP/31329/2005. 2006. Available at http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001386.jsp&mid=WC0b01ac058002958c. Last Accessed 16 April 2018.

EMA. Assessment Report for ARANESP, Doc. Ref No. EMEA/478499/2008. 2008. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000332/WC500026148.pdf. Accessed 15 April 2018.

EMA. Guideline on non-clinical and clinical development of similar biological medicinal products containing recombinant erythropoietins (Revision), EMEA/CHMP/BMWP/301636/2008 Corr. 2010. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/04/WC500089474.pdf. Last Accessed 16 April 2018.

EMA. EPAR for REMSIMA, June 2013, EMA/CHMP/589317/2013. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002576/WC500151486.pdf. Last Accessed 16 April 2018.

EMA. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. EMEA/CHMP/BMWP/42832/2005 Rev1. Committee for Medicinal Products for Human Use (CHMP). 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf. Last Accessed 8 January 2018.

EMA. EPAR for BENEPALI, EMA/CHMP/819219/2015. 2015. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004007/WC500200380.pdf. Last Accessed 16 April 2018.

EMA. EPAR for FLIXABI, April 2016, EMA/CHMP/272283/2016. 2016a. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004020/WC500208358.pdf. Last Accessed 16 April 2018.

EMA. EPAR for MOVYMIA, EMA/88527/2017. 2016b. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004368/WC500223669.pdf. Last Accessed 16 April 2018.

EMA. Guideline on Immunogenicity assessment of therapeutic proteins (EMEA/CHMP/BMWP/14327/2006 Rev 1) Committee for Medicinal Products for Human Use (CHMP). 2017a. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/06/WC500228861.pdf. Last Accessed 8 January 2018.

EMA. EPAR for IMRALDI, June 2017, EMA/CHMP/559383/2017. 2017b. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004279/WC500233922.pdf. Last Accessed 16 April 2018.

EMA. Variation Assessment Report for IV Orencia, June 2017, EMA/455579/2017. 2017c. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000701/WC500233148.pdf. Last Accessed 16 April 2018.
EMA. EPARs for RIXATHON EMA/303207/2017. 2017d. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/003903/WC500232462.pdf. Last Accessed 16 April 2018.
EMA. EPAR for TRUXIMA; EMA/CHMP/75695/2017. 2017e. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004112/WC500222695.pdf. Last Accessed 16 April 2018.
EMA. EPAR for GP2015 ERELZI; April 2017, EMA/CHMP/302222/2017. 2017f. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004192/WC500230144.pdf. Last Accessed 16 April 2018.
Emery P, Vencovský J, Sylwestrzak A, Leszczyński P, Porawska W, Stasiuk B, Hilt J, Mosterova Z, Cheong SY, Ghil J. Long-term efficacy and safety in patients with rheumatoid arthritis continuing on SB4 or switching from reference etanercept to SB4. Ann Rheum Dis. 2017. https://doi.org/10.1136/annrheumdis-2017-211591.
FDA. Guidance for industry: immunogenicity assessment for therapeutic protein products. 2014. Available at https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm338856.pdf. Last Accessed 16 April 2018.
FDA. Transcript of FDA Adcom for Zarzio, held on 7 January 2015. 2015a. Available at https://wayback.archive-it.org/7993/20170404153112/https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM436387.pdf. Last Accessed 10 May 2018.
FDA. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. FDA, April 2015. 2015b. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf. Last Accessed 16 April 2018.
FDA. FDA summary review for supplementary BLA (Humira 100mg/mL). 2015c. Available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/125057orig1s394sumr.pdf. Last Accessed 16 April 2018.
FDA. Guidance for industry: clinical pharmacology data to support a demonstration of biosimilarity to a reference product, FDA. 2016a. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm397017.pdf Last Accessed 8 January 2018.
FDA. FDA Briefing Document for Arthritis Advisory Committee meeting held on 9 February 2016 (Table 16) CT-P13. 2016b. Available at https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/ucm484858.htm. Last Accessed 16 April 2018.
FDA. Cross discipline team leader review for CT-P13 review. 2016c. Available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/125544orig1s000crossr.pdf. Last Accessed 16 April 2018.
FDA. FDA summary review for CT-P13 review. 2016d. Available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/125544orig1s000sumr.pdf. Last Accessed 16 April 2018.
FDA. FDA addendum to primary clinical review for CT-P13 review. 2016e. Available at https://www.fda.gov/downloads/drugs/developmentapprovalprocess/developmentresources/ucm510208.pdf. Last Accessed 16 April 2018.
FDA. FDA Backgrounder for GP2015 AC, 13 July 2016, Page 23. 2016f. Available at https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/ucm510492.htm. Last Accessed 16 April 2018.

FDA. FDA Briefing Document for Arthritis Advisory Committee meeting for ABP 501 dated 12 July 2016. 2016g. Available at https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM510293.pdf. Last Accessed 10 May 2018.

FDA. Draft Guidance for industry: considerations in demonstrating interchangeability with a reference product. 2017a. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf.

FDA. FDA clinical review for Samsung 351(k) BLA submission for SB2 (RENFLEXIS). 2017b. Available at https://www.fda.gov/downloads/drugs/developmentapprovalprocess/developmentresources/ucm557884.pdf. Last Accessed 18 April 2018.

FDA. Cross-discipline team leader review/division director summary for Samsung 351(k) BLA submission for SB2. 2017c. Available at https://www.fda.gov/downloads/drugs/developmentapprovalprocess/developmentresources/ucm557883.pdf.

Federici M, Lubiniecki A, Manikwar P, Volkin DB. Analytical lessons learned from selected therapeutic protein drug comparability studies. Biologicals. 2013;41(3):131–47. https://doi.org/10.1016/j.biologicals.2012.10.001.

Fineberg SE, Kawabata TT, Finco-Kent D, Fountaine RJ, Finch GL, Krasner AS. Immunological responses to exogenous insulin. Endocr Rev. 2007;28(6):625–52.

Fiorino G, Ruiz-Argüello MB, Maguregui A, Nagore D, Correale C, Radice S, Gilardi D, Allocca M, Furfaro F, Martínez A, Danese S. Full interchangeability in regard to immunogenicity between the infliximab reference biologic and biosimilars CT-P13 and SB2 in inflammatory bowel disease. Inflamm Bowel Dis. 2018;24(3):601–6. https://doi.org/10.1093/ibd/izx086.

Glintborg B, Kringelbach T, Bolstad N, Warren DJ, Eng G, Sørensen IJ, Loft AG, Hendricks O, Hansen I, Linauskas A, Nordin H, Kristensen S, Lindegaard H, Jensen DV, Goll GL, Høgdall E, Gehin J, Enevold C, Nielsen CH, Krogh NS, Johansen JS, Hetland ML. Drug concentrations and anti-drug antibodies during treatment with biosimilar infliximab (CT-P13) in routine care. Scand J Rheumatol. 2018;9:1–4. https://doi.org/10.1080/03009742.2017.1376110.

Haile LA, Puig M, Kelley-Baker L, Verthelyi D. Detection of innate immune response modulating impurities in therapeutic proteins. PLoS One. 2015;10(4):e0125078.

Haile LA, et al. Cell based assay identifies TLR2 and TLR4 stimulating impurities in interferon beta. Sci Rep. 2017;7(1):10490.

Harding FA, Stickler MM, Razo J, DuBridge RB. The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions. MAbs. 2010;2(3):256–65.

Hassett B, Singh E, Mahgoub E, O'Brien J, Vicik SM, Fitzpatrick B. Manufacturing history of etanercept (Enbrel®): consistency of product quality through major process revisions. MAbs. 2018;10(1):159–65. https://doi.org/10.1080/19420862.2017.1388483.

Hawe A, Romeijn S, Filipe V, Jiskoot W. Asymmetrical flow field-flow fractionation method for the analysis of submicron protein aggregates. J Pharm Sci. 2012;101(11):4129–39. https://doi.org/10.1002/jps.23298.

Homann A, Röckendorf N, Kromminga A, Frey A, Platts-Mills TA, Jappe U. Glycan and peptide IgE epitopes of the TNF-alpha blockers infliximab and adalimumab - precision diagnostics by cross-reactivity immune profiling of patient sera. Theranostics. 2017;7(19):4699–709. https://doi.org/10.7150/thno.20654.

ICH. ICH comparability of biotechnological/biological products subject to changes in their manufacturing process. Q5E. ICH, June 2005. CPMP/ICH/5721/03. 2005. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002805.pdf. Last Accessed 8 January.

Inotai A, Prins CPJ, Csanádi M, Vitezic D, Codreanu C, Kaló Z. Is there a reason for concern or is it just hype? A systematic literature review of the clinical consequences of switching from originator biologics to biosimilars. Expert Opin Biol Ther. 2017;17(8):915–26.

Jawa V, Cousens LP, Awwad M, Wakshull E, Kropshofer H, De Groot AS. T-cell dependent immunogenicity of protein therapeutics: preclinical assessment and mitigation. Clin Immunol. 2013;149(3):534–55. https://doi.org/10.1016/j.clim.2013.09.006.

Jørgensen KK, Olsen IC, Goll GL, Lorentzen M, Bolstad N, Haavardsholm EA, Lundin KEA, Mørk C, Jahnsen J. Kvien TK; NOR-SWITCH study group. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. Lancet. 2017;389(10086):2304–16.

Joubert MK, Deshpande M, Yang J, Reynolds H, Bryson C, Fogg M, Baker MP, Herskovitz J, Goletz TJ, Zhou L, Moxness M, Flynn GC, Narhi LO, Jawa V. Use of in vitro assays to assess immunogenicity risk of antibody-based biotherapeutics. PLoS One. 2016;11(8):e0159328. https://doi.org/10.1371/journal.pone.0159328.

Jung SK, Lee KH, Jeon JW, Lee JW, Kwon BO, Kim YJ, Bae JS, Kim DI, Lee SY, Chang SJ. Physicochemical characterization of Remsima. MAbs. 2014;6(5):1163–77. https://doi.org/10.4161/mabs.32221.

Karsten Roth NM, Wessels H, Hoefler J, Jankowsky R. Comparability of pharmacodynamics and immunogenicity of B12019, a proposed pegfilgrastim biosimilar to neulasta. Blood. 2017;130(Suppl 1):1002.

Kim JS, Kim SH, Kwon B, Hong S. Comparison of immunogenicity test methods used in clinical studies of infliximab and its biosimilar (CT-P13). Expert Rev Clin Immunol. 2015;11(Suppl 1):S33–41. https://doi.org/10.1586/1744666X.2015.1090312.

Kim HS, Kim I, Zheng L, Vernes JM, Meng YG, Spiess C. Evading pre-existing anti-hinge antibody binding by hinge engineering. MAbs. 2016;8(8):1536–47.

Krieckaert CL, Nurmohamed MT, Wolbink GJ. Methotrexate reduces immunogenicity in adalimumab treated rheumatoid arthritis patients in a dose dependent manner. Ann Rheum Dis. 2012;71(11):1914–5. https://doi.org/10.1136/annrheumdis-2012-201544.

Kurki P, van Aerts L, Wolff-Holz E, Giezen T, Skibeli V, Weise M. Interchangeability of biosimilars: a European perspective. BioDrugs. 2017;31(2):83–91.

Lee KH, Lee J, Bae JS, Kim YJ, Kang HA, Kim SH, Lee SJ, Lim KJ, Lee JW, Jung SK, Chang SJ. Analytical similarity assessment of rituximab biosimilar CT-P10 to reference medicinal product. MAbs. 2018;10(3):380–96. https://doi.org/10.1080/19420862.2018.1433976.

Li J, Yang C, Xia Y, Bertino A, Glaspy J, Roberts M, Kuter DJ. Thrombocytopenia caused by the development of antibodies to thrombopoietin. Blood. 2001;98(12):3241–8.

Liu YD, van Erik JZ, Flynn GC. Human antibody Fc deamidation in vivo. Biologicals. 2009;37(5):313–22.

Liu J, Eris T, Li C, Cao S, Kuhns S. Assessing analytical similarity of proposed amgen biosimilar ABP 501 to adalimumab. BioDrugs. 2016;30(4):321–38. https://doi.org/10.1007/s40259-016-0184-3.

Loumaye E, Dreano M, Galazka A, Howles C, Ham L, Munafo A. Recombinant human FSH product development group, Ares-Serono. Recombinant follicle stimulating hormone: development of the first biotechnology product for the treatment of fertility. Hum Reprod Update. 1998;4(6):862–81.

Luo Y, Lu Z, Raso SW, Entrican C, Tangarone B. Dimers and multimers of monoclonal IgG1 exhibit higher in vitro binding affinities to Fc γ receptors. MAbs. 2009;1(5):491–504.

Macdougall IC, Casadevall N, Locatelli F, Combe C, London GM, Di Paolo S, Kribben A, Fliser D, Messner H, McNeil J, Stevens P, Santoro A, De Francisco AL, Percheson P, Potamianou A, Foucher A, Fife D, Mérit V. Vercammen E; PRIMS study group. Incidence of erythropoietin antibody-mediated pure red cell aplasia: the Prospective Immunogenicity Surveillance Registry (PRIMS). Nephrol Dial Transplant. 2015;30(3):451–60. https://doi.org/10.1093/ndt/gfu297.

Manivannan V, Decker WW, Stead LG, Li JT, Campbell RL. Visual representation of National Institute of Allergy and Infectious Disease and Food Allergy and Anaphylaxis Network criteria for anaphylaxis. Int J Emerg Med. 2009;2:3–5.

McKinnon RA, Cook M, Liauw W, Marabani M, Marschner IC, Packer NH, Prins JB. Biosimilarity and Interchangeability: principles and evidence: a systematic review. BioDrugs. 2018;32(1):27–52. https://doi.org/10.1007/s40259-017-0256-z.

Minghetti P, Rocco P, Schellekens H. The constrained prescription, interchangeability and substitution of biosimilars. Nat Biotechnol. 2015;33(7):688–9. https://doi.org/10.1038/nbt.3272.

Moots R, Azevedo V, Coindreau JL, Dörner T, Mahgoub E, Mysler E, Scheinberg M, Marshall L. Switching between reference biologics and biosimilars for the treatment of rheumatology, gastroenterology, and dermatology inflammatory conditions: considerations for the clinician. Curr Rheumatol Rep. 2017;19(6):37. https://doi.org/10.1007/s11926-017-0658-4.

Murdaca G, Spanò F, Contatore M, Guastalla A, Penza E, Magnani O, Puppo F. Immunogenicity of infliximab and adalimumab: what is its role in hypersensitivity and modulation of therapeutic efficacy and safety? Expert Opin Drug Saf. 2016;15(1):43–52. https://doi.org/10.1517/14740338.2016.1112375.

Park W, Yoo DH, Miranda P, Brzosko M, Wiland P, Gutierrez-Ureña S, et al. Efficacy and safety of switching from reference infliximab to CT-P13 compared with maintenance of CT-P13 in ankylosing spondylitis: 102-week data from the PLANETAS extension study. Ann Rheum Dis. 2017;76:346–54. https://doi.org/10.1136/annrheumdis-2015-208783.

Pfizer. Genotropin® prescribing information, December 2016, Pfizer. Sourced on 10 May 2018. 2016. http://labeling.pfizer.com/showlabeling.aspx?id=577.

Pisupati K, Benet A, Tian Y, Okbazghi S, Kang J, Ford M, Saveliev S, Sen KI, Carlson E, Tolbert TJ, Ruotolo BT, Schwendeman SP, Schwendeman A. Biosimilarity under stress: a forced degradation study of Remicade® and Remsima™. MAbs. 2017;9(7):1197–209. https://doi.org/10.1080/19420862.2017.1347741.

Pulsipher MA, et al. Lower risk for serious adverse events and no increased risk for cancer after PBSC vs BM donation. Blood. 2014;123(23):3655–63.

Radstake TR, Svenson M, Eijsbouts AM, van den Hoogen FH, Enevold C, van Riel PL, Bendtzen K. Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical responses in rheumatoid arthritis. Ann Rheum Dis. 2009;68(11):1739–45. https://doi.org/10.1136/ard.2008.092833.

Reinisch W, Jahnsen J, Schreiber S, Danese S, Panés J, Balsa A, Park W, Kim J, Lee JU, Yoo DH. Evaluation of the cross-reactivity of antidrug antibodies to CT-P13 and infliximab reference product (remicade): an analysis using immunoassays tagged with both agents. BioDrugs. 2017;31(3):223–37. https://doi.org/10.1007/s40259-017-0219-4.

Reinivuori T, Kurki P, Chamberlain P. Immunogenicity assessment of biosimilars. Pharm Med. 2018;32:103. https://doi.org/10.1007/s40290-018-0231-0.

Rombach-Riegraf V, Karle AC, Wolf B, Sordé L, Koepke S, Gottlieb S, Krieg J, Djidja MC, Baban A, Spindeldreher S, Koulov AV, Kiessling A. Aggregation of human recombinant monoclonal antibodies influences the capacity of dendritic cells to stimulate adaptive T-cell responses in vitro. PLoS One. 2014;9(1):e86322. https://doi.org/10.1371/journal.pone.0086322.

Rosenberg AS. Immunogenicity of biological therapeutics: a hierarchy of concerns. Dev Biol. 2003;112:15–21.

Rubic-Schneider T, Kuwana M, Christen B, Aßenmacher M, Hainzl O, Zimmermann F, Fischer R, Koppenburg V, Chibout SD, Wright TM, Seidl A, Kammüller M. T-cell assays confirm immunogenicity of tungsten-induced erythropoietin aggregates associated with pure red cell aplasia. Blood Adv. 2017;1(6):367–79. https://doi.org/10.1182/bloodadvances.2016001842.

Rup B, Pallardy M, Sikkema D, Albert T, Allez M, Broet P, Carini C, Creeke P, Davidson J, De Vries N, Finco D, Fogdell-Hahn A, Havrdova E, Hincelin-Mery A, Holland MC, Jensen PEH, Jury EC, Kirby H, Kramer D, Lacroix-Desmazes S, Legrand J, Maggi E, Maillère B, Mariette X, Mauri C, Mikol V, Mulleman D, Oldenburg J, Paintaud G, Pedersen CR, Ruperto N, Seitz R, Spindeldreher S, Deisenhammer F, ABIRISK Consortium. Standardizing terms, definitions and concepts for describing and interpreting unwanted immunogenicity of biopharmaceuticals: recommendations of the Innovative Medicines Initiative ABIRISK consortium. Clin Exp Immunol. 2015;181(3):385–400.

Ryding J, Stahl M, Ullmann M. Demonstrating biosimilar and originator antidrug antibody binding comparability in antidrug antibody assays: a practical approach. Bioanalysis. 2017;9(18):1395–406. https://doi.org/10.4155/bio-2017-0111.

Sandoz. Sandoz backgrounder for FDA Advisory Committee for GP2015 AC, 13 July 2016, Page 23. 2016. Available at https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/ucm510492.htm. Last Accessed 16 April 2018.

Sauerborn M, Brinks V, Jiskoot W, Schellekens H. Immunological mechanism underlying the immune response to recombinant human protein therapeutics. Trends Pharmacol Sci. 2010;31(2):53–9. https://doi.org/10.1016/j.tips.2009.11.001.

Scherlinger M, Schaeverbeke T, Truchetet ME. Serum sickness-like disease after switching to biosimilar infliximab. Rheumatology. 2017;56(11):2032–4. https://doi.org/10.1093/rheumatology/kex268.

Schiestl M, Stangler T, Torella C, Cepeljnik T, Toll H, Grau R. Acceptable changes in quality attributes of glycosylated biopharmaceuticals. Nat Biotechnol. 2011;29(4):310–2. https://doi.org/10.1038/nbt.1839.

Schmidt E, Hennig K, Mengede C, Zillikens D, Kromminga A. Immunogenicity of rituximab in patients with severe pemphigus. Clin Immunol. 2009;132(3):334–41.

Schultz HS, Reedtz-Runge SL, Bäckström BT, Lamberth K, Pedersen CR, Kvarnhammar AM, ABIRISK Consortium. Quantitative analysis of the CD4+ T cell response to therapeutic antibodies in healthy donors using a novel T cell:PBMC assay. PLoS One. 2017;12(5):e0178544. https://doi.org/10.1371/journal.pone.0178544.

Seidl A, Hainzl O, Richter M, Fischer R, Böhm S, Deutel B, Hartinger M, Windisch J, Casadevall N, London GM, Macdougall I. Tungsten-induced denaturation and aggregation of epoetin alfa during primary packaging as a cause of immunogenicity. Pharm Res. 2012;29(6):1454–67. https://doi.org/10.1007/s11095-011-0621-4.

Seo N, Polozova A, Zhang M, Yates Z, Cao S, Li H, Kuhns S, Maher G, McBride HJ, Liu J. Analytical and functional similarity of Amgen biosimilar ABP 215 to bevacizumab. MAbs. 2018;10(4):678–91. https://doi.org/10.1080/19420862.2018.1452580.

Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A, Quarmby V, Richards S, Schneider CK, Subramanyam M, Swanson S, Verthelyi D, Yim S, American Association of Pharmaceutical Scientists. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. AAPS J. 2014;16(4):658–73.

Sherman MR, Williams LD, Sobczyk MA, Michaels SJ, Saifer MG. Role of the methoxy group in immune responses to mPEG-protein conjugates. Bioconjug Chem. 2012;23(3):485–99. https://doi.org/10.1021/bc200551b.

Smith HW, Butterfield A, Sun D. Detection of antibodies against therapeutic proteins in the presence of residual therapeutic protein using a solid-phase extraction with acid dissociation (SPEAD) sample treatment prior to ELISA. Regul Toxicol Pharmacol. 2007;49(3):230–7.

Smolen JS, Choe J-Y, Prodanovic N, Niebrzydowski J, Staykov I, Dokoupilova E, Rho YH. Safety, immunogenicity and efficacy after switching from reference infliximab to biosimilar SB2 compared with continuing reference infliximab and SB2 in patients with rheumatoid arthritis: results of a randomised, double-blind, phase III transition study. Ann Rheum Dis. 2018;77(2):234–40. https://doi.org/10.1136/annrheumdis-2017-211741.

Somerville L, Bardelas J, Viegas A, D'Andrea P, Blogg M, Peachey G. Immunogenicity and safety of omalizumab in pre-filled syringes in patients with allergic (IgE-mediated) asthma. Curr Med Res Opin. 2014;30(1):59–66.

Stickler M, Valdes AM, Gebel W, Razo OJ, Faravashi N, Chin R, et al. The HLA-DR2 haplotype is associated with an increased proliferative response to the immunodominant CD4(+) T-cell epitope in human interferon-beta. Genes Immun. 2004;5(1):1–7.

Strik AS, van de Vrie W, Bloemsaat-Minekus JPJ, Nurmohamed M, Bossuyt PJJ, Bodelier A, Rispens T, van Megen YJB, D'Haens GR, SECURE Study Group. Serum concentrations after switching from originator infliximab to the biosimilar CT-P13 in patients with quiescent inflammatory bowel disease (SECURE): an open-label, multicentre, phase 4 non-inferiority trial. Lancet Gastroenterol Hepatol. 2018;3(6):404–12.

Subramanyam M. Case study: immunogenicity of natalizumab. In: Weert M, Møller EH, editors. Immunogenicity of biopharmaceuticals. Biotechnology: pharmaceutical aspects, vol. VIII. New York: Springer; 2008.

Talotta R, Berzi A, Doria A, Batticciotto A, Ditto MC, Atzeni F, Sarzi-Puttini P, Trabattoni D. The immunogenicity of branded and biosimilar infliximab in rheumatoid arthritis according to Th9-related responses. Int J Mol Sci. 2017;18(10):E2127. https://doi.org/10.3390/ijms18102127.

Tangri S, Mothé BR, Eisenbraun J, Sidney J, Southwood S, Briggs K, Zinckgraf J, Bilsel P, Newman M, Chesnut R, Licalsi C, Sette A. Rationally engineered therapeutic proteins with reduced immunogenicity. J Immunol. 2005;174(6):3187–96.

Tebbey PW, Varga A, Naill M, Clewell J, Venema J. Consistency of quality attributes for the glycosylated monoclonal antibody Humira® (adalimumab). MAbs. 2015;7(5):805–11. https://doi.org/10.1080/19420862.2015.1073429.

Torosantucci R, Schöneich C, Jiskoot W. Oxidation of therapeutic proteins and peptides: structural and biological consequences. Pharm Res. 2014;31(3):541–53.

Van Aerts LA, De Smet K, Reichmann G, van der Laan JW, Schneider CK. Biosimilars entering the clinic without animal studies. A paradigm shift in Europe. MAbs. 2014;6(5):1155–62.

van Schie KA, Wolbink GJ, Rispens T. Cross-reactive and pre-existing antibodies to therapeutic antibodies–Effects on treatment and immunogenicity. MAbs. 2015a;7(4):662–71. https://doi.org/10.1080/19420862.2015.1048411.

van Schie KA, Hart MH, de Groot ER, Kruithof S, Aarden LA, Wolbink GJ, Rispens T. The antibody response against human and chimeric anti-TNF therapeutic antibodies primarily targets the TNF binding region. Ann Rheum Dis. 2015b;74(1):311–4. https://doi.org/10.1136/annrheumdis-2014-206237.

van Schie KA, Kruithof S, van Schouwenburg PA, Vennegoor A, Killestein J, Wolbink G, Rispens T. Neutralizing capacity of monoclonal and polyclonal anti-natalizumab antibodies: the immune response to antibody therapeutics preferentially targets the antigen-binding site. J Allergy Clin Immunol. 2017;139(3):1035–1037.e6. https://doi.org/10.1016/j.jaci.2016.09.014.

van Schouwenburg PA, van de Stadt LA, de Jong RN, van Buren EE, Kruithof S, de Groot E, Hart M, van Ham SM, Rispens T, Aarden L, Wolbink GJ, Wouters D. Adalimumab elicits a restricted anti-idiotypic antibody response in autoimmune patients resulting in functional neutralisation. Ann Rheum Dis. 2013a;72(1):104–9. https://doi.org/10.1136/annrheumdis-2012-201445.

van Schouwenburg PA, Krieckaert CL, Rispens T, Aarden L, Wolbink GJ, Wouters D. Long-term measurement of anti-adalimumab using pH-shift-anti-idiotype antigen binding test shows predictive value and transient antibody formation. Ann Rheum Dis. 2013b;72(10):1680–6. https://doi.org/10.1136/annrheumdis-2012-202407.

van Schouwenburg PA, Kruithof S, Votsmeier C, van Schie K, Hart MH, de Jong RN, van Buren EE, van Ham M, Aarden L, Wolbink G, Wouters D, Rispens T. Functional analysis of the anti-adalimumab response using patient-derived monoclonal antibodies. J Biol Chem. 2014;289(50):34482–8. https://doi.org/10.1074/jbc.M114.615500.

van Schouwenburg PA, Kruithof S, Wolbink G, Wouters D, Rispens T. Using monoclonal antibodies as an international standard for the measurement of anti-adalimumab antibodies. J Pharm Biomed Anal. 2016;120:198–201. https://doi.org/10.1016/j.jpba.2015.12.040.

Verma A, Ngundi MM, Burns DL. Mechanistic analysis of the effect of deamidation on the immunogenicity of anthrax protective antigen. Clin Vaccine Immunol. 2016;23(5):396–402.

Verthelyi D, Wang V. Trace levels of innate immune response modulating impurities (IIRMIs) synergize to break tolerance to therapeutic proteins. PLoS One. 2010;5(12):e15252.

Vezér B, Buzás Z, Sebeszta M, Zrubka Z. Authorized manufacturing changes for therapeutic monoclonal antibodies (mAbs) in European Public Assessment Report (EPAR) documents. Curr Med Res Opin. 2016;32(5):829–34. https://doi.org/10.1185/03007995.2016.1145579.

Visser J, Feuerstein I, Stangler T, Schmiederer T, Fritsch C, Schiestl M. Physicochemical and functional comparability between the proposed biosimilar rituximab GP2013 and originator rituximab. BioDrugs. 2013;27(5):495–507. https://doi.org/10.1007/s40259-013-0036-3.

Waller C, Ranganna GM, Pennella E, Mattano LA, Loa CC, Donnelly C, Liu MS, Watson H, St-Jean J, Chatterjee S, Nayak V, Sengupta N, Kothekar MA, Barve A. Comparison of immunogenicity between the proposed pegfilgrastim biosimilar MYL-1401H and reference pegfilgrastim. Blood. 2017;130(Suppl 1):3568.

Wang YM, Fang L, Zhou L, Wang J, Ahn HY. A survey of applications of biological products for drug interference of immunogenicity assays. Pharm Res. 2012;29(12):3384–92.

Weber CA, Mehta PJ, Ardito M, Moise L, Martin B, De Groot AS. T cell epitope: friend or foe? Immunogenicity of biologics in context. Adv Drug Deliv Rev. 2009;61(11):965–76. https://doi.org/10.1016/j.addr.2009.07.001.

Weinbuch D, Zölls S, Wiggenhorn M, Friess W, Winter G, Jiskoot W, Hawe A. Micro-flow imaging and resonant mass measurement (Archimedes)–complementary methods to quantitatively differentiate protein particles and silicone oil droplets. J Pharm Sci. 2013;102(7):2152–65. https://doi.org/10.1002/jps.23552.

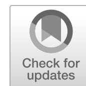
Weise M, Kurki P, Wolff-Holz E, Bielsky MC, Schneider CK. Biosimilars: the science of extrapolation. Blood. 2014;124(22):3191–6. https://doi.org/10.1182/blood-2014-06-583617.

Wolbink GJ, Voskuyl AE, Lems WF, de Groot E, Nurmohamed MT, Tak PP, et al. Relationship between serum trough infliximab levels, pretreatment C reactive protein levels, and clinical response to infliximab treatment in patients with rheumatoid arthritis. Ann Rheum Dis. 2005;64:704–7.

Wolbink GJ, Vis M, Lems W, de Groot E, Nurmohamed MT, Stapel S, et al. Development of antiinfliximab antibodies and relationship to clinical response in patients with rheumatoid arthritis. Arthritis Rheum. 2006;54:711–5.

Yoo DH, Prodanovic N, Jaworski J, Miranda P, Ramiterre E, Lanzon A, Park W. Efficacy and safety of CT-P13 (biosimilar infliximab) in patients with rheumatoid arthritis: comparison between switching from reference infliximab to CT-P13 and continuing CT-P13 in the PLANETRA extension study. Ann Rheum Dis. 2017;76(2):355–63. https://doi.org/10.1136/annrheumdis-2015-208786.

Chapter 20
Interchangeability Study Design and Analysis



Harry Yang, Dongyue Fu, and Lorin Roskos

Abstract The fact that many best-besting biologics are soon going off patient, coupled with the establishment of abbreviated regulatory pathways for the approval of biosimilar products, has fueled the development of biosimilars. However, in many countries, to realize the full benefits of a biosimilar, the product needs to attain the status of interchangeability, wherein the product is not only "highly similar" to the originator medicine but also is expected to produce the same clinical results for any given patient as the original product. The designation also requires demonstration that switching back and forth between the biosimilar and originator products would not increase the risk associated with the originator product when used alone. Regulatory approaches to interchangeability vary from region to region. While in some countries, an approved biosimilar is automatically granted to be interchangeable; FDA guidance requires data from a switching study or studies in one of more appropriate conditions of use. In this chapter, we discuss current regulatory thinking on interchangeability and focus primarily on considerations on interchangeability study design and analysis in fulfillment of the FDA requirements.

Keywords Biosimilars · Interchangeability · Pharmacokinetics · Substitution · Switching

H. Yang (✉) · D. Fu
Statistical Sciences, MedImmune, LLC, Gaithersburg, MD, USA
e-mail: yangh@medimmune.com; fud@medimmune.com

L. Roskos
Clinical Pharmacology & DMPK, MedImmune, LLC, Gaithersburg, MD, USA
e-mail: roskosl@medimmune.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_20

Introduction

The fact that many best-besting biologics are soon going off patient, coupled with the establishment of abbreviated regulatory pathways for the approval of biosimilar products, has fueled the development of biosimilars. However, in many countries, to realize the full benefits of a biosimilar, the product needs to attain the status of interchangeability, wherein the product is not only "highly similar" to the originator medicine but also is expected to produce the same clinical results for any given patient as the original product. The designation also requires demonstration that switching back and forth between the biosimilar and originator products would not increase the risk associated with the originator product when used alone. Regulatory approaches to interchangeability vary from region to region. While in some countries, an approved biosimilar is automatically granted to be interchangeable; FDA guidance requires data from a switching study or studies in one of more appropriate conditions of use. In this chapter, we discuss current regulatory thinking on interchangeability and focus primarily on considerations on interchangeability study design and analysis in fulfillment of the FDA requirements.

Background

Biosimilars are biopharmaceutical products that are shown to be of similar quality, efficacy and safety to the originator products. A biosimilar can be used as an alternative to the originator product after the patent of the latter product expires. The advances in analytical technologies and regulatory policies and regulations, which have created regulatory pathways for biosimilar across the world, have powered the biosimilar development, making it one of the fastest-growing sectors (EMA 2013, 2014; EMA and Europeam Commission 2017; FDA 2015a,b,c,d, 2017a,b). However, because of their large size, complex structure, and complicated manufacturing process, many scientific challenges remain for biosimilar development. Demonstration of biosimilarity requires data from comparative studies focused on both analytical characterization and clinical efficacy and safety. However, for a biosimilar product to be designated as interchangeable with the reference product, additional evidence needs to be generated. Although EMA was the first to create an abbreviated regulatory pathway for biosimilars (EMA 2013, 2014, EMA and European Commission 2017), it does not have a legal or regulatory framework for interchangeability; the designation resides in the health authorities of its Member States. By contrast, the US Biologics Pricing Competition and Innovation (BPCI) Act clearly lays down the conditions for interchangeability approval. Specifically, it states *interchangeable* or *interchangeability* means that the biological product "is biosimilar to the reference product" and "can be expected to produce the same clinical result as the reference product in any given patient" and that "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 alternate or switch" (FDA 2017a, b).

The newly issued FDA guidance on demonstration of interchangeability lays down the pathway for interchangeability approval. However, the lack of prior experience with interchangeability study design poses a host of challenges for the sponsors who seek the interchangeability designation for their biosimilar. So far, none of the approved biosimilars has been designated as interchangeable in the US.

Interchangeability, Switching and Substitution

Interchangeability, switching, and substitution are very different concepts which have caused much confusion among stakeholders. The formal definitions of these terms are given in the European Commission document (European Commission 2014) and listed in Table 20.1.

In general, interchangeability is a health or regulatory authority designation of a biosimilar product; whereas switching is a decision made by an attending physician and substitution is an action taken by a pharmacist. In the US, interchangeable means "the biological product may be substituted for the reference product without the intervention of the health care provider who prescribed the reference product." (FDA 2017a, b). Therefore, the US concept of interchangeability corresponds to the terminology of substitution in EU.

Interchangeability in the European Union

As discussed above, interchangeability is not a legal designation at the EU level. Furthermore, EMA does not take an official stance on interchangeability, nor does

Table 20.1 Definitions of interchangeability, switching and substitution

Terminology	Definition
Interchangeability	Medical practice of changing one medicine for another that is expected to achieve the same clinical effect in a given clinical setting and in any patient on the initiative, or with the agreement of the prescriber
Switching	Decision by the treating physician to exchange one medicine for another medicine with the same therapeutic intent in patients who are undergoing treatment
Substitution	Practice of dispensing one medicine instead of another equivalent and interchangeable medicine at the pharmacy level without consulting the prescriber

it have any plans to establish legal or regulatory requirements for interchangeability studies. (Kurki et al. 2017). Designation of interchangeability is a matter of the Member States in the European Union. As pointed out by Welch and Gutka in Chap. 2, there has been growing acceptance of a single switch between the reference and biosimilar products owing to publications of several switching studies, which revealed no significant safety concerns. Most recently several officials from the Finnish Medicines Agency, Netherlands' Medicines Evaluation Board, the Norwegian Medicines Agency, and Germany's Paul-Ehrlich-Institut argued that "a state-of-art demonstration of biosimilarity, together with intensified post-marketing surveillance, is a sufficient and realistic way of ensuring interchangeability of EU-approved biosimilars under supervision of the prescriber. In short, they concluded that "biosimilars licensed in the EU are interchangeable if the patient is clinically monitored" (Kurki et al. 2017).

Interchangeability in the United States

Contrasting to EMA's view, interchangeability in the US is a legal designation and the same as automatic substitution in EU. The prerequisite for interchangeability is that product is highly similar to the reference product. In addition, it can be expected to produce the same clinical result as the reference product in any given patient. The designation also requires a proof that a patient whose treatment is switched back and forth between the biosimilar and reference product would not have increased risk in safety and efficacy. In its draft guidance, the FDA recommends a switching study or studies, in which patients are switched between the biosimilar and reference product and monitored for any clinical meaningful difference, to support a demonstration of interchangeability. The guidance also recommends careful evaluation of the risks associated with differences in container closure system(s) and/or delivery device. Taken together, these additional data will provide evidence and information for interchangeability approval. However, an interchangeable biosimilar can only be substituted for the reference product at the pharmacy level if the state laws permit it.

Interchangeability Worldwide

Regulatory thinking and framework regarding interchangeability varies from region to region. A recent survey by Mendoza et al. (2015) indicates that in most countries, biosimilar substitution is currently not allowed or recommended. The decision of treating a patient with a reference biologic or biosimilar remains with the prescribing physician although automatic substitution occurs for economic reasons in some countries where no specific provisions or guidelines apply. It is foreseeable that changes regarding switching, interchangeability and substitution may be expected as these issues are under considerations by many regulatory bodies.

Switching Study

In this section, we are solely focused on discussions of switching study design and analysis per FDA requirements. A switching study is intended to evaluate the risk in terms of safety or diminished efficacy as a result of alternating or switching between the use of the biosimilar and reference products. To this end, it is necessary to use a comparative design for the study and most sensitive endpoints for detecting differences between the switching and non-switching arms. Since most biological products are of a long course, it is necessary to take into account dropouts in sizing the study so as to have sufficient power to demonstrate comparability between the two arms. Per the FDA guidance, if the interchangeability is demonstrated in an appropriate condition of use, the sponsor may seek licensure of the product as an interchangeable product for other indications for which the reference product is approved. Therefore, careful considerations should be given to the selection of indication for the switching study to maximize the sponsor's ability to extrapolate the product beyond the studied indication.

Study Endpoints

The FDA guidance (FDA 2017a, b) recommends the use of clinical pharmacokinetics and pharmacodynamics (if available) as the primary endpoints in a switching study. PK and PD measures are in general more sensitive to immunogenicity and exposure when compared to efficacy endpoints. Although efficacy endpoints may be supportive, the impact of immunogenicity and exposure on efficacy endpoints may not manifest in a study of limited duration and with a limited number of switches. Since measures of PK and PD parameters rely on accurate analytical test methods, it is also important to develop and validate these tests to ensure they are fit for their intended purpose. It is recommended that the switching study should include immunogenicity and safety assessments.

Study Design

Per the FDA guidance, a switching study should begin with a lead-in period of treatment with the reference product, followed by a randomization of patients into either one arm with alternating use of the biosimilar and reference product (switching arm) or the other arm in which patients continue receiving the reference product (non-switching arm). A diagram of a switching design is provided in Fig. 20.1.

In addition to the lead-in period, the study design includes a switching, a washout, and an intensive PK sampling period, which are discussed as follows:

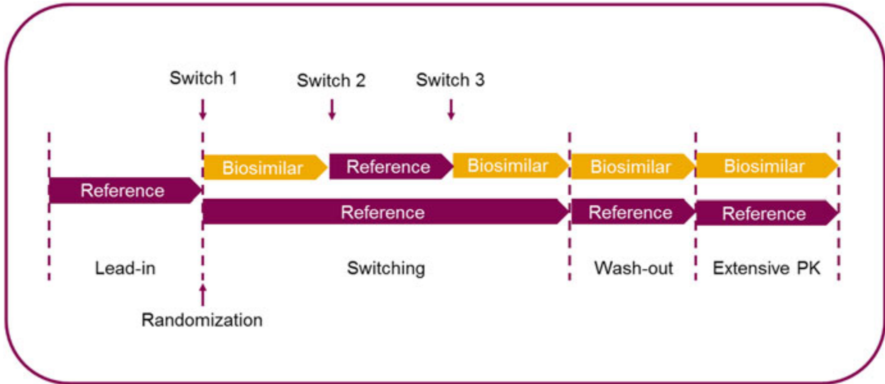


Fig. 20.1 Diagram of a switching study consisting of a lead-in period, switching, wash-out, and intensive PK

Lead-in Period

Per the FDA guidance on interchangeability (2017a, b), the lead-in period should be sufficiently long to ensure an adequate baseline with respect to the study (e.g., steady state of pharmacokinetics) before randomization to the switching period of the study. In the following, we demonstrate how to determine the duration of the lead-in period based on time needed to achieve steady state of pharmacokinetics. For the sake of simplicity, it is assumed that the half-life of the reference product is 3 weeks and the therapy consists of every 3-week dosing schedule. Therefore, the lead-in period is a multiple of 3 weeks. It is further assumed that the reference product is known to have linear PK, that is, the serum drug concentration C of the product follows the first order exponential model:

$$C = C_0 e^{-\lambda t} \tag{20.1}$$

where C_0 is the peak drug concentration and λ is the clearance rate.

From (20.1), it is evident that

$$C_0/2 = C_0 e^{-\lambda t_{1/2}}. \tag{20.2}$$

Solving (20.2) for λ, we obtain

$$\lambda = -\frac{\ln 2}{21}. \tag{20.3}$$

Since the dosing cycle is every 3 weeks, at the end of the first dose, the serum drug concentration is $0.5C_0$. The peak serum drug concentration is $C_0 + 0.5C_0$ after the second dose. Because of the property of linear PK of the reference product, the

serum drug concentration after the second dosing also follows first order exponential model below with the same clearance rate given in (20.3). That is,

$$C = (C_0 + 0.5C_0)\, e^{-(\ln 2/21)t}. \tag{20.4}$$

Following the same argument, it can be readily derived that the serum drug concentration at the end of the dose n is given by

$$C = \sum_{i=1}^{n} (0.5)^i C_0 = \left(1 - 0.5^n\right) C_0. \tag{20.5}$$

The serum drug concentration profile of the repeated dosing is shown in Fig. 20.2 along with the peak and trough values.

From (20.5), the trough serum concentrations are calculated and presented in Table 20.2.

As seen from the table, the steady state is achieved by week 18 where the overall intake of a drug is in equilibrium with its elimination evidenced by the trough level achieving its maximum. In addition, at the end of the week 12 or cycle 4, the trough serum concentration is a 94% of the steady state. Therefore, the steady state is approximately achieved. As a result, it is sufficient to use 0–12 weeks as the lead-in period.

Fig. 20.2 The time course of serum drug concentration after repeated dosing

Table 20.2 Trough serum drug concentration by dose cycle

Week	Dosing cycle	Trough
3	1	$0.5 \times C_0$
6	2	$0.75 \times C_0$
9	3	$0.88 \times C_0$
12	4	$0.94 \times C_0$
15	5	$0.97 \times C_0$
≥18	≥8	$1.00 \times C_0$

While the above example relies on the linear PK assumption, in practice, steady state is generally reached after 4 to 5 half-lives post the first dose (Dhillon and Gill 2017).

Number of Switching and Duration

It is recommended in the FDA guidance (FDA 2017a, b) that there should be at least three switches with each switch crossing over to the alternate product. In addition, the last switch should be from the reference product to the biosimilar. To mitigate the carryover effect of the reference product, it is suggested that the duration of the last switch should be at least three or more half-lives to allow for washout of the reference product before the intensive PK sampling period starts.

PK and Immunogenicity Sampling

Serum samples should be collected pre-dose for each infusion to establish the baseline PK measure and serum drug trough levels. Since the intensive PK sampling does not take place until the last cycle of treatment, care must be taken to ensure sampling uniformity. For example, a biological product administered through infusion is usually given over a shorter duration if the initial infusions were well tolerated. As pointed by Hussein et al. (1993), infusion duration may have a significant effect on some key PK parameters such as Cmax. Therefore it is important to specifically define the duration of infusion for the last intensive PK sampling period of the switching study. Of equal importance is to ensure samples are collected at adequate frequency to warrant accurate estimate of C_{max} and AUC.

Immunogenicity assessment is a key component of the switching study. It is intended to demonstrate that the alternating use of the biosimilar and reference product would not increase undesirable immune responses. Development of both neutralizing and non-neutralizing antibodies as a result of immunogenicity can cause loss of efficacy, resulting in diminished clinical benefits, disease progression, and even death. It may also alternate the pharmacokinetic properties of the treatment, thus prompting the need for dose modification. The impact of immunogenicity on safety is wide-ranged. They include potential allergic reactions, cytokine release syndrome, and cross-reactivity to endogenous proteins. However, clinical relevance of immunogenicity events observed from the switching study can only be elucidated through correlation with clinical responses. Therefore, it is crucial that samples for immunogenicity assessment are collected both at the baseline and the times when serum samples are collected for PK and PD analysis, and efficacy assessments. Such data would make it possible to evaluate correlation between immunogenicity occurrence and drug exposure, safety and efficacy outcomes, using the statistical techniques described by Yang et al. (2016).

Sample Size

Since the primary objective of the study is to establish PK equivalence, in terms of AUC_{tau} and C_{max}, between the switching and non-switching arms, the sample size should be based primarily on PK considerations. Per the FDA guidance, PK equivalence is established if the 90% confidence interval for the geometric mean ratio between the switching and non-switching arms is within 80–125%. This method is operational equivalent to two-one-sided t-tests (TOST) used to test the hypotheses:

$$H_0 : \mu_s/\mu_{ns} \leq 0.8 \text{ and } \mu_n/\mu_{ns} \geq 1.25 \quad \text{vs.} \quad H_a : 0.8 < \mu_s/\mu_{ns} < 1.25 \tag{20.6}$$

where μ_s and μ_{ns} are the mean PK parameters of the switching and non-switching arms, respectively.

Let $\mathbf{Y}_S = \left(Y_{S1}, \ldots, Y_{Sn_S}\right)$ and $\mathbf{Y}_{NS} = \left(Y_{NS1}, \ldots, Y_{NSn_{NS}}\right)$ denote two sets of values of a PK parameter determined from PK samples collected during the intensive PK sampling period from n subjects in the switching and non-switching arms, respectively. It is also assumed that $Y_{ij}(i = S \text{ or } NS, j = 1, \ldots, n)$ are distributed according to log-normal distributions such that

$$\ln\left(Y_{ij}\right) \sim N\left(\nu_i, \sigma_i^2\right). \tag{20.7}$$

To demonstrate how the sample size is calculated, we assume that $n_S = n_{NS} = n$ and $\sigma_S^2 = \sigma_{NS}^2$ for the sake of simplicity. Let CV denote the coefficient of variation of AUC_{tau}.

It can be derived that the mean and variance of Y_{ij} is given by

$$E\left[Y_{ij}\right] = \mu_i = e^{\nu_i + \sigma^2/2}$$

$$Var\left[Y_{ij}\right] = \left(e^{\sigma^2} - 1\right) e^{\nu_i + \sigma^2/2}. \tag{20.8}$$

From the first equation in (20.8), it becomes evident that

$$\mu_s/\mu_{ns} = e^{\nu_s - \nu_{ns}}. \tag{20.9}$$

Let $X_{S.l} = (\ln Y_{S1}, \ldots, \ln Y_{Sn})$ and $X_{NS.l} = (\ln Y_{NS1}, \ldots, \ln Y_{NSn})$. Note that the 90% confidence interval of $\nu_s - \nu_{ns}$ is given by

$$\overline{X}_{S.l} - \overline{X}_{NS.l} \pm t_{0.95}(2n - 2)\, s/\sqrt{2/n} \tag{20.10}$$

where $t_{0.95}(2n - 2)$ is the 95th percentile of the standard t-distribution with $2n - 2$ degrees of freedom, $\overline{X}_{S.l}$, $\overline{X}_{NS.l}$, and s are sample means of the switching and non-switching arms and pooled sample variance, respectively, given by

$$\overline{X}_{S.l} = \frac{\sum_{i=1}^{n} \ln\left[Y_{Si}\right]}{n}$$

$$\overline{X}_{NS.l} = \frac{\sum_{i=1}^{n} \ln\left[Y_{NSi}\right]}{n}$$

$$s = \frac{\sum_{i=1}^{n} \left\{\ln\left[Y_{Si}\right] - \overline{X}_{S.l}\right\}^2 / (n-1) + \sum_{i=1}^{n} \left\{\ln\left[Y_{NSi}\right] - \overline{X}_{NS.l}\right\}^2 / (n-1)}{2n-2}. \tag{20.11}$$

Combining (20.10) and (20.11), the 90% confidence interval of μ_s/μ_{ns} can be obtained as

$$e^{\overline{X}_{S.l} - \overline{X}_{NS.l} \pm t_{0.95}(2n-2)s/\sqrt{2/n}}. \tag{20.12}$$

It can be readily seen that

$$e^{\overline{X}_{S.l} - \overline{X}_{NS.l}} = \frac{\sqrt[n]{\prod_{i=1}^{n} Y_{Si}}}{\sqrt[n]{\prod_{i=1}^{n} Y_{NSi}}} \tag{20.13}$$

which is the ratio of the geometric means between the switching and non-switching arms.

From (20.8)

$$CV\left[Y_{ij}\right] = \sqrt{Var\left[Y_{ij}\right]}/E\left[Y_{ij}\right] = \left(e^{\sigma^2} - 1\right). \tag{20.14}$$

Solving the above equation for σ, we have

$$\sigma = \sqrt{\ln\left\{CV\left[Y_{ij}\right]^2 + 1\right\}}. \tag{20.15}$$

Assuming equal variability for the switching and non-switching arms, $CV[Y_{ij}] = CV$, Type I error of α, and power of $1 - \beta$, the sample size n can be determined by (Chow et al. 2008),

$$n = \frac{2\left(z_\alpha + z_{\beta/2}\right)^2 CV^2}{\delta^2}. \tag{20.16}$$

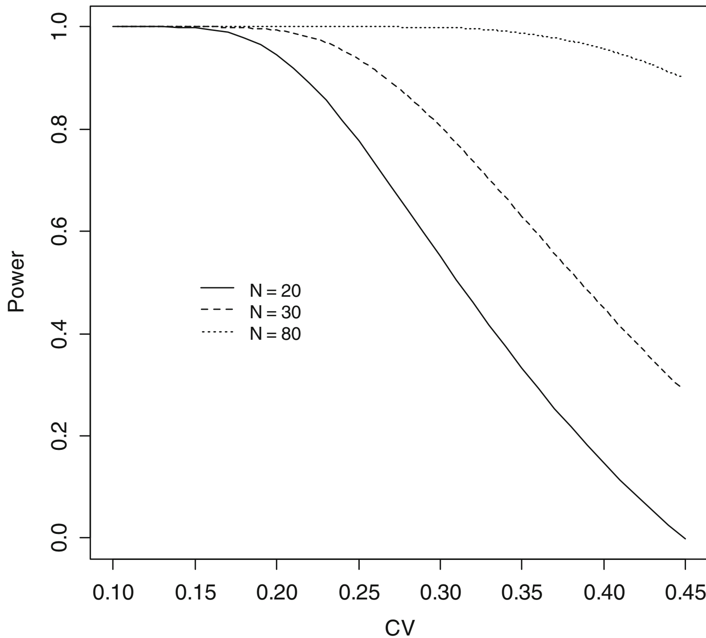


Fig. 20.3 Power curves of PK equivalence studies

where δ is the absolute value of the logarithmic of the equivalence limits 80–125%, which is equal to 0.223.

Therefore, the sample size of the switching study is strictly influenced by the variability of the PK parameters, characterized by the CVs. Figure 20.3 shows a plot of power curves for sample sizes n = 20, 30, and 80 versus various CV values.

From the plot, the higher the CV is, the lower the power. To achieve the same level of power, 90%, sample sizes of 20, 30, and 80 per arm are required for CVs of 24.4%, 30.1%, and 51.1%, respectively.

The sample size calculated using the formula in (20.16) provides power of $1 - \beta$ for demonstration of equivalence in one PK parameter. Since there are two primary PK endpoints AUC_{tau} and C_{max}, the sample size needs to be chosen to provide adequate overall study power. The overall power is the probability, *P*, for rejecting H_0^* below when H_a^* is true.

$$H_0^* : \mu_{s.i}/\mu_{ns.i} \le 0.8 \text{ and } \mu_{s.i}/\mu_{ns.i} \ge 1.25 \text{ vs. } H_a : 0.8 < \mu_{s.i}/\mu_{ns.i} < 1.25 \qquad (20.17)$$

where $i = AUC_{tau}$ and C_{max}.

Therefore, the power P can be expressed as:

$$P = P\left[\overline{Y}_{S.l.i} - \overline{Y}_{NS.l.i} \pm t_{0.95}\,(2n-2)\, s_i/\sqrt{2/n} \in (\ln 0.8,\ \ \ln 1.25)\right.$$
$$\left.\text{for}\ \ i = \text{AUC and } \text{C}_{\max}\right] \tag{20.18}$$

where $\overline{Y}_{S.l.i}$, $\overline{Y}_{NS.l.i}$, and s_i are sample means of the switching and non-switching arms and pooled sample variance, for $i = \text{AUC}_{\text{tau}}$ and $\text{C}_{\max}$, respectively.

When the variance-covariance matrix of $Y_{NS.\,AUC}$ and $Y_{NS.C_{\max}}$ is known and can be estimated from the historical data, the intervals $\overline{Y}_{S.l.i} - \overline{Y}_{NS.l.i} \pm t_{0.95}\,(2n-2)\, s_i/\sqrt{2/n}$ can be simultaneously simulated. From each iteration of simulation, the condition $\overline{Y}_{S.l.i} - \overline{Y}_{NS.l.i} \pm t_{0.95}\,(2n-2)\, s_i/\sqrt{2/n} \in (\ln 0.8,\ \ln 1.25)$ for $i = \text{AUC}_{\text{tau}}$ and $\text{C}_{\max}$ can be assessed. The power as a function of n can be estimated as the percent of times when the above condition is met. The sample size n is determined such that P is greater or equal to a pre-specified limit such as 80%. However, in practice, it is often assumed that P calculated of AUC_{tau} and $\text{C}_{\max}$, $Y_{NS.\,AUC}$ and $Y_{NS.C_{\max}}$ are independent. Under such an assumption,

$$\begin{aligned} P &= P\left[\overline{Y}_{S.l.AUC} - \overline{Y}_{NS.l.AUC} \pm t_{0.95}\,(2n-2)\, s_{AUC}/\sqrt{2/n} \in (\ln 0.8,\ \ln 1.25)\right] \\ &\quad \times P\left[\overline{Y}_{S.l.C_{\max}} - \overline{Y}_{NS.l.C_{\max}} \pm t_{0.95}\,(2n-2)\, s_{C_{\max}}/\sqrt{2/n} \in (\ln 0.8,\ \ln 1.25)\right] \\ &= P_{AUC} \times P_{C_{\max}} \end{aligned} \tag{20.19}$$

where P_{AUC} and $P_{C_{\max}}$ are power of demonstrating equivalence in AUC_{tau} and $\text{C}_{\max}$, respectively, for sample size n.

The sample size can be chosen such that the over power exceeds a pre-specified limit, say, 80%. To illustrate, consider $CV_{AUC} = 25\%$ and $CV_{C_{\max}} = 30\%$. It can be determined that with $n = 40$, $P_{AUC} \approx 96\%$ and $P_{C_{\max}} \approx 94\%$. Therefore, the overall power $P = 96\,\% \times 94\,\% \approx 80\,\%$.

Estimation of CV

As demonstrated above, the variability of PK parameter has a direct impact on the sample size and power of the switching study. A robust estimate of the CV lends the ability to sufficiently size the study for demonstration of interchangeability. In general, multiple PK studies were conducted for the marketing approval of the reference products. The results of these studies can be either extracted from published literature or regulatory documents which are in the public domain. Table 20.3 lists CV data from 6 PK studies of a marketed biological product.

Table 20.3 CV estimates of AUC from historical data

Study	Dose	N	CV
1	1	5	26%
	3	5	55%
	10	5	19%
2	3	5	15%
	3	3	19%
	3	4	27%
3	10	4	28%
4	3	15	33%
	10	17	31%
	20	41	24%
	7.5	16	37%
	15	30	31%
5	5	32	33%
	10	34	26%
6	15	28	25%

There are various ways to estimate the CV. The most common is the meta-analysis, which provides an overall estimate of CV based on results from multiple historical studies (Glass 1976; Stangl and Berry 2000; Whitehead 2002). The CVs from historical study can be viewed either as fixed effects or random effects. A fixed effects model assumes

$$CV_i = CV + \varepsilon_i \tag{20.20}$$

Where ε_i is an error term following a normal distribution $N(0, \tau^2)$.

A weighted regression analysis is carried out to provide a weighted estimate of the CV:

$$\overline{CV} = \frac{\sum_{i=1}^{I} w_i CV_i}{\sum_{i=1}^{I} w_i} \tag{20.21}$$

where CV_i is the CV_i from the ith study with sample size of n_i and w_i's are weights of the studies.

There are several choices of w_i. One way is to use the sample size of the study. In this case, the CV can be calculated as a weighted average of CVs from historical studies:

$$\overline{CV} = \frac{\sum_{i=1}^{I} n_i CV_i}{\sum_{i=1}^{I} n_i}. \tag{20.22}$$

As an example, the CV based on data in Table 20.3 is estimated to be $\overline{CV} = \frac{5\times 0.265\times 0.55+\cdots+28\times 0.16}{5+5+\cdots+28} \approx 29\%$.

Alternatively, one may estimate the overall CV through a random effect model, where CV_i's are viewed as observed random effects from a distribution (DerSimonian and Laird 1986; DerSimonian and Kacker 2007):

$$CV_i = CV + \mu_{CV_i} + \varepsilon_i \quad (20.23)$$

where CV is the true coefficient of variation, μ_{CV_i} are random effects to describe heterogenicity due to difference in study.

Sometimes, the data concerning the CVs of the reference product may be spare. For example, there might be only one published study that contains information of the CV in the form as shown in Table 20.4.

In such a circumstance, it is of interest to estimate the upper limit of the μ_{CV}. There are several methods suggested in the literature for constructing a confidence interval for CV. Lehman (1996) derived the sample distribution of CV for the purpose to calculate an exact confidence interval. Suppose X_i are measured PK parameters, say, AUC, from n ($i = 1, \ldots$, n) individuals. It is further assumed that X_i are identically and independently distributed according to a normal distribution with mean μ and variance σ^2. Hence the statistic $\overline{X}/s\sqrt{n}$ follows a non-central t-distribution with n − 1 degrees of freedom and non-centrality parameter of

$$\tau = \mu\sqrt{n}/\sigma. \quad (20.24)$$

That is,

$$\overline{X}/s\sqrt{n} \sim t_{n-1}(\tau). \quad (20.25)$$

Based on the above distribution, an exact 95% lower confidence limit of τ is given by τ_L such that

Table 20.4 AUC estimates of 44 patients from a single historical study

Patient	AUC	Patient	AUC	Patient	AUC	Patient	AUC
1	2150	12	2230	23	2280	34	2340
2	1880	13	1920	24	1970	35	2220
3	2540	14	1760	25	2170	36	1870
4	1570	15	2810	26	3010	37	2260
5	2820	16	2750	27	2360	38	2330
6	1990	17	1420	28	2380	39	1530
7	1990	18	2130	29	2100	40	2880
8	2360	19	2140	30	2630	41	2600
9	2890	20	1720	31	3100	42	2720
10	2080	21	2320	32	2200	43	2240
11	1950	22	2720	33	1980	44	2710

$$P\left[t_{n-1}\left(\tau_L\right) < \overline{X}/s\sqrt{n}\right] = 0.95. \tag{20.26}$$

Form (20.24), the CV, σ/μ, can be expressed as

$$\sigma/\mu = \sqrt{n}/\tau. \tag{20.27}$$

Consequently, the 95% upper confidence limit for σ/μ is given by

$$CV_U = \sqrt{n}/\tau_L. \tag{20.28}$$

Both statistical software packages SAS and R provide subroutines and functions for calculating τ_L. Based on the data in Table 20.4, the CV of AUC can be calculated as 17.9%. Applying the above method, using function *ci.cv* in R package MBESS, the upper 95% confidence limit of CV is estimated to be 21.9%. Erring to the safe side, it is advisable to use the upper limit of CV to size the switching study.

An alternate method based on generalized pivotal quantity analysis can also be used to obtain an interval estimate of the CV (Tsui and Weerahandi 1989; Weerahandi 1993).

Dropout

One of the factors that impact the size of the switching study is the dropout rate, which is defined as percent of patients who do not complete the intensive PK sampling. Because the long course of therapy of the switching study, it is conceivable that dropout rate would be high. This many diminish the probability for demonstration of PK equivalence. To address this issue, it is important to estimate the dropout rate and factor it into the sample size calculation. Let r denote the dropout rate. The total number patients that need to enter into the study is given by

$$N = 2n/\left(1 - r\right). \tag{20.29}$$

Patients may drop out of a study due to variety of reasons such as disease progression, death, and intolerance to toxicities caused by the treatment or others. Data from historical studies of the reference product in indication used for the switching study may be used to estimate the dropout rate. Information related to median time of on treatment and percent of patients who stayed on treatment for other fixed amounts of time is most useful. These data can be used to predict the dropout rate for the switching study based on statistical modeling. Among potential models that can be used for this purposes are exponential and Weibull models, which are often utilized for describing time to event data. Let X denote the duration from randomization to the time when the patient's treatment is terminated. When X follows an exponential distribution, it implies

$$Y(t) = P\left[X > t\right] = e^{-\theta t}. \tag{20.30}$$

Table 20.5 Treatment termination data

Week	Cycle	Y(t)
12	4	0.75
21	7	0.5
48	16	0.2

Oftentimes, the model parameter θ is unknown and needs to be estimated. Assuming that from the historical data, we obtain observations $Y(t_i)$ at times t_i, $i = 1, \ldots, k$. The least square estimate of θ is given by

$$\widehat{\theta} = -\frac{\sum_{i=1}^{k} \ln\left[Y\left(t_i\right)\right] t_i}{\sum_{i=1}^{k} t_i^2}. \tag{20.31}$$

To demonstrate, we assume that the following data are available for an oncology product that is administered every 3 weeks (Table 20.5).

From the table,

$$\widehat{\theta} = -\frac{(\ln 0.75) \times 12 + \ln 0.5 \times 21 + \ln 0.2 \times 48}{12^2 + 21^2 + 48^2} \approx 0.033. \tag{20.32}$$

Suppose that a switching study is designed with a 12-week lead-in period, 3 switches, the first and second of which last for 3 weeks each, and the last for 9 weeks to warrant sufficient washout of the reference product, and a 3-week intensive PK sampling period. The total study duration is 30 weeks. From (20.12) and (20.14), the dropout rate by the end of week 30 is given by,

$$r = 1 - Y(30) = 1 - e^{-0.033 \times 30} \approx 62.8\%. \tag{20.33}$$

It is very common that only median duration of treatment, t_M, is reported. In such case, the model parameter θ may be calculated as a solution to the following equation:

$$0.5 = P\left[X > t_M\right] = e^{-\theta t_M}. \tag{20.34}$$

Solving the above equation results in

$$\widehat{\theta} = -\frac{\ln 2}{t_M}. \tag{20.35}$$

From Table 20.3, $t_M = 21$. Substituting this value for t_M in (20.17), we obtain

$$\widehat{\theta} = -\frac{\ln 2}{21} \approx 0.033. \tag{20.36}$$

As a result, the dropout rate at the end of week 30 for the above example is given by

$$r = 1 - e^{-0.033 \times 30} \approx 62.9\%. \tag{20.37}$$

Alternatively, one may use a Weibull model to describe the time to treatment termination variable X. That is,

$$Y(t) = e^{-(\theta t)^p} \tag{20.38}$$

where the parameters $\theta > 0$ and $p > 0$. p is called the shape parameter. When $p = 1$, the model is reduced to the exponential model in (20.34). $p < 1$, $p = 1$, and $p > 1$ correspond to the situations that the treatment termination rate decreases, remains constant, and increases, respectively.

Note that

$$\ln\left[-\ln Y(t)\right] = p_0 + p \ln t \tag{20.39}$$

where

$$p_0 = p \ln \theta. \tag{20.40}$$

Because of the linear relationship in (20.19), the least squared estimates of parameters p_0 and p can be calculated, using the following formula,

$$\widehat{p} = -\frac{\sum_{i=1}^{k} \left\{\ln\left[-\ln Y(t_i)\right] - \overline{Y}_{ll}\right\} \times \left(\ln t_i - \bar{t}_{ll}\right)}{\sum_{i=1}^{k} \left(\ln t_i - \bar{t}_{ll}\right)^2}$$

$$\widehat{p}_0 = \overline{Y}_{ll} - \widehat{p}\bar{t}_{ll} \tag{20.41}$$

where

$$\overline{Y}_{ll} = \frac{\sum_{i=1}^{k} \ln\left[-\ln Y(t_i)\right]}{k}$$

$$\bar{t}_{ll} = \frac{\sum_{i=1}^{k} \ln t_i}{k}. \tag{20.42}$$

Combining (20.40) and (20.41), θ can be estimated by

$$\widehat{\theta} = e^{(\widehat{p}_0/\widehat{p})}. \tag{20.43}$$

For the example discussed above, now we assume that the time to treatment termination variable X has a Weibull distribution. From (20.42),

$$\overline{Y}_{ll} = \frac{\ln[-\ln 0.75] + \ln[-\ln 0.5] + \ln[-\ln 0.25]}{3} \approx -0.379$$

$$\bar{t}_{ll} = \frac{\ln 7 + \ln 21 + \ln 48}{3} = 3.134. \tag{20.44}$$

From (20.40) and (20.42),

$$\widehat{p} = -\frac{\{\ln[-\ln 0.75] - (-0.397)\} \times (\ln 7 - 3.134)}{(\ln 7 - 3.134)^2 + (\ln 21 - 3.134)^2 + (\ln 48 - 3.134)^2}$$

$$-\frac{\{\ln[-\ln 0.5] - (-0.397)\} \times (\ln 21 - 3.134)}{(\ln 7 - 3.134)^2 + (\ln 21 - 3.134)^2 + (\ln 48 - 3.134)^2}$$

$$-\frac{\{\ln[-\ln 0.25] - (-0.397)\} \times (\ln 48 - 3.134)}{(\ln 7 - 3.134)^2 + (\ln 21 - 3.134)^2 + (\ln 48 - 3.134)^2}$$

$$\approx 1.225$$

$$\widehat{p}_0 = -0.379 - 1.225 \times 3.134 = -4.218$$

$$\widehat{\theta} = e^{(-4.218/1.225)} \approx 0.387. \tag{20.45}$$

Hence, the dropout rate at the end of week 30 is obtained as

$$r = 1 - Y(30) = e^{-(0.387 \times 30)^{-4.218}} \approx 61.3\%. \tag{20.46}$$

Comparing (20.33), (20.37), and (20.46), the three methods provide similar estimates for the dropout rate for the example provided in Table 20.3.

Patient Population and Condition of Use

For the switching study, the FDA guidance suggests that the patient population should be adequately sensitive to allow for detection of differences in pharmacokinetics and/or pharmacodynamics, common adverse events, and immunogenicity between the switching and non-switching arms. Although the FDA is open for conducting the switching study in a patient population which is different from those used for the licensure of the reference product, it expects the sponsor to provide adequate justification. Similarly, the FDA guidance encourages selection of a condition of use for which the reference product is licensed. There are several advantages of following these recommendations for the switching study: (1) The switching study would mimic how the proposed interchangeable product will be used in clinical practice; (2) Although the switching study is adequately sized and powered to address the primary objective of PK comparison between the switching and non-switching arms, comparison of safety and immunogenicity profiles between the switching and non-switching arms is of great importance. Historical clinical study data may provide a context for assessing clinical and statistical significance if any residual uncertainty arises due to a difference in safety and immunogenicity observed from the switching study; (3) Use of the same patient population and indication for the switching study would support extrapolation of data to other conditions of use.

Stratification

There are patient baseline factors that impact on PK parameters. Of particular note is gender. If the products are intended to be used in combination of other standard care such as chemotherapies, to encourage recruitment, the sponsor may render the sites the flexibility of using the standard care in accordance to local policies and principal investigator's preference. Under such circumstance, it is necessary to use stratified randomization to ensure balance of patients between the two arms.

Data Analysis

Primary Analysis

As previously discussed, the primary objective of the study is to establish PK equivalence between the switching and non-switching arms. For this purpose, PK data from the intensive PK sampling period are used to construct the 90% confidence intervals (CIs) of AUC_{tau} and C_{max} for the ratio of geometric means. This can be

accomplished through performance a one-way analysis of the variance (ANOVA) on log-transformed data, using the following model:

$$\ln\left(Y_{ij}\right) = v_i + \varepsilon_{ij} \tag{20.47}$$

where Y_{ij} is observed value of the jth patient in the ith treatment arm $i = S, NS$; $j = 1, \ldots, n_i$; v_i is mean of treatment arm i and ε_{ij} is residual error, following a normal distribution $N\left(0, \sigma_i^2\right)$.

There are two ways to construct the 90% C is depending on if the variance from the two arms are equal or not (Burdick et al. 2017). The assessment of equal variance can be assessed either through graphical display of the data by arm or a statistical test. However, the probability of rejecting the equal variance claim depends heavily on the sample size, $n_1 + n_2$. Therefore, graphical examination is a preferable method for this evaluation. When the variances are equal, the 90% CI for the ratio of geometric means is given by:

$$e^{\overline{X}_S - \overline{X}_{NS} \pm t_{0.95}(v)\sqrt{s_p^2(1/n_S + 1/n_{NS})}}, \tag{20.48}$$

where $\overline{X}_S$ and $\overline{X}_{NS}$ are sample means of $X_{ij} = \ln\left(Y_{ij}\right)$, $i = S, NS$, and s_p^2 is the pooled sample variance given by

$$s_p^2 = \frac{(n_s - 1)\, s_S^2 + (n_{NS} - 1)\, s_{NS}^2}{n_s + n_{Ns} - 2} \tag{20.49}$$

with s_S^2 and s_{NS}^2 being the sample variance of the switching and non-switching arms, respectively, and $t_{0.95}(v)$ is the 95 percentile of the standard t-distribution with $v = n_S + n_{NS} - 2$ degrees of freedom.

When the variances of the two arms are not equal, the 90% CI for the ratio of geometric means is obtained by

$$e^{\overline{X}_S - \overline{X}_{NS} \pm t_{0.95}(v)\sqrt{s_N^2/n_S + s_{NS}^2/n_{NS}}}, \tag{20.50}$$

where the degrees of freedom ν estimated from the Satterthwaite's approximation (Satterthwaite 1946):

$$\nu = \frac{\left(s_T^2/N_T + s_R^2/N_R\right)^2}{\frac{s_R^4}{(N_R - 1)N_T^2} + \frac{s_T^4}{(N_T - 1)N_R^2}}. \tag{20.51}$$

The PK equivalence is established if the 90% CIs for both AUC_{tau} and C_{max} are contained within (0.80, 1.25). To illustrate, we simulated a set of AUC_{tau}

values based on X_{ij} $N(v_i, \sigma^2)$, $i = 1$ or 2, $j = 1 \ldots 80$, and the assumptions: (1) $n_S = n_{NS} = 40$; (2) The geometric means are 51,200 h μg/mL and 49,500 h μg/mL for the for switching non-switching arms, respectively; (3) The CV is 0.3 and the same for the 2 arms. Thus $\sigma = \sqrt{\ln\left(CV^2 + 1\right)} = 0.29$. The data are presented in Table 20.6.

Based on the simulated data set, summary for AUC_{tau} by treatment arm including 95% CI for the geometric mean ratio is provide in Table 20.7 below.

The 90% CI for the geometric mean ratio of AUC_{tau} is (1.008, 1.231). Therefore, it can be concluded that equivalence in AUC_{tau} is demonstrated.

It is known that some patient characteristics such as weight and gender may have an impact on PK parameters. Therefore, imbalance in these parameters may negatively affect the demonstration of PK equivalence. To mitigate such risk, these baseline measurements can be used either as stratification factors and/or covariate in the PK comparison. The 90% confidence intervals discussed above can be constructed based on multi-way ANOVA or analysis of covariance (ANCOVA). As an example, we simulated random sample of baseline weight using a normal distribution $N(75, 15)$ to augment the data in Table 20.6 and carried out the ANCOVA analysis using the data. The 90% CI for the geometric mean ratio of AUCtau, adjusted for baseline weight, is (1.009, 1.229). The interval is slightly narrower than that previously determined.

Secondary Analysis

Per FDA guidance, other PK data (C_{trough} and T_{max}) should be obtained and analyzed as secondary endpoints. Safety (adverse events, clinical laboratory data, etc.), immunogenicity, and efficacy should be also analyzed as secondary endpoints, using descriptive summaries. Since no formal hypotheses are tested concerning these secondary endpoints, visual display of these data often provide sufficient information for assessing comparability of the data between the switching and non-switching arms. For example, occurrence of adverse events (AEs) can be summarized and visually displayed by treatment group, in terms of percent of patients experiencing AEs, having grade 3 or higher AEs, serious AEs (SAEs), AEs leading to discontinuation of treatment, or AEs resulting in death. Figure 20.4 shows one example of this graphical comparison. It is evident that the two arms are of a similar safety profile. The same visual examination can be carried out for AEs common to the reference product and AEs of special interest, etc.

Table 20.6 Simulated AUC_{tau} dataset

Treatment arm	AUC_{tau}	$Ln(AUC_{tau})$	Treatment arm	AUC_{tau}	$Ln(AUC_{tau})$
Switching	55,330	10.92	Non-switching	68,019	11.13
Switching	70,192	11.16	Non-switching	42,789	10.66
Switching	65,095	11.08	Non-switching	35,203	10.47
Switching	43,531	10.68	Non-switching	43,213	10.67
Switching	80,468	11.30	Non-switching	65,322	11.09
Switching	35,642	10.48	Non-switching	27,733	10.23
Switching	49,116	10.80	Non-switching	43,612	10.68
Switching	69,521	11.15	Non-switching	50,184	10.82
Switching	52,198	10.86	Non-switching	36,477	10.50
Switching	73,364	11.20	Non-switching	62,457	11.04
Switching	49,019	10.80	Non-switching	55,506	10.92
Switching	55,752	10.93	Non-switching	39,361	10.58
Switching	36,537	10.51	Non-switching	48,445	10.79
Switching	44,437	10.70	Non-switching	74,720	11.22
Switching	44,825	10.71	Non-switching	62,153	11.04
Switching	55,486	10.92	Non-switching	73,095	11.20
Switching	47,940	10.78	Non-switching	54,613	10.91
Switching	54,593	10.91	Non-switching	48,096	10.78
Switching	45,261	10.72	Non-switching	69,191	11.14
Switching	55,040	10.92	Non-switching	35,019	10.46
Switching	66,838	11.11	Non-switching	54,782	10.91
Switching	79,176	11.28	Non-switching	36,799	10.51
Switching	83,429	11.33	Non-switching	51,731	10.85
Switching	31,247	10.35	Non-switching	60,601	11.01
Switching	40,948	10.62	Non-switching	41,573	10.64
Switching	61,586	11.03	Non-switching	36,523	10.51
Switching	77,962	11.26	Non-switching	49,083	10.80
Switching	56,915	10.95	Non-switching	53,202	10.88
Switching	50,862	10.84	Non-switching	41,455	10.63
Switching	58,729	10.98	Non-switching	28,230	10.25
Switching	63,715	11.06	Non-switching	50,323	10.83
Switching	43,476	10.68	Non-switching	33,902	10.43
Switching	76,442	11.24	Non-switching	62,015	11.04
Switching	52,482	10.87	Non-switching	73,289	11.20
Switching	34,818	10.46	Non-switching	38,295	10.55
Switching	29,041	10.28	Non-switching	48,328	10.79
Switching	61,072	11.02	Non-switching	50,300	10.83
Switching	65,081	11.08	Non-switching	32,908	10.40
Switching	69,583	11.15	Non-switching	61,137	11.02
Switching	38,200	10.55	Non-switching	56,395	10.94

Table 20.7 Summary for AUC_{tau} by treatment arm

Statistics	Switching arm ($n_S = 40$)	Non-switching arm ($n_{NS} = 40$)
Geometric mean	53,739	48,257
$\overline{x}_i$	10.89	10.78
S_i	0.271	0.266
Difference		
Geometric mean ratio	1.11	
$\overline{x}_S - \overline{x}_{NS}$	0.11	
Standard error $SE_{\overline{x}_S - \overline{x}_{NS}} = s_p\sqrt{2/n}$	0.06	
95% CI		
Log scale	(0.008, 0.207)	
Original scale (i.e. CI for the geometric mean ratio)	(1.008, 1.231)	

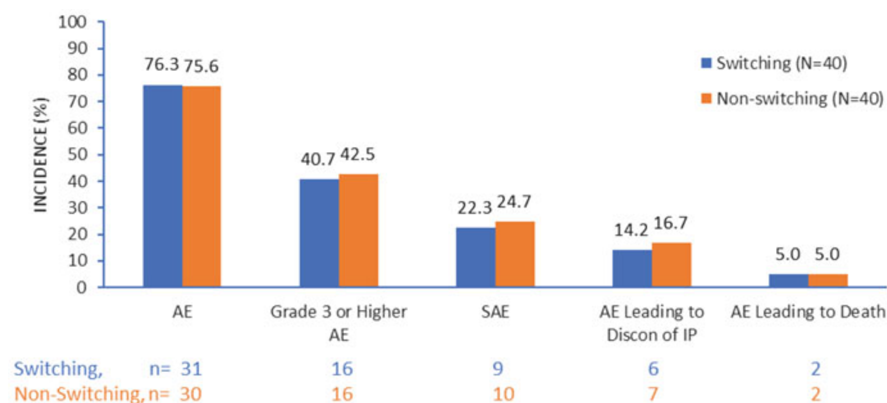


Fig. 20.4 Comparison of AE profiles between switching and non-switching arms

Other Considerations

Drug Supply

Continued and adequate drug supply is a key component of conducting a successful clinical trial. For biosimilar studies, drug supply presents a host of unique challenges. First of all, for each reference product, there are usually several biosimilars in clinical development. Therefore, sourcing the reference product can be very competitive. Secondly, since biological products are often intended for repeated or long-term use, large quantities of the reference drug are needed over an extended period. Thirdly, regulatory authorities for certain region require use of the reference product in the clinical studies which is specifically produced for that region. This limits the number of procurement sources. Advanced planning and strategic thinking

are key to making the interchangeability study less costly and complete accordingly to the plan. It is also advantageous to break down drug supply by calendar time so that the procurement of the drug can be planned accordingly.

Modeling Approach

In an oncology study, drug supply can be estimated in terms of expected total number of cycles of treatment needed for each month for the entire duration of the study. When this estimate is available, the amount of drug can be calculated with the input of other parameters such as dose levels, e.g., 15 mg/kg, and average weight of the patients in the study. A key factor in the estimation of number of cycles of treatment for a given month is the conditional probability for a patient to stay on the treatment for the month. This can be accomplished through modeling the time to treatment discontinuation, T. Assume T follows an exponential distribution $P[T > t] = e^{-\lambda t}$. Let t_0 be the time at the beginning of the month, and t ($>t_0$) a future time and $d = t - t_0$. It can be derived that the conditional probability for a patient to receive treatment at times from t_0 to t can be derived

$$Y(t) = P(T > t|T > t_0) = \frac{Prob\,(T > t\ and\ T > t_0)}{Prob\ (T > t_0)} = \frac{Prob\,(T > t)}{Prob\,(T > t_0)} = \frac{e^{-\lambda t}}{e^{\lambda t_0}} = e^{-\lambda d} \tag{20.52}$$

From (20.52), the total number of cycles needed from $t_0$to t is calculated as

$$\frac{d}{cycle\ duration} \times (N_a + N_{dr}/2)\,, \tag{20.53}$$

where N_{dr} is the number of dropouts from t_0 to t and N_a is the number of active subjects at time t:

$$N_{dr} = \text{No. of active subjects at } t_0 \times [1 - Y(t)]$$

$$N_a = \text{No. of randomized subjects at t} - \text{No. of total dropouts at t.}$$

Note that we assume that the distribution of dropout is uniform in the time interval $[t_0, t]$. Therefore, on the average, these dropout patients only have half of the planned cycles of treatment. It is also important to point out that in order to estimate the number of active subjects at the beginning of a time interval, the recruitment rate needs to be specified.

Example

In the following, we illustrate the use of the above method with an example. It is assumed that 80 subjects are planned to be randomized to the study within 6 months with the first subject being randomized in December, 2017. It is further assumed that the rate of recruitment is constant and drugs are administered in each 14-day cycle. Drug supplies are planned to be provided until January, 2020 when the study is complete. The median duration of treatment time is 190 days (about 6.24 months) for the reference product. Based on above-mentioned the exponential model, the model parameter is estimated by

$$\lambda = -\frac{\ln(0.5)}{6.24} = 0.111$$

From (20.52)

$$Y(t) = e^{-0.111d} \tag{20.54}$$

Based on (20.53), the number of treatment dropouts and cycles by calendar time can be estimated iteratively. The results are presented in Table 20.8.

From the table, there are 62 active subjects by 30-Jun-18. The probability of subjects who are still active by 31-Jul-18 is 0.893. Therefore, the number of dropouts from 30-Jun-18 to 31-Jul-18 is calculated as $62 \times (1 - 0.893) = 7$. Since by 31-Jul-18, 80 subjects are randomized and the total number of dropouts is 24. Therefore, the number of active subjects by 31-Jul-18 is 80 – 24 = 56. From (20.53), the total number of cycles until 31-Jan-20 is 1591 in this case. We could calculate the number of kits/vials needed at certain time or total based on this estimate and further calculate the corresponding cost at study plan stage. In practice, certain percent of overage, for example, 15–20%, is usually included to ensure the enough drug supplies. If needed, the post-study drug supplies can be also estimated.

Concluding Remarks

In recent years, thanks to the advances in regulations, several abbreviated licensure pathways have been created for biosimilars. As a result, various biosimilar products have been approved for marketing in EU, US, and other regions. However, to realize the full benefit of biosimilars, it is important to gain the designation of interchangeability, which allows the biosimilars to be substituted for the reference products at the pharmacy level. In the US, the first approved interchangeable product also enjoys one year exclusivity. However, regulations that govern such interchangeability status are sparse and vary from region to region. In 2017, the US FDA issued guidance on demonstration of interchangeability. Chief among

Table 20.8 Estimate of number of cycles of treatment needed by calendar time

Calendar time	d (mons)	$Y(t)$	No. of randomized subjects	No. of drop-out	Active	Total dropouts	No. of cycles
31-Dec-17			1	0	1	0	
31-Jan-18	1	0.893	13	0	13	0	29
28-Feb-18	1	0.903	26	1	25	1	51
31-Mar-18	1	0.893	39	3	35	4	80
30-Apr-18	1	0.896	52	4	44	8	99
31-May-18	1	0.893	65	5	53	12	122
30-Jun-18	1	0.896	80	5	62	18	139
31-Jul-18	1	0.893	80	7	56	24	130
31-Aug-18	1	0.893	80	6	50	30	116
30-Sep-18	1	0.896	80	5	44	36	101
31-Oct-18	1	0.893	80	5	40	40	93
30-Nov-18	1	0.896	80	4	36	44	81
31-Dec-18	1	0.893	80	4	32	48	75
31-Jan-19	1	0.893	80	3	28	52	67
28-Feb-19	1	0.903	80	3	26	54	54
31-Mar-19	1	0.893	80	3	23	57	54
30-Apr-19	1	0.896	80	2	21	59	46
31-May-19	1	0.893	80	2	18	62	43
30-Jun-19	1	0.896	80	2	16	64	37
31-Jul-19	1	0.893	80	2	15	65	34
31-Aug-19	1	0.893	80	2	13	67	31
30-Sep-19	1	0.896	80	1	12	68	27
31-Oct-19	1	0.893	80	1	10	70	25
30-Nov-19	1	0.896	80	1	9	71	21
31-Dec-19	1	0.893	80	1	8	72	20
31-Jan-20	1	0.893	80	1	7	73	18
Total							1591

the requirements is a switching study or studies designed to show that alternating use of a proposed biosimilar and the reference product would not incur more risk than the use of the reference product alone. This chapter is primarily focused on considerations for the design and analysis of the switching study. Regardless of patient population and selection of condition of use, the statistical treatments discussed are in general directly applicable for the switching study design.

Acknowledgement The authors would like to thank Drs. Balaji Agoram and Mei Tang for helpful discussions on the interchangeability study design.

References

Burdick RK, et al. Statistical applications for chemistry, manufacturing and controls (CMC) in the pharmaceutical industry. Cham: Springer; 2017.
Chow SC, Shao J, and Wang H. Sample size calculations in clinical research. 2nd ed. New York: Chapman & Hall/CRC; 2008.
DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. Cotemp Clin Trials. 2007;28(2):105–14.
DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.
Dhillon S, Gill K. Basic pharmacokinetics. 2017. http://www.dandybooksellers.com/acatalog/9780853695714.pdf. Assessed 11 Nov 2017.
EMA. Biosimilar medicinal products, brochure. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2011/03/WC500104228.pdf. Assessed 11 Nov 2017.
EMA. Guidelines on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/01/WC500180219.pdf. Assessed 11 Nov 2017.
EMA and European Commission. Biosimilars in the EU: information guide for healthcare professionals. 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2017/05/WC500226648.pdf. Assessed 11 Nov 2017.
European Commission. What you need to know about biosimilar medicinal products: a consensus information document. 2014. http://ec.europa.eu/DocsRoom/documents/8242. Assessed 11 Nov 2017.
European Medicines Agency [EMA] and U.S. Food and Drug Association [FDA]. General principles EMA-FDA parallel scientific advice (human medicinal products). 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2009/11/WC500014868.pdf. Assessed 11 Nov 2017.
FDA. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. Silver Spring: The United States Food and Drug Administration; 2015a.
FDA. Guidance for industry: quality considerations in demonstrating biosimilarity to a reference product. Silver Spring: The United States Food and Drug Administration; 2015b.
FDA. Guidance for industry: biosimilars: questions and answers regarding implementation of the biologics price competition and innovation act of 2009. Silver Spring: The United States Food and Drug Administration; 2015c.
FDA. FDA Briefing Document for the Oncologic Drugs Advisory Committee (ODAC) meeting held on January 7, 2015. 2015d. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM428781.pdf.
FDA. Guidance for industry: consideration in demonstrating interchangeability to a reference product. Silver Spring: The United States Food and Drug Administration; 2017a.
FDA. Guidance for industry: statistical approaches to evaluation analytical similarity. Silver Spring: The United States Food and Drug Administration; 2017b.
Glass GV. Primary, secondary, and meta-analysis of research. Educ Res. 1976;10:3–8.
Hussein Z, Patterson KE, Lamm JE, Kavanauch JH, Grannemen R. Effects of infusion duration on valproate pharmacokinetics. Biopharm Drug Dispos. 1993;14:389–99.
Kurki P, van Aerts L, Wolff-Holtz E, Giezen T, Skibeli V, Weise M. Interchangeability of biosimilars: a European perspective. BioDrugs. 2017;31(2):83–91.
Lehman EL. Testing hypothesis. 2nd ed. New York: Wiley; 1996.
Medoza C, Ionescu D, Radiére G, Rémazat C, Young KE, Toumi M. Biosimilar substitution policies an overview value. Health. 2015;18(7):A525.
Satterthwaite FE. An approximate distribution of estimates of variance components. Biom Bull. 1946;2:110–4.

Stangl D, Berry D. Meta-analysis in medicine and health policy. New York: Chapman & Hall/CRC/Press; 2000.

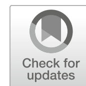
Tsui K, Weerahandi S. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. J Am Stat Assoc. 1989;84:602–7. Corrections: 86 (1991), p. 256).

Weerahandi S. Generalized confidence intervals. J Am Stat Assoc. 1993;88:899–905. Corrections: 89 (1994), 726.

Whitehead A. Meta-analysis of controlled clinical trials. Chichester: Wiley; 2002.

Yang H, Zhang J, Yu B, Zhao W. Statistical methods for immunogenicity assessment. New York: Chapman & Hall, CRC/Press; 2016.

Chapter 21
Extrapolation



Peijuan Zhu and Andrej Skerjanec

Abstract A major advantage of the biosimilar 351(k) BLA pathway compared to the 351(a) BLA innovator biologic pathway is the term "extrapolation". Unlike 351(a), when pursuing 351(k) pathway the applicant need not conduct clinical studies in every indication that the proposed biosimilar product is intended to treat. Instead, the applicant usually conducts clinical evaluations in one or two indications, then provide scientific justification for extrapolating clinical data to support a determination of biosimilarity for each condition of use for which licensure if sought. This approach is being adopted by many regulatory agencies including US FDA, European EMA, Japanese PMDA and Canadian CFIA are also adopting similar approaches.

The underlying rationale behind the concept of extrapolation is the scientific principle that protein structure determines the molecular function and ultimately clinical PK/PD, efficacy and safety of the proposed biosimilar. Based on the scientific principle, the key considerations of extrapolation include: 1) the magnitude of residual uncertainty or the analytical/functional differences between the proposed biosimilar product and the reference product; 2) mechanism of action of the biologic product in each indication and the justification that the residual uncertainty will not lead to any significant difference in clinical efficacy and safety in indications sought by extrapolation; 3) PK characteristics and distribution/clearance mechanism in each indication and the justification that the residual uncertainty will not lead to any significant difference in clinical PK or bio-distribution in indications sought by extrapolation; 4) safety and immunogenicity profile/mechanism in each indication and the justification that the residual uncertainty will not lead to any significant difference in clinical safety and immunogenicity in indications sought by extrapolation.

P. Zhu (✉)
Janssen Research and Development, Inc., Raritan, NJ, USA
e-mail: pzhu9@its.jnj.com

A. Skerjanec
Sandoz AG, Basel, Switzerland
e-mail: andrej.skerjanec@sandoz.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_21

In this chapter, we will discuss the above concepts of extrapolation in details, providing scientific background and real-world examples that showcase how extrapolation is justified in approved biosimilar products.

Keywords Biosimilar · Extrapolation · Regulatory · Clinical pharmacokinetics (PK) · Pharmacodynamics (PD) · Residual uncertainty · Mechanism of action (MOA) · Safety · Immunogenicity

Introduction

When applying for a license of a biologic, a major advantage of the biosimilar 351(k) biologics license application (BLA) pathway compared to the 351(a) BLA originator biologic pathway is the opportunity for “extrapolation” of indications beyond those directly studied. Unlike 351(a), when pursuing 351(k) pathway the sponsor need not conduct clinical studies in every indication(s) that the proposed biosimilar product is intended to treat. Instead, the sponsor “may seek licensure of the proposed product for one or more additional conditions of use for which the reference product is licensed”, if the sponsor provides “sufficient scientific justification for extrapolating clinical data to support a determination of biosimilarity for each condition of use for which licensure is sought” (FDA’s Guidance for Industry: Scientific considerations in demonstrating biosimilarity to a reference product) (FDA 2015). In addition to FDA, most regulatory agencies such as EMA, Japanese PMDA and Canadian CFIA, etc. are adopting similar approaches.

The underlying rationale behind extrapolation is the scientific principle that protein structure determines the molecular function and ultimately clinical PK/PD, efficacy and safety of the biologic drug. With the advancement of state-of-the-art physical- and chemical-analytical technologies, protein structures can be characterized using close examination of structural features, and even the minute differences can be detected. Using these technologies, a proposed biosimilar product can be thoroughly characterized and compared with the reference product to demonstrate structural (biochemical and biophysical) similarity and to identify any minor differences, if any, termed as the residual uncertainty. The residual uncertainty is then assessed using *in vitro* binding and functional assays to show that it does not impact molecular function and that the proposed biosimilar product is highly similar in terms of ligand binding, off-target binding and downstream cell signaling, etc. In addition, clinical pharmacokinetic (PK) and pharmacodynamic (PD) equivalence test and in most cases efficacy and safety data in a sensitive patient population, i.e. a population most likely to show potential differences if any, between the proposed biosimilar product and the reference product, will confirm the biosimilarity observed at the analytical and functional levels. If it is known that the different indications for which the reference product is licensed have a similar proposed mechanism(s) of actions (MOAs), and the proposed biosimilar product

is shown to be highly similar to the reference product, a strong justification of extrapolation to these indications can be established.

It is important to note that extrapolation is not directly between the indications, but rather is based on structural-functional similarity and the scientific understanding of how the structural and functional properties affect the MOAs in different indications. Even if the MOAs are different between indications or some of MOAs are unknown in some of the indications, it does not automatically preclude a biosimilar from being licensed for all indications that the reference product is licensed for based on high similarity in physical-chemical and functional data. However, it is well acknowledged that in such cases the justification for extrapolation is more difficult and most likely requires extremely close analytical similarity in physical-chemical and functional data with no or minimal residual uncertainty.

The scientific principles forming the basis for extrapolation include:

- The magnitude of residual uncertainty or the analytical/functional differences between the proposed biosimilar product and the reference product.
- Degree to which MOA is known for the reference product, and the similarity of MOAs of the reference drug in each indication as a basis for assessing similarity in clinical PD/efficacy/safety in tested indications; and the justification that the residual uncertainty will not lead to any significant difference in clinical efficacy and safety in indications sought by extrapolation.
- PK characteristics and distribution/clearance mechanism of the reference drug in each indication, high similarity in clinical PK/PD comparisons of the proposed biosimilar and the reference product in tested indications, and the justification that the residual uncertainty will not lead to any significant difference in clinical PK or bio-distribution in indications sought by extrapolation.
- Safety and immunogenicity profile/mechanism of the reference drug in each indication, sufficient similarity in clinical safety/immunogenicity comparisons of the proposed biosimilar and the reference product in tested indications, and the justification that the residual uncertainty will not lead to any significant difference in clinical safety and immunogenicity in indications sought by extrapolation.

In this chapter, we will discuss each of the above mentioned aspects in more detail, provide scientific background and introduce real world examples that illustrate how the justifications of extrapolation is established for approved biosimilar products such as Celltrion's infliximab biosimilar and Amgen's adalimumab biosimilar. In addition, we will also compare different opinions among regulatory bodies across the world and share their explanations and application of the extrapolation concept. We will use these differences to explain why Health Canada approved only a subset of indications at the initial approval of Celltrion's infliximab biosimilar, but subsequently added other indications to enable extrapolation to all indications that the reference product is licensed in.

The Regulatory Requirement for Extrapolation

FDA

In FDA's guidance for industry (FDA 2015), it is clearly stated that "If the proposed product meets the statutory requirements for licensure as a biosimilar product under section 351(k) of the PHS Act based on, among other things, data derived from a clinical study or studies sufficient to demonstrate safety, purity, and potency (similarity) in an appropriate condition of use, the sponsor may seek licensure of the proposed product for one or more additional conditions of use for which the reference product is licensed". The scientific justifications for extrapolation may include the follow items:

- The MOA(s) in each condition of use for which licensure is sought and this may include:
  - The target/receptor(s) for each relevant activity/function of the product
  - The binding, dose/concentration response, and pattern of molecular signaling upon engagement of target/receptor(s)
  - The relationships between product structure and target/receptor interactions
  - The location and expression of the target/receptor(s)
- The PK and bio-distribution of the product in different patient populations (relevant PD measures may also provide important information on the MOA.)
- The immunogenicity of the product in different patient populations
- Differences in expected toxicities in each condition of use and patient population (including whether expected toxicities are related to the pharmacological activity of the product or to off-target activities)
- Any other factor that may affect the safety or efficacy of the product in each condition of use and patient population for which licensure is sought

Although FDA emphasized the evaluation of potential differences between indications, FDA further explained that "differences between conditions of use with respect to the factors described above do not necessarily preclude extrapolation." FDA is seeking a "totality of the evidence" approach on extrapolation and is willing to accept scientific justifications for extrapolations, even though potential differences between indications may exist.

FDA also recommend that sponsors conduct clinical studies for biosimilar in "a condition of use that would be adequately sensitive to detect clinically meaningful differences between the two products". This can be used to justify extrapolation of biosimilarity for conditions of use that are less or equally sensitive. On the other hand, it will be difficult to extrapolate to more sensitive indications when clinical studies of biosimilarity were conducted in less sensitive indications due to the fact that potential clinical differences may become more visible in the more sensitive indications.

EMA

EMA's regulatory guidance for monoclonal antibodies (mAb) (EMA 2012) outlined a similar concept to extrapolation as that of FDA. It states that "Extrapolation of clinical efficacy and safety data to other indications of the reference mAb, not specifically studied during the clinical development of the biosimilar mAb, is possible based on the overall evidence of comparability provided from the comparability exercise and with adequate justification. If pivotal evidence for comparability is based on PD and for the claimed indications different mechanisms of action are relevant (or uncertainty exists), then applicants should provide relevant data to support extrapolation to all claimed clinical indications. Applicants should support such extrapolations with a comprehensive discussion of available literature including the involved antigen receptor(s) and mechanism(s) of action."

An important point that the EMA guidance stresses, is that if a reference antibody drug is licensed both as an immuno-modulator and as an anticancer antibody, e.g. rituximab, the scientific justification for extrapolation is more challenging. The basis for such extrapolation requires an extensive quality and non-clinical database, including potency assay(s) and *in vitro* assays that cover the functionality of the molecule, supplemented by relevant clinical data which may be needed in both the inflammatory disease and oncology indications. In the case of Rixathon (the biosimilar rituximab), its EMA approval and extrapolation to all indications heavily relied on the clinical PK/PD, safety and efficacy data in both rheumatoid arthritis (PK/PD as primary endpoint and efficacy as secondary endpoint) as well as follicular lymphoma (efficacy as primary endpoint and PK/PD as secondary endpoint) (Jurczak et al. 2016; Jurczak et al. 2017; Smolen et al. 2017). The extrapolation of safety including immunogenicity data also requires significant data package, which includes actively controlled safety data collected pre-authorization and, if available, pharmacovigilance from regions where the biosimilar is already marketed. With regards to MOAs, several mechanisms may play a role in the different clinical conditions, e.g. ADCC may be more important in some indications than in others, and this will be the case for the biosimilar just as it is the case for the originator. Nonetheless, the more we understand the MOAs in different clinical conditions scientifically, the better we may be able to design and conduct non-clinical and clinical studies to demonstrate biosimilarity and justify extrapolation.

Health Canada

Health Canada's Guidance document outlines a position that is similar to FDA and EMA (Canada 2015). It states that "extrapolation should be justified based on: mechanism(s) of action; pathophysiological mechanism(s) of the disease(s) or conditions involved; safety profile in the respective conditions and/or populations; and, clinical experience with the reference biologic drug. A detailed scientific rationale that appropriately addresses the benefits and risks of such a proposal

should be provided to adequately support the data extrapolation". A case study involving Health Canada will be discussed in section "Case Study 1: Infliximab Biosimilar CT-P13, Biosimilar of Remicade®".

Scientific Considerations on Extrapolations

Consideration on Mechanism of Actions (MOAs) in Different Indications and Extrapolation of Efficacy and Safety

Mechanism of Actions for Hormonal Protein Drugs

Since hormonal protein drugs, such as human growth hormone (hGH) somatropin, generally have highly similar structure and function as the corresponding endogenous hormones, their mechanism of actions are usually identical to that of the endogenous hormones with the same binding receptor and down-steam biological effects.

Mechanism of actions (MOAs) for antibody drugs

The five major mechanism of actions for antibody drugs are well summarized in a review paper by Suzuki et al. (2015) (Fig. 21.1).

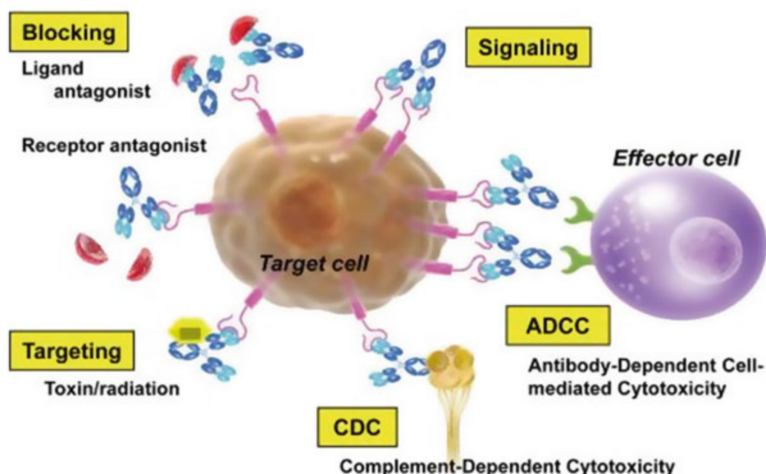


Fig. 21.1 Mechanisms of action of therapeutic antibodies (Suzuki et al. 2015). Note: Reproduced with permission of the Japanese Society of Toxicologic Pathology from Suzuki, Masami et al. Therapeutic antibodies: their mechanisms of action and the pathological findings they induce in toxicity studies. J Toxicol Pathol 28: 233-39, 2015

Neutralization: the act of neutralization blocks the pathophysiological function of the target molecules. In this case, antibodies bind to the ligand or receptor that is expressed on the cell surface and block the target signaling pathway. When the signaling in the tumor through these ligands or receptors is diminished, it can result in cellular activity being lost, proliferation being inhibited, pro-apoptotic programs being activated, or cells being re-sensitized to cytotoxic agents (Cavallo et al. 2007; Suzuki et al. 2015).

ADCC: To trigger ADCC, the Fv binding domain of an antibody binds to a specific antigen expressed on the surface of a target cell. The antibody is then able to recruit immune-effector cells (such as macrophages and NK cells) that express various receptors able to bind to the Fc region of the antibody and thus activate the immune-effector cells to lyse the target cell (Zafir-Lavie et al. 2007; Suzuki et al. 2015).

CDC: CDC is triggered when the C1 complex binds the antibody–antigen complex, activates a cascade of complement proteins, and causes a formed complex to attack the membrane of target cells, resulting in lysis of the target cell (Zafir-Lavie et al. 2007; Suzuki et al. 2015). Both ADCC and CDC are interactions that involve components of the host immune system and, among the therapeutic antibodies being developed for cancer, there are presumably products that utilize more than one mechanism (ADCC, CDC, and neutralizing functions) in their pharmacological actions.

Drug delivery carrier: Antibodies can be applied as drug delivery carriers when conjugated to radioisotopes, toxins, drugs or cytokines (Zafir-Lavie et al. 2007; Suzuki et al. 2015). The advantage of these conjugates over conventional drugs is that cytotoxic agents can be delivered directly and at higher local concentrations to tumor tissues, without causing damage to normal cells.

Antibodies that bind or cross-link two target molecules or two target cells and thus stimulate downstream biological activities are also under research.[1] The BiTEs® (bispecific T cell engagers) are biologics in this category and blinatumomab is one example that has been commercialized.

Structural Residual Uncertainty that Might Impact the MOAs

Due to the complexity of antibody structure, especially the complexity of post-translational modifications such as glycosylation, yielding many different variations of the same antibody, each antibody drug is essentially a mixture of different variations whose relative abundance could vary slightly from batch to batch. Therefore, it is technically impossible to make an exact copy of an antibody drug. Residual uncertainties in the antibody structure, minor or major, could be discovered during the physical-chemical characterization. Many glycosylations have potential impacts on the MOA as well as pharmacokinetics of the antibody. Table 21.1 illustrates

[1] BiTEs® is a trademark of Amgen Inc., Thousand Oaks, CA, USA.

Table 21.1 Summary of key impacts of glycosylation on the PK and PD of mAb and Fc-fusion proteins (Liu 2015)

Glycan	Impact
Mannose	• Increases the clearance of mAb • Enhances FcγRIIIa binding/ADCC of mAb • Reduces C1q binding/CDC of mAb
Fucose	• Interferes with binding to FcγRIIIa • Defucosylation enhances FcγRIIIa binding/ADCC activity
Galactose	• Exposed galactose may increase the clearance of mAb • Enhances CDC of mAb
GlcNac	• Bisecting GlcNAc enhance FcγRIIIa binding/ADCC • Increases the clearance of Fc-fusion proteins
Sialic acid NANA	• Critical for reducing the clearance of Fc-fusion proteins • Anti-inflammatory activity
Sialic acid NGNA	• Interferes with FcγRIIIa binding and reduce ADCC activity of mAb • May be immunogenic in humans
Galα1–3Galβ1–4GlcNAc-R	• Immunogenic in humans and may induce anaphylaxis

the important glycosylations and their potential impact on the pharmacology and pharmacokinetics of mAbs (Liu 2015).

MOAs of Efficacy and Safety in Different Indications, a Major Challenge for Extrapolation

The existence of different MOAs and various combinations of them in different indications could lead to significant complexity in the justification for extrapolation. It is not only that the major MOAs for efficacy and safety may be different from indication to indication; but also it is possible that multiple MOAs are involved, with the extent of their contributions being different between indications. In a worst case scenario, the exact MOAs may be unknown or less understood in some indications, making the justification of extrapolation more difficult, especially when residual uncertainty was identified for some key quality attributes that might trigger concerns.

However, it is important to note that the justification based on MOA is only to the extent of what is known for the reference. The biosimilar sponsor is not required to establish a MOA that is unknown to the reference. And the extrapolation is not directly between the indications, but rather is based on structural-functional similarity and the scientific understanding of how the structural and functional properties affect the MOAs in different indications. Even if the MOAs are different between indications or some of MOAs are unknown in some of the indications, it does not automatically preclude a biosimilar from being licensed for all indications the reference product is licensed based on high similarity in physical-chemical and functional data. However, it is well acknowledged that in such cases the justification

for extrapolation is more difficult and most likely requires extremely close analytical similarity in physical-chemical and functional data with no or minimal residual uncertainty.

For example, in the case of infliximab (Remicade®),[2] it is known that neutralization of TNFα is adequate to drive a therapeutic response in rheumatoid arthritis (RA), ankylosing Spondylitis (AS), Psoriatic Arthritis (PsA), and Plaque Psoriasis (PsO). However the accurate MOAs of infliximab in inflammatory bowel diseases (IBD) such as Crohn's disease and ulcerative colitis are less understood. At the same time, Celltrion's biosimilar Remsima®[3]/Inflectra®[4] had a key residual uncertainty of slightly lower ADCC readout attributed to the slightly lower level of G0, afucosylated glycans. In Celltrion's initial submission to EMA and Health Canada, Celltrion's strategy was to argue that ADCC is not an important mediator of the efficacy of their product (or of Remicade®) in inflammatory bowel diseases. This was accepted by EMA, which granted Remsima®/Inflectra® approval for all indications that Remicade® was licensed in. However, Health Canada had a different opinion and did not grant the extrapolation to inflammatory bowel diseases (see section "Regulatory Response to the Extrapolation Justification and Celltrion's Follow-up Data Package"). It was only after Celltrion showed a positive risk-benefit of Remsima®/Inflectra® in patients with inflammatory bowel diseases in a subsequent submission in 2015 did Health Canada finally approve Remsima®/Inflectra® for inflammatory bowel diseases.

Considerations on Extrapolation of Pharmacokinetics (PK)

If the PK profiles of the reference biologics drug is similar across indications with common absorption, distribution and clearance mechanisms, and if minimal residual uncertainty is identified during the physical-chemical and functional comparison, the justification of PK extrapolation would be straight-forward.

However, some antibody drugs may have slightly different PK profiles in different indications. In that case, the extrapolation of PK equivalence to other indications can potentially be justified based on the structural and functional similarity between the proposed biosimilar and the reference drug, as well as the scientific understanding of how the structure and function of antibody drugs affect their pharmacokinetics profiles. The clearance of antibody drugs is usually through two pathways: the non-specific immunoglobulin (IgG) clearance pathways and the target mediated drug disposition (TMDD) pathway. Some antibody drug's clearance involves significant TMDD while others may be eliminated only by the non-specific pathways.

[2]Remicade® is a trademark of Janssen Biotech Inc., Horsham PA, USA.

[3]Remsima® is a trademark of Celltrion Inc., Incheon, South Korea.

[4]Inflectra® is a trademark of Celltrion Inc., Incheon, South Korea and Pfizer Inc, New York City, NY, USA.

The non-specific immunoglobulin G (IgG) clearance pathways are reported to be mediated by binding to the neonatal Fc receptor (FcRn), a major histocompatibility complex class-1-related receptor (salvage pathway) or binding to various Fc receptors (FcγRs) expressed by various phagocytic cells (effector function pathway) (Tabrizi et al. 2006). If the proposed biosimilar and the reference drug are highly similar in structure and they bind to FcRn and various FcγRs with similar affinity, their elimination by the non-specific IgG clearance pathways is expected to be comparable regardless of patient populations.

The TMDD pathway is mediated by binding to the target-receptor on target cells. It has been reported that *in vitro* target binding is predictive of *in vivo* drug elimination via TMDD for antibody drugs (Singh et al. 2015). Since TMDD is mediated by target binding, the similarity in target binding and FcRn and FcγRs binding combined with comparability of one or more appropriate PD markers downstream of the target binding might provide substantial evidence that the target mediated drug disposition (TMDD), as a consequence of target binding and the target-binding-triggered downstream events, is also expected to be comparable regardless of patient populations.

Clinically, if PK comparability (equivalence test and/or descriptive data) is demonstrated in more than one indication, these data will provide additional evidence that PK equivalence can be extrapolated to different indications.

Considerations on Extrapolation of Immunogenicity

Immunogenicity comparisons are usually conducted descriptively in clinical studies, especially for biologics that have low immunogenicity incidences, due to the difficulty with which to sufficiently power the immunogenicity endpoints for equivalence test. Structural residual uncertainty that can potentially increase immunogenicity, combined with a descriptively higher incidence of immunogenicity in one or more clinical studies, which may be a chance event, can trigger a lot of concerns and will be difficult to defuse. Therefore, such residual uncertainty should be avoided as much as possible during the CMC and product development. The key quality attributes to pay attention to include but are not limited to amino acid sequence, aggregates, folding, disulfide bridges, free cysteines, degradation, host cell proteins, leachables/extractables, glycosylation, and α-1,3-galactose.

If no major immunogenicity concern is raised from structural and functional comparisons, the extrapolation of immunogenicity can usually be justified with descriptive clinical immunogenicity comparisons in one or more indications. It is recommended to assess immunogenicity in all clinical studies, and it is important to make sure that at least one study is conducted in an indication sensitive for immunogenicity assessment, e.g. an indication that has a relatively high immunogenicity rate and the patients are not immuno-compromised.

It is not only the incidence of immunogenicity that matters, but also the nature of the immunogenicity, e.g. neutralizing or non-neutralizing. Immunogenicity is

not to be evaluated alone, but rather to be combined with an assessment of its potential impact on clinical PK/PD, efficacy and safety. For example, the difference in the incidence of neutralizing antidrug antibodies can trigger concerns because neutralizing antidrug antibodies can reduce clinical PD/efficacy, and therefore additional scrutiny on the PD and efficacy data may be required.

Using Modeling and Simulation as Potential Justifications for Extrapolation

Population modeling and simulation could potentially be helpful for extrapolation justifications, especially when the model is mechanism based and is well established in the disease/patient population.

One example is extrapolation of PK to different indications. In certain indications, the expression levels of the target varies significantly, e.g. CD20 expression in different oncology indications, and therefore the extent of target mediated drug disposition (TMDD) are different, leading to different PK profiles in different patient populations. In such case, if the proposed biosimilar has the residual uncertainty of a slightly different target binding affinity, Health Authorities might question the extrapolation of PK from an indication with low extent of TMDD to another with high extent of TMDD. To address this issue, a mechanism based PK model incorporating the target levels and the binding kinetics can be established for the proposed biosimilar as well as the reference product, respectively, using existing clinical PK data in tested indications. The model for the reference product can then be adapted for the untested indication with high extent of TMDD by modifying the parameters that reflect the different target levels. This adapted model can be validated using published PK data of the reference drug in the untested indication. Once the adaptation process is validated, the PK model for the proposed biosimilar can be adapted in the same way, and then used to simulate the expected PK profile for the proposed biosimilar in the untested indication, and to help evaluate how the slight different in target binding affinity might impact clinical PK of the proposed biosimilar in the untested indications. If the impact is minimal and high PK similarity is predicted between the proposed biosimilar and the reference product for the untested indication, it can be used as potential justification that the PK equivalence be extrapolated to other indications despite differences in target expression level and the extent of TMDD.

Extrapolation of PD/efficacy using modeling and simulation can also be implemented provided the disease mechanisms are well understood and well-established mechanistic PD/efficacy models exist. The key parameters of residual uncertainty can be built into the PD/efficacy models for different indications and validated and tested using a similar process as described for the PK case described above. If the simulation results show that the key parameters of residual uncertainty are not expected to result in significant difference in PD/efficacy readouts in other untested indications, they can be used as an argument for the extrapolation of

PD/efficacy to the other indications. The potential limitation of this approach is that well-established mechanistic PD/efficacy models are often not available for many diseases.

Case Study 1: Infliximab Biosimilar CT-P13, Biosimilar of Remicade®

Remicade® (infliximab) was approved for inflammatory diseases such as rheumatoid arthritis (RA), ankylosing Spondylitis (AS), Psoriatic Arthritis (PsA), and Plaque Psoriasis (PsO), as well as inflammatory bowel diseases (IBD) such as Crohn's disease and ulcerative colitis. Clinical program for the proposed biosimilar Inflectra®/Remsima® (CT-P13) was conducted in RA and AS, and the extrapolation to IBD was a significant challenge, especially given the structural residual uncertainty of lower levels of afucosylated glycans which might lead to lower ADCC activities. The data discussed in this session are from CT-P13 Summary of Product Characteristics (EMA 2013) and the briefing book for FDA Advisory Committee Meeting (Celltrion 2016; FDA 2016a)

Analytical and Functional Characterization and Residual Uncertainties

Celltrion conducted an extensive 3-way evaluation of physicochemical and structural attributes of CT-P13 (proposed biosimilar) and US and EU sourced reference product, using over 20 test methods to analyze attributes of each product. The structural similarity studies included an extensive comparative analysis of primary, secondary, and tertiary structure, multimers and fragments, charge variants, glycan profiles, and other post-translational modifications. Excipients were also analyzed. Statistical analysis of the physicochemical test data from 3-way biosimilarity studies was carried out and it was concluded that CT-P13 was highly similar to US and EU sourced reference product in most physicochemical attributes, despite some residual uncertainties.

Since these residual uncertainties could potentially have an impact on the biologic activity or immunogenicity, they were further investigated using rational approaches, either by appropriate *in vitro* studies or by assessment in clinical studies, to demonstrate that these residual uncertainties did not have any significant impact in target binding or biological activities, and would not alter clinical efficacy or safety. The list below illustrates in more details what each residual uncertainty was, and how each residual uncertainty was investigated and justified as not compromising biosimilarity.

- A lower level of intact Immunoglobulin G (IgG) was mainly due to a higher proportion of non-assembled forms. The main fragment (H2L1) was partially purified and tested using *in vitro* assays to show similar biological activities.
- A lower level of charge variants with 2 C-terminal lysine residues and more isoforms with no C-terminal lysine or a single C-terminal lysine residue were found. Charge variants were purified and characterized using *in vitro* assays to show similar biological activities
- A slightly lower level of monomer and corresponding higher level of High Molecular Weight (HMW) forms (0.8% for CT-P13 and 0.2% for US-sourced reference product) were observed. The slightly higher level of multimers had no impact on TNFα binding or biological activities based on *in vitro* assays. In addition immunogenicity assessment was incorporated in all clinical studies and no differences were observed.
- A slightly higher level of glycation was observed for CT-P13. The slightly higher level of glycation was shown to have no impact on TNFα binding or biological activities based on *in vitro* assays.
- A slightly lower level of G0, afucosylated glycans was observed for CT-P13. Subsequent investigation in biological assays indicated that the lower level of afucosylated glycans in CT-P13 resulted in a slightly lower mean binding affinity to Fc-gamma Receptor 3a (FcγRIIIa). This difference was thought to have resulted in the lower ADCC activity in the most sensitive experimental *in vitro* model using NK cells of patients suffering from Crohn disease (CD) and with high affinity genotypes (V/V and V/F). This residual uncertainty was by far the most important residual uncertainty from the point of the Healthy Authorities because the binding affinity to Fc-gamma Receptor 3a (FcγRIIIa) is known to be linked to effector functions and has a potential impact on important mechanisms of actions such as ADCC or CDC. Celltrion undertook additional efforts that included a series of pre-clinical as well as clinical studies to address this residual uncertainty. These additional studies are discussed in greater details in sections "Additional Non-clinical Evaluation to Address the Residual Uncertainty of Afucosylated Glycans" and "Clinical Program".

Additional Non-clinical Evaluation to Address the Residual Uncertainty of Afucosylated Glycans

In order to address the residual uncertainty around the lower amount of afucosylated species, which resulted in lower binding to FcγRIIIa and hence lower ADCC activity in the most sensitive experimental *in vitro* model using NK cells of patients suffering from Crohn's disease (CD) and with high affinity genotypes (V/V and V/F), a series of additional *in vitro* experiments were conducted using experimental models that represented the patho-physiological conditions of specific indications and putative mechanisms of action of infliximab.

Celltrion repeated the experiments in the presence of serum of a CD patient, by using peripheral blood mononuclear cells preparations (rather than isolated NK cells) or by using whole blood. It was shown that in whole blood, which represents a more physiologically relevant environment, the differences in binding to FcγRIIIa and in ADCC activity were abolished by competition from plasma IgGs, soluble factors, immune complexes and the presence of mixed cell populations expressing multiple FcRs. These results suggest that the difference in ADCC seen in the most sensitive *in vitro* model may not translate to any difference in clinical outcomes.

Celltrion also conducted *in vitro* experiments in an experimental system representative of the inflammatory conditions *in vivo*, using LPS-stimulated monocytes as target cells and PBMCs as effector cells. In these experiments, no ADCC response was detected with either products and regardless of the donor cells used (healthy volunteer or CD patient).

Experiments using mixed lymphocyte reaction from FcγRIIIa genotype matched PBMCs (either healthy donors or CD patients) were conducted to show that no differences in the proportion of regulatory macrophages induced and in the inhibition of T-cell proliferation could be detected between CT-P13 and and the reference product, regardless of the FcγRIIIa genotype of the donor cells.

In addition to these experiments intended to compare Fc-related functions, Celltrion performed experiments using a model relevant to IBD, intended to compare Fab-related functions. The data showed a dose-dependent suppression of cytokine secretion from human epithelial cells, stimulated by a mixture of stimulators and this effect was comparable between both products (Table 21.2).

Clinical Program

The clinical development program to show biosimilarity between CT-P13 and reference product consisted of two pivotal trials. The two studies are discussed in more detail in subsequent sections (Table 21.3).

CT-P13 1.1 PK Equivalence Study in AS Patients

The primary objective of the CT-P13 1.1 study was to demonstrate comparable PK of CT-P13 and the reference product at steady state with the dosing regimen of 5 mg/kg dose at Weeks 0, 2, and 6 and then every 8 weeks up to Week 54. Co-primary PK endpoints were set as AUC_{τ} and $C_{max,ss}$ between Weeks 22 and 30. The 90% CIs of the geometric means ratios for both AUC_{τ} and $C_{max,ss}$ lied between 93% and 116%, well contained within the standard bioequivalence interval of 80–125%; this demonstrated that the PK of infliximab is equivalent between CT-P13 and Remicade® at the dose of 5 mg/kg. The main secondary PK endpoints such as T_{max}, $C_{min,ss}$, $T_{½}$, CL_{ss}, V_{ss} between Weeks 22 and 30, as well as C_{max} and C_{min} after the 9 treatment doses, were also comparable in the CT-P13 and the reference treatment groups.

Table 21.2 Additional *in vitro* studies to address the residual uncertainty of lower amount of afucosylated species in CT-P13 (Celltrion 2016; FDA 2016a)

Test method		Key findings
Comparative antibody-dependent cell-mediated cytotoxicity (ADCC) of Remsima and Remicade using tmhTNFα-Jurkat cells as target cells and human PBMC as effector cells		Remsima and Remicade had comparable ADCC activity and no statistically significant differences were detected
Comparative ADCC of Remsima and Remicade using tmhTNFα-Jurkat cells as target cells and NK cells from healthy donor as effector cells		Comparable ADCC for Remsima and Remicade when NK cells from a healthy donor (genotype V/F) were used as effector cells
Evaluation of regulatory macrophage function	Suppression of T cell proliferation by induced regulatory macrophages in mixed lymphocyte reaction (MLR) assay	Inhibition of T cell proliferation of PBMCs from healthy donors and CD patients was shown to be comparable and dose dependent for Remsima and Remicade
	Quantitation of the induced regulatory macrophages by FACS analysis	Induction of regulatory macrophages in a 2-way allogenenic MLR using FcγRIIIa genotype matched PBMCs, from either healthy donors or CD patients, was shown to be comparable for Remsima and Remicade
	Induced regulatory macrophage-mediated wound healing of colorectal epithelium cells	Promotion of *in vitro* wound healing of colorectal epithelial cells by regulatory macrophages from healthy donors and CD patients (induced by Remsima or Remicade) in the MLR assay was comparable (Celltrion 2016; FDA 2016a)

Comparison of ADCC activity between Remsima and Remicade using transfected Jurkat cells as target cells and either PBMCs or NK cells from CD patients as effector cells		No differences in ADCC activity were detected using PBMC from CD patients (V/F or F/F genotype). Differences in ADCC with Remsima and Remicade were seen when NK cells from CD patients were used as effector cells. Effect was FcγRIIIa genotype specific; differences were observed with V/V and V/F, but not F/F genotypes
Comparison of ADCC effect between Remsima and Remicade using transfected Jurkat cells as target cells and whole blood from healthy donor or CD patients as effector cells		No differences in ADCC were seen between various batches of Remsima and Remicade
Comparison of ADCC between Remsima and Remicade using LPS-stimulated monocytes from healthy donor or CD patient as target cells and PBMC as effector cells		No ADCC activity was seen with Remsima and Remicade when PBMCs from a healthy donor (V/F) or a CD patient (V/F) were used as effector cells and LPS-stimulated monocytes were used as target cells
Effect of blocking soluble TNFα *in vitro* IBD model	Suppression of cytokine secretion in epithelial cell line by blocking soluble TNFα	Suppression of pro-inflammatory cytokine (IL-6 and IL-8) secretion from co-stimulated epithelial cell line was shown to be comparable and dose dependent for Remsima and Remicade; no statistical difference in pro-inflammatory cytokines suppression was found
	Suppression of apoptosis in epithelial cell line cells by blocking soluble TNFα	Suppression of epithelial cell line apoptosis was shown to be comparable for Remsima and Remicade

Table 21.3 Overview of clinical studies in the clinical development program (Celltrion 2016; FDA 2016a)

Protocol	Design	Objectives	Treatment	Study population
CT-P13 1.1 PK equivalence (study name: PLANET AS)	Prospective Phase 1, randomised, double-blind, multicentre, multiple single-dose i.v. infusion, parallel-group	Primary: to demonstrate comparable PK at steady state in terms AUCτ, Cmax,ss between CT-P13 and Remicade determined between Weeks 22 and 30 Secondary: long-term efficacy, PK and overall safety up to Week 54	CT-P13 or Remicade	AS patients with active disease Planned: 246 (ratio: 1:1) Randomised: 250 CT-P13: 125 Remicade: 125
CT-P13 3.1 Therapeutic equivalence (study name: PLANET RA)	Prospective phase 3, randomised, double-blind, multicentre, multiple single-dose i.v. infusion, parallel-group	Primary: to demonstrate that CT-P13 is equivalent to Remicade, in terms of efficacy as determined by clinical response according to ACR20 at Week 30 Secondary: long-term efficacy, PK, PD, and overall safety up to Week 54	CT-P13 plus MTX or Remicade plus MTX	RA patients with active disease while receiving MTX Planned: 584 (ratio: 1:1) Randomised: 606 CT-P13: 302 Remicade: 304

With regards to immunogenicity, 44/128 patients (34.4%) in the CT-P13 arm and 39/122 patients (32.0%) in the reference product arm were reported to have at least one positive anti-drug antibody (ADA) immunogenicity test result at any time point up to week 54. Almost all antibodies were found to be neutralizing, which was expected considering the mouse (Fab)-human chimeric nature of infliximab.

In addition, secondary efficacy endpoint was assessed at Week 14 and 30. Although this efficacy evaluation was considered supportive in AS patients as it was not powered to show therapeutic equivalence, the results were comparable between treatment arms, as reflected by ASAS20 and ASAS40 responses and decreases in BASDAI and BASFI, all of which showed comparable results when comparing CT-P13 and the reference product.

CT-P13 3.1 Efficacy and Safety Comparison in Patients with Active RA

Study CT-P13 3.1 was a randomized, double-blind, parallel-group study designed to demonstrate equivalence in efficacy and safety between CT-P13 and reference product when co-administered with methotrexate (MTX) in patients with active rheumatoid arthritis. The dosing regimen was in agreement with the Remicade® US package insert and was 3 mg/kg at 0, 2 and 6 weeks then every 8 weeks increasing up to 10 mg/kg or treating every 4 weeks. The trial was designed to show equivalence of the test and reference products if the 95% CI for the difference between treatments was entirely within −15% to +15%. This clinical model was considered sufficiently sensitive to enable the detection of differences between the two products. The choice of the patient population was based on the effect size (infliximab vs. placebo) observed in the pivotal Remicade® trials, which appeared larger in the ATTRACT trial (patients with inadequate response to MTX) than in the ASPIRE trial (with MTX in the first line).

The pivotal efficacy trial (Study CT-P13 3.1) evaluated the therapeutic equivalence of CT-P13 compared to reference product in patients with active RA. The primary efficacy endpoint , defined as the proportion of patients achieving clinical response in accordance to the ACR criteria of 20% improvement (ACR20) at Week 30 was met, indicating therapeutic equivalence between the treatment arms. All secondary efficacy endpoints evaluated were also comparable between CP-P13 and the reference product.

In addition, supportive PK data were generated in the pivotal study CT-P13 3.1, providing estimates of C_{min}, C_{max}, and T_{max} in RA patients. The PK population consisted of 578 patients, i.e. 290 for CT-P13 and 288 for reference product. Generally, peak and trough levels measured after 9 doses of 3 mg/kg were similar between CT-P13 and Remicade® arms.

The size of the safety database and duration of exposure was considered appropriate for the evaluation of the general safety profile of CT-P13. The treatment-emergent adverse event patterns observed in the clinical studies were similar between the CT-P13 and the reference product arms and appeared in line with the well-characterized safety profile of Remicade®.

With regards to immunogenicity, the proportion of patients with positive immunogenicity results was similar between CT-P13 and EU-sourced reference product groups, with the proportion of ADA positive patients increasing up to Week 30, and maintained thereafter up to Week 54. The incidences of neutralizing antibody (NAb) and ADA/NAb titer levels were similar between CT-P13 and EU Remicade®.

Comprehensive Justification for Extrapolation to Indications Besides AS and RA

Justifying Extrapolation of PK Based on Clinical PK Data and Immunogenicity

Celltrion argued that there was no evidence of notable differences in the PK of infliximab across its various indications. With the two studies in AS and RA, Celltrion covered the two recommended doses of infliximab (3 and 5 mg/kg). Moreover, PK data were generated under conditions of monotherapy and in combination with MTX. Therefore, from a PK perspective, it was considered that sufficient data are available to support the extrapolation to all indications of Remicade®. In addition, the patient populations selected (AS and RA) are considered sensitive to evaluate immunogenicity; even in combination with MTX, a high level of immune response was demonstrated.

Extrapolation of Efficacy and Safety

Celltrion provided a review of the literature on the role of TNFα in disorders covered by therapeutic indications of Remicade® and the potential mechanisms of action of the various anti-TNFs as follows.

The primary mode of action is known to result from blockade of TNF receptor-mediated biological activities. Infliximab binds to soluble (s) or transmembrane (tm) TNF, thereby blocking their capacities to bind TNFR1 or TNFR2 and hence preventing cellular functions such as cell activation, cell proliferation, cytokine and chemokine production, which in turn inhibits cell recruitment, inflammation, immune regulation, angiogenesis, and extracellular matrix degradation. Several other potential mechanisms are induced by the binding of infliximab to tmTNF and include reverse signaling (inducing apoptosis or cytokine suppression) or cytotoxicity of the tmTNF-bearing cell by CDC or ADCC. While binding to sTNF and tmTNF involves the Fab region of infliximab, the latter mechanisms involve binding of the molecule to complement or effector cells through its Fc region.

It is believed that neutralization of sTNF and tmTNF is responsible of its efficacy in RA by preventing TNF from inducing TNFR-mediated cellular functions. It can also be accepted that the effects of infliximab blockade on synovial inflammation are comparable in different forms of arthritis. Such effects are believed to play a role in psoriasis plaques. However, more mechanisms were likely involved in inflammatory bowel diseases (IBD), which are related to its binding to tmTNFα and include reverse signaling and Fc-related effector functions. The relative contribution of these various effects was unknown.

The results of the extensive comparability exercise showed that the only difference between CT-P13 and Remicade® was a lower amount of afucosylated species, which resulted in lower binding to FcγRIIIa and hence lower ADCC activity in the

most sensitive experimental *in vitro* model using NK cells of patients suffering from Crohn disease (CD) and with high affinity genotypes (V/V and V/F). To address this residual uncertainty, Celltrion provided a range of arguments to conclude that the differences in the level of afucosylation and binding to FcγRIIIa are not clinically relevant. The arguments are listed below:

- In blood, the physiological environment, the differences in binding to FcγRIIIa and in ADCC activity were abolished. Celltrion demonstrated this by repeating the experiments in the presence of serum of a CD patient, by using peripheral blood mononuclear cells preparations (rather than isolated NK cells) or by using whole blood. The difference in binding affinity was overcome by competition from plasma IgGs, soluble factors, immune complexes and the presence of mixed cell populations expressing multiple FcRs. At inflammatory sites, the vascular permeability is increased, which allows for many blood components to enter the extravascular space.
- Celltrion conducted *in vitro* experiment in a system representative of the inflammatory focus *in vivo* using LPS-stimulated monocytes as target cells and PBMCs as effector cells, no ADCC response was detected with either products and regardless of the donor cells used (healthy volunteer or CD patient). Indeed, LPS-stimulated monocytes expressed much lower levels of tmTNFα compared with transfected Jurkat cells and these were not sufficient to elicit an effective ADCC response. This meant that ADCC was likely to be limited in inflammatory settings *in vivo*. It is acknowledged that to date there are no published reports describing the induction of ADCC by TNF antagonists in a patient.
- Infliximab was shown to induce a subset of regulatory macrophages (regMø), an effect that has been postulated to promote gut mucosal wound healing in IBD. This effect reflected the ability of infliximab to bind to macrophages (which express both FcγRI and FcRIIIa) through its Fc region and to activated T-cells expressing tmTNFα through its Fab region. Upon this binding, a distinct macrophage subset was induced with immunosuppressive capacities, including the production of anti-inflammatory cytokines and inhibition of T-cell proliferation. To address this MOA, Celltrion conducted an experiment using mixed lymphocyte reaction from FcγRIIIa genotype matched PBMCs (either healthy donors or CD patients) to show that no differences in the proportion of regulatory macrophages induced and in the inhibition of T-cell proliferation could be detected between CT-P13 and reference product, regardless of the FcγRIIIa genotype of the donor cells. This indicated that the difference in binding affinity did not affect the induction of regulatory macrophages. Furthermore, the monocyte/macrophages induced by CT-P13 or reference product showed the same ability to promote healing of an artificial wound made in a culture of colon epithelial cells.
- In addition to these experiments intended to compare Fc-related functions of CT-P13 and Remicade®, Celltrion performed experiments using an *in vitro* model relevant to IBD intended to compare Fab-related functions. The data showed a dose-dependent suppression of cytokine secretion from human epithelial cells

stimulated by a mixture of stimulators and this effect was comparable with both products. Likewise, CT-P13 and reference product were similarly able to suppress apoptosis of human epithelial cells by blocking soluble TNFα.

By using the above mentioned experimental models to represent the pathophysiological conditions and putative mechanisms of action of infliximab, Celltrion argued that the difference detected in the amount of afucosylated species had no clinically relevant impact on the efficacy and safety of CT-P13, in particular in IBD.

Regulatory Response to the Extrapolation Justification and Celltrion's Follow-up Data Package

Based on the above extrapolation justification, EMA and Japanese PMDA approved CT-P13 for all indications on Remicade® label.

However, Health Canada had a different opinion and did not approve the extrapolation to inflammatory bowel diseases. In Health Canada's Summary Basis of Decision (SBD) for CT-P13 (Canada 2014), it is stated that: "The sponsor provided rationale to support their position that ADCC is not an important mediator of the efficacy of their product (or of Remicade®)" in inflammatory bowel diseases; "however, after review of the sponsor's rationale for extrapolation and of literature regarding this mechanism of action, it was concluded that ADCC cannot be ruled out as a mechanism of action in the inflammatory bowel diseases (IBD). This position is supported by the observation that certolizumab pegol, another anti-TNF that lacks the ability to induce ADCC, displays only marginal efficacy in Crohn's patients compared to other anti-TNFs, namely infliximab. Therefore, since differences in ADCC have been observed between the two products and because ADCC may be an active mechanism of action for infliximab in the setting of IBD, but not in the setting of rheumatic disease (the studied populations), extrapolation from the settings of rheumatoid arthritis and ankylosing spondylitis to IBD cannot be recommended due to the absence of clinical studies in IBD." It was only after Celltrion provided additional clinical data in 2015 to show a positive risk-benefit of CT-P13 in patients known as Crohn's disease (CD) and ulcerative colitis (UC), both inflammatory bowel diseases, did Health Canada finally agreed to the extrapolation to inflammatory bowel diseases.

FDA's initial response also voiced concerns on the differences in ADCC and requested additional *in vitro* characterization of ADCC as well as "an adequate justification, including an evaluation of the role of ADCC particularly in the setting of inflammatory bowel disease, that the observed difference in ADCC is not relevant to clinical activity", which led to the delay of the FDA Advisory Committee meeting for almost one year, from March 17, 2015 to Feb 9, 2016. At the FDA Advisory Committee meeting in 2016, the additional clinical data from patients with inflammatory bowel diseases served as important evidence to help justify to

the panel the extrapolation of indications. Eventually FDA approved CT-P13 for all indications that were licensed for Remicade® as of 2016.

Case Study 2: Adalimumab Biosimilar ABP501, Biosimilar of Humira®

Humira®[5] was approved in RA, Juvenile idiopathic arthritis (JIA) in patients 4 years of age and older, PsA, AS, adult CD, UC, and PsO, while the clinical program for the proposed biosimilar ABP501 was conducted in RA and PsO. Since the physical-chemical characterization identified no major residual uncertainties, the justification of extrapolation was more straightforward in comparison with the case for CT-P13. The data discussed in this session are from the briefing books for FDA Advisory Committee meeting (Amgen 2016; FDA 2016b)

Analytical and Functional Similarity and Residual Uncertainty

The ABP501 analytical similarity assessment demonstrated a high degree of similarity with minimal analytical differences between ABP501 and the reference product. Furthermore, the functional similarity comparison, which included assessments of multiple TNFα-dependent functions that are relevant to all indications of use for adalimumab, also demonstrated high similarity.

The functional similarity of ABP501 compared to reference product was demonstrated with respect to binding and neutralizing TNFα. The binding kinetics to soluble TNFα were also demonstrated to be similar. In addition, further characterization assays were performed to confirm similarity in neutralizing TNFα activity in both NFκB-dependent and NFκB-independent signaling pathways. Thus, ABP501 was highly similar to reference product in binding to and neutralizing soluble TNFα, inclusive of both reported downstream signaling pathways. This provided a key component of scientific justification supporting extrapolation to all indications based on an understanding of the primary mechanism of action.

For the inflammatory bowel disease indications, binding of adalimumab to transmembrane TNFα might contribute to clinical efficacy via cell depletion (ADCC and CDC), decreased proliferation of transmembrane TNFα-expressing cells, or a combination of the different mechanisms. Amgen applied appropriate functional assays to test not only binding to transmembrane TNFα but also the induction of the effector functions, ADCC and CDC, and decreased proliferation in a mixed lymphocyte reaction. Similarity in these functions, along with similar binding

[5] Humira® is a trademark of AbbVie Inc., North Chicago IL, USA.

and neutralization of soluble TNFα, provided justification for extrapolation to the inflammatory bowel disease indications (ulcerative colitis and Crohn's disease).

Clinical Program

The clinical evidence supporting the similarity of ABP501 to reference product includes the following studies:

- a single-dose, PK similarity study—Study 20110217 in healthy subjects

  Healthy subjects were selected to assess PK similarity since these subjects do not receive concomitant medications and do not have medical conditions that could potentially affect PK. The approved dose of adalimumab in most adult indications is 40 mg, and therefore a 40 mg dose was selected for the study. The study demonstrated PK equivalence as assessed by AUCinf and Cmax between ABP501 and US-source reference product, as well as PK similarity between ABP501 and EU-sourced reference product and between US-sourced reference product and EU-sourced reference product.
- Study 20120262 in subjects with moderately to severely active rheumatoid arthritis

  This study demonstrated the clinical similarity in safety, efficacy, and immunogenicity for ABP501 and adalimumab in the primary efficacy analysis of ACR20 at week 24. The 90% confidence interval for risk ratio of ACR20 between ABP501 and reference product was fully contained within the pre-defined equivalence margin (0.738, 1/0.738 [i.e., 1.355]).
- Study 20120263 in subjects with moderate to severe chronic plaque psoriasis

  This study demonstrated the clinical equivalence of ABP501 and reference product in the primary efficacy analysis of PASI percent improvement from baseline at week 16, with the 95% confidence interval for the treatment difference fully contained within the pre-defined equivalence margin of $\pm$15%. The study also found that there is no increased risk with respect to safety, efficacy, or immunogenicity associated with the single transition from reference product to ABP501

Clinical Pharmacology and Justification for Extrapolation of PK

Amgen provided literature evaluation on the adalimumab PK profile in healthy subjects and across therapeutic indications for factors that could influence exposure. Of the factors investigated, the presence of anti-drug antibodies, concomitant use of MTX, body weight, and serum albumin levels were determined to affect adalimumab PK. The disease type is not a statistically significant factor affecting PK variability and the mean trough concentrations across populations were consistent (Fig. 21.2).

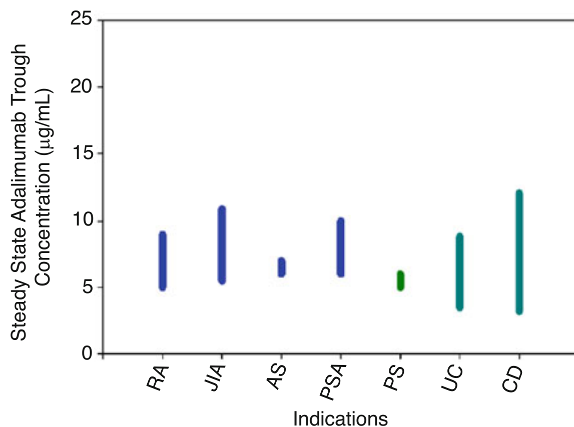


Fig. 21.2 Steady-state trough adalimumab concentration ranges in different patient populations (Amgen 2016; FDA 2016b)

The PK similarity of ABP501 to reference product in healthy subjects was demonstrated by meeting all bioequivalence endpoints. Additionally, trough PK samples that were incorporated into the clinical studies in the rheumatoid arthritis and plaque psoriasis populations also showed similarity between ABP501 and reference product. Amgen argued that these comparative PK data in sensitive and representative populations, combined with the knowledge of the PK profiles of adalimumab in different patient populations indicated that ABP501 retained PK properties similar to reference product in all indications for which ABP501 licensure is sought.

Justification for Extrapolation of Efficacy

The primary MOA of adalimumab is direct binding and blocking of TNF receptor mediated biological activities. Adalimumab binds to both soluble(s) and transmembrane (tm) TNF, thus blocking TNF binding to its receptors TNFR1 and TNFR2 and the resulting downstream pro-inflammatory cascade of events. The scientific literature indicates that this MOA is the primary MOA in RA, JIA, PsA, AS, and PsO. The data provided by Amgen showed similar TNF binding and potency to neutralize TNF-α, supporting the demonstration of analytical similarity pertinent to this MOA. Therefore, based on the above considerations, Amgen proposed to extrapolate conclusions regarding similar efficacy and safety of ABP501 and reference product in RA and PsO to JIA, PsA and AS.

TNF plays a central role in the pathogenesis of the IBD indications (Crohn's Disease and ulcerative colitis), and TNF inhibition is important in treating the diseases, as evidenced by the efficacy of the approved anti-TNF monoclonal antibodies, but the detailed cellular and molecular mechanisms involved have not been fully elucidated. However, the available scientific evidence suggests that for TNF inhibitors in IBD, in addition to binding and neutralization of sTNF, other

MOA may play a role. Binding to sTNF and tmTNF involves the Fab region of the antibody, while the other plausible mechanisms of action involve the Fc region of the molecule. To address these MOAs, Amgen provided experimental data supporting a demonstration that ABP501 and reference product were highly similar based on extensive structural and functional analytical characterization. Based on the submitted robust analytical data that evaluated attributes of ABP501 that might potentially influence its performance in IBD (i.e., the extensive structural characterization, other functional assays, binding to mTNF-α, and evaluation of a related IBD mechanism based on the activation of regulatory macrophages) , no differences were found that would preclude the conclusion that ABP501 was highly similar to reference product. Based on all above, Amgen proposed extrapolation of conclusions regarding similar efficacy and safety of ABP501 and reference product in RA and PsO to IBD.

Clinically, the response rates to adalimumab were comparable across all arthritis's indications, and the dosing regimens across the adult indications in arthritis's were identical. Study 20120262 in subjects with moderately to severely active rheumatoid arthritis demonstrated clinical equivalence in efficacy between ABP501 and reference product. Amgen argued that these efficacy results are considered predictive of similar efficacy in all other arthritic conditions of use for which ABP501 licensure was sought.

Study 20120263 in subjects with moderate to severe chronic plaque psoriasis provided a younger population with fewer comorbidities and concomitant medications, and was also a sensitive and appropriate model to detect clinically meaningful differences in efficacy between ABP501 and reference product. Clinical equivalence between ABP501 and reference product was confirmed and the efficacy results were considered predictive of similar efficacy in all other dermatologic conditions of use for which ABP501 licensure is sought.

Amgen argued that given the efficacy of ABP501 and reference product was found to be similar in the respective studies conducted in rheumatoid arthritis and plaque psoriasis populations, in addition to the highly similar analytical and functional characteristics and PK similarity, ABP501 was expected to exhibit similar efficacy in all conditions of use for which ABP501 licensure is sought.

Justification for Extrapolation of Immunogenicity

The incidence rate of anti-drug antibodies against adalimumab was generally similar across conditions of use when compared using the same immunoassay and considering the use of immuno-suppressants, though small differences are reported for different populations. The two ABP501 clinical studies provided data in rheumatoid arthritis population (with concomitant immunosuppressant therapy (methotrexate) and a population without concomitant immunosuppressant therapy (plaque psoriasis). In both studies, the expected high rates of anti-drug antibody formation based on the knowledge of adalimumab were demonstrated, providing a robust ability to detect potential differences, if they existed, between ABP501 and

reference product. Amgen argued that the immunogenicity results demonstrated similar rates of binding and neutralizing anti-drug antibody formation, covering 2 different treatment paradigms (with and without immunosuppression) and different populations of patients. Therefore Amgen concluded the immunogenicity of ABP501 is expected to be similar to reference product in all populations and indications of use for which licensure is sought.

Justification for Extrapolation of Safety

Study 20120262 in subjects with moderately to severely active rheumatoid arthritis receiving concomitant methotrexate was considered sensitive and relevant for assessing potential differences in safety profiles of ABP501 and reference product. Study 20120263, in subjects with moderate to severe chronic plaque psoriasis, assessed safety in younger subjects with fewer comorbidities and without concomitant immunosuppressive therapy, and was also considered relevant and informative for assessing differences in safety profiles of ABP501 and reference product. In addition, subjects in Study 20120263 were administered a loading dose of 80 mg of study drug, similar to the dosing regimen for Crohn's disease and ulcerative colitis, thus providing additional justification for expected similarity in indications requiring higher doses to be administered. No clinically meaningful differences in toxicities were observed between treatment groups in the rheumatoid arthritis and plaque psoriasis studies, and the adverse events from both studies were in agreement with the known safety profile of adalimumab. Given the consistency of the safety profiles in adalimumab's approved indications, and the similarity in the safety profiles in the two ABP501 clinical studies, Amgen argued that the ABP501 safety profile is expected to be the same for all indications for which ABP501 licensure is sought.

Regulatory Response to the Extrapolation Justification

Based on the totality of evidence including high similarity in physical chemical property, highly similar functional test results, highly comparable clinical PK, efficacy, safety and immunogenicity data, both FDA and EMA approved ABP501 for all indications that Humira® was licensed in.

Conclusions

Regulatory agencies such as FDA and EMA have determined that differences between conditions of use do not necessarily preclude extrapolation. A scientific justification for the proposed extrapolation can be acceptable if the totality of data provide strong evidence that no significant differences in PK, PD, efficacy, safety

and immunogenicity is expected between the proposed biosimilar and the reference product in the indications sought by extrapolation. The strength of the data package is determined by several factors.

First and foremost, the analytical characterization of high similarity coupled with high similarity in functional testing are the foundation for a solid extrapolation justification. Residual uncertainties are the key issues to be addressed before extrapolations can be possible. Therefore, reducing the residual uncertainty to the minimum would greatly facilitate the extrapolation argument.

Secondly, it is critical to thoroughly understand MOAs of the reference drug in each indication if known, in order to design the appropriate testing assays as well as to set up sensitive clinical studies to address any residual uncertainties, and justify that the residual uncertainty will not lead to any significant difference in clinical efficacy and safety in indications sought by extrapolation. Sometimes, the MOAs of the reference drugs are not fully understood and the biosimilar sponsor is usually not required to establish a MOA that is unknown to the reference. However, it may help the sponsor of the proposed biosimilar to show evidence that the MOA mediated by a quality attribute that has a residual uncertainty is not relevant in the indication sought by extrapolation, and therefore minimize the concern that the residual uncertainty may preclude extrapolation. Celltrion had adopted this approach arguing that ADCC was not a significant MOA in IBD, which was accepted by EMA and PMDA.

Thirdly, it is important to fully understand PK characteristics and distribution/clearance mechanism of the reference drug in each indication in order to design the most sensitive clinical pharmacology studies to show clinical PK (/PD) equivalence in a sensitive patient population, and then combine the two to justify that the residual uncertainty will not lead to any significant difference in clinical PK or bio-distribution in indications sought by extrapolation.

Lastly, it is important to understand the safety and immunogenicity profile/mechanism of the reference drug in each indication, design appropriate clinical evaluations to demonstrate similarity in clinical safety/immunogenicity in comparison of the proposed biosimilar and the reference product in a sensitive patient population, and justify that the residual uncertainty will not lead to any significant difference in clinical safety and immunogenicity in indications sought by extrapolation.

Extrapolation is the most important reason for pursuing the biosimilar path. Its justification varies from case to case and it requires extensive background research, careful thinking, planning and execution in order to succeed. In a way, extrapolation hinges on the totality of evidence concept, and very often it requires the entire dossier to justify extrapolation.

The two case studies mentioned in this chapter only provide a tip-of-the-iceberg glimpse of the challenges associated with extrapolation. More challenges lie ahead. A few questions that might be interesting to ask include: would it be possible to extrapolate to oncology indications when the clinical studies were conducted in inflammatory indications only, (or vice versa)? If so what additional data/justifications are required? How do we take best advantage of the PK/PD

modeling and simulation approaches to help sponsors of biosimilars justify extrapolation and whether regulatory agencies will be receptive to such approaches? How about extrapolation to additional new indications for an approved biosimilar when the reference product acquires new label indications and could there be an abbreviated path for it? As more and more proposed biosimilars are heading toward full development/submission stage, more cases for extrapolation will be available, which will provide impetus for regulatory requirement of extrapolation to evolve.

References

Amgen. Amgen briefing book for adalimumab biosimilar FDA Advisory Committee Meeting. 2016. https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM510295.pdf. Accessed Jan 18, 2017.

Canada H. Summary basis of decision (SBD) for inflectra. 2014. http://www.hc-sc.gc.ca/dhp-mps/prodpharma/sbd-smd/drug-med/sbd_smd_2014_inflectra_159493-eng.php. Accessed Jan 18, 2017.

Canada H. Guidance document: information and submission requirements for subsequent entry biologics (SEBs). 2015. http://hc-sc.gc.ca/dhp-mps/consultation/biolog/submission-seb-exigences-pbu-eng.php. Accessed Jan 18, 2017.

Cavallo F, Calogero RA, Forni G. Are oncoantigens suitable targets for anti-tumour therapy? Nat Rev Cancer. 2007;7(9):707–13.

Celltrion. Celltrion briefing book for infliximab biosimilar FDA Advisory Committee Meeting. 2016. https://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/arthritisadvisorycommittee/ucm484860.pdf. Accessed Jan 18, 2017.

EMA. Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues. 2012. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf. Accessed Jan 18, 2017.

EMA. CHMP assessment report remsima. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002576/WC500151486.pdf. Accessed Jan 18, 2017.

FDA. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. 2015. https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf.

FDA. FDA briefing book for CT-P13 infliximab biosimilar FDA Arthritis Advisory Committee Meeting. 2016a. https://www.fda.gov/downloads/%E2%80%A6/UCM484859.pdf. Accessed Jan 18, 2017.

FDA. FDA briefing book for ABP 501 adalimumab biosimilar FDA Arthritis Advisory Committee Meeting. 2016b. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm397017.pdf. Accessed Jan 18, 2017.

Jurczak W, Ilidia M, Govindbabu KS, Eduardo M, Ascuncion EM, Giri P, Nelson C, Juliana P, Luiza A, Sergey A, Dzhelil O, Pultar P, Cherfi A, Zhu P, Amersdorffer J. A phase III efficacy and safety study of the proposed rituximab biosimilar GP2013 versus rituximab in patients with previously untreated advanced follicular lymphoma. Blood. 2016;128(22):1809.

Jurczak W, Moreira I, Kanakasetty GB, Munhoz E, Echeveste MA, Giri P, Castro N, Pereira J, Akria L, Alexeev S, Osmanov E, Zhu P, Alexandrova S, Zubel A, Harlin O, Amersdorffer J. Rituximab biosimilar and reference rituximab in patients with previously untreated advanced follicular lymphoma (ASSIST-FL): primary results from a confirmatory phase 3, double-blind, randomised, controlled study. Lancet Haematol. 2017;4(8):e350–61.

Liu L. Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. J Pharm Sci. 2015;104(6):1866–84.

Singh AP, Krzyzanski W, Martin SW, Weber G, Betts A, Ahmad A, Abraham A, Zutshi A, Lin J, Singh P. Quantitative prediction of human pharmacokinetics for mAbs exhibiting target-mediated disposition. AAPS J. 2015;17(2):389–99.

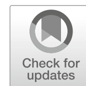
Smolen JS, Cohen SB, Tony H-P, Scheinberg M, Kivitz A, Balanescu A, Gomez-Reino J, Cen L, Zhu P, Shisha T. A randomised, double-blind trial to demonstrate bioequivalence of GP2013 and reference rituximab combined with methotrexate in patients with active rheumatoid arthritis. Ann Rheum Dis. 2017;76(9):1598–602.

Suzuki MK, Kato C, Kato A. Therapeutic antibodies: their mechanisms of action and the pathological findings they induce in toxicity studies. J Toxicol Pathol. 2015;28(3):133–9.

Tabrizi MA, Tseng C-ML, Roskos LK. Elimination mechanisms of therapeutic monoclonal antibodies. Drug Discov Today. 2006;11(1–2):81–8.

Zafir-Lavie I, Michaeli Y, Reiter Y. Novel antibodies as anticancer agents. Oncogene. 2007;26(25):3714–33.

Chapter 22
Totality of Evidence and the Role of Clinical Studies in Establishing Biosimilarity



Hillel P. Cohen, William C. Lamanna, and Martin Schiestl

Abstract The totality of evidence describes the sum of analytical, non-clinical and clinical studies used to justify regulatory approval of a biosimilar. The foundation of this approach is a detailed analytical comparison of the biosimilar and reference medicine to establish molecular sameness by use of physicochemical and functional assays. By leveraging established knowledge and experience that the reference medicine is safe, pure and potent, an abbreviated clinical program is sufficient to establish that the biosimilar is highly similar to the reference medicine and will exhibit the same safety and efficacy in all approved indications. The extent of clinical studies required for the demonstration of biosimilarity is product specific, depending on the degree of molecular similarity and remaining residual uncertainty following analytical (physico-chemical and functional) analyses. This chapter aims to illustrate the scientific basis for the "Totality of Evidence" concept and to provide insight into the role of clinical trials for the verification of biosimilarity, using real world examples.

Keywords Biosimilarity · Bridging · Clinical · Competitive intelligence · Guidelines · Regulatory pathways · Non-clinical · Target Product Profile · Totality of Evidence · Strategy

H. P. Cohen (✉)
Sandoz Inc., Princeton, NJ, USA
e-mail: hillel.cohen@sandoz.com

W. C. Lamanna · M. Schiestl
Sandoz Biopharmaceuticals, Sandoz GmbH, Kundl, Austria
e-mail: william.lamanna@sandoz.com; martin.schiestl@sandoz.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_22

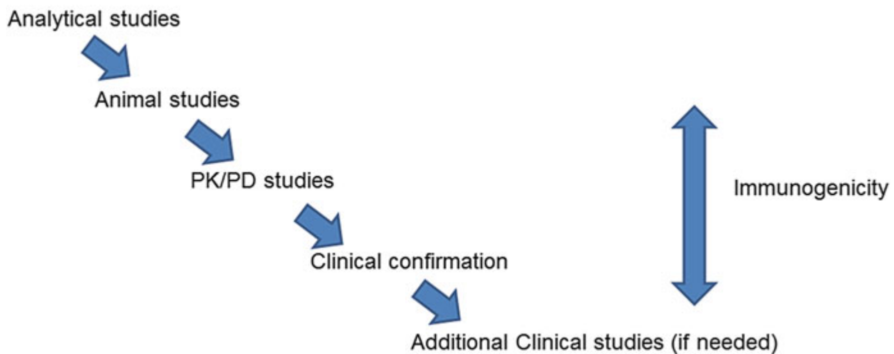


Fig. 22.1 Stepwise approach to address and reduce residual uncertainty. Adapted from (FDA 2015a), FDA Guidance for Industry: Scientific Considerations in Demonstrating Biosimilarity to a Reference Product (April 2015) (accessed March 1, 2018)

Totality of Evidence is the Regulatory Paradigm for Biosimilar Approval

Totality of evidence is a stepwise approach used to gather evidence that a proposed biosimilar will have the same quality, safety and efficacy as its respective reference medicine (FDA 2015a; Strand et al. 2017; Markus et al. 2017). At each of the steps, biosimilar developers seek to learn as much as possible about how closely the characteristics of the proposed biosimilar and the respective reference medicine match each other (Fig. 22.1). This information helps manufacturer assess how much residual uncertainty remains, which in turn serves to help the design the next set of head to head comparisons.

Focusing on novel biopharmaceuticals, they have historically been approved around the world based on a well-established paradigm whereby clinical trials are used to generate the data that establishes a favorable benefit:risk ratio for use of the drug to treat a given indication. Each indication is evaluated separately, with typically distinct and unique clinical studies. Drug development of biopharmaceuticals proceeds along a progressive sequence of steps, beginning with drug design and manufacturing process development, and then progressing through animals studies, and then Phase 1, Phase 2 and Phase 3 clinical studies (Fig. 22.2) (Lim and Christl 2017). The large-scale Phase 3 clinical trials are the most critical evidence to establish the efficacy and safety of the proposed drug in the intended population for the intended clinical indication.

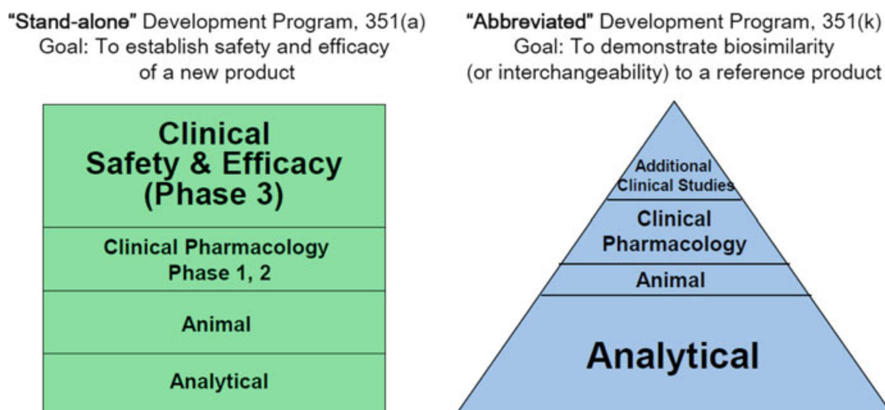


Fig. 22.2 Comparison of the standard (351(a)) regulatory pathway and the biosimilar (351(k)) regulatory pathway (Adapted from (Lim and Christl 2017))

In contrast, the goal in developing a biosimilar is not to reestablish the safety and efficacy of a biopharmaceutical de novo, but instead to establish molecular and functional similarity to a previously approved product. This goal is achieved using extensive analytical (physico-chemical and functional) comparisons to assess molecular similarity, followed by an abbreviated nonclinical and clinical program (as warranted) to confirm comparable safety and efficacy (Fig. 22.1). The basis for this alternative developmental and regulatory approach is the fact that a biosimilar is developed to match the structure, function and physiological impact of an approved biopharmaceutical to the degree that is possible. The underlying premise is that molecules that are structurally the same will elicit the same physiological responses in vivo. Analytical assays form the foundation of the totality of evidence approach due to their ability to sensitively and accurately detect molecular differences. These are followed by nonclinical and clinical assessments whose expanse and scope will depend on the extent of demonstrated molecular similarity and remaining residual uncertainty as to whether any detected differences are clinically meaningful (Fig. 22.1). The concept of Totality of Evidence" has been accepted by health authorities, and is incorporated into guidance documents issued by the Food and Drug Administration (FDA), the European Medicines Agency (EMA) and the World Health Organization (WHO) (FDA 2015a; European Medicines Agency 2014; World Health Organisation 2009).

The first and most foundational step in establishing the totality of evidence is the analytical and functional (*in vitro*) evidence (Section IV of this book, Chaps. 11, 12, 13, 14, 15, 16, and 17). Multiple orthogonal methods assess multiple critical quality attributes, and different batches of the biosimilar candidate are compared head-to-head with the reference product batches. The results are typically compared using visual and descriptive statistical methods (Chap. 17). FDA issued a draft guidance on the analytical statistics for biosimilars which includes equivalence testing for

the most critical quality attributes (FDA 2017a). However, the nature of analytical methodology to be used is still under discussion and the future role of inferential statistical approaches is not clear (Regulations.gov 2018). In the draft guidance as well as in public presentations (Lim and Christl 2017; FDA 2017a), the FDA has been very clear that the final conclusion of biosimilarity does not depend on passing or failing a statistical analysis of quality attributes, but on the totality of evidence including the impact of an enhanced control strategy in the manufacturing process.

Implicit in the concept of totality of evidence is the fact that there is no one pivotal assay or study that is solely decisive for the final conclusion on biosimilarity (Christl and Lim 2018). Given the biological nature of the manufacturing process, it is likely that some differences will be seen when comparing the reference medicine and a proposed biosimilar. It is important to assess whether or not the differences are detected in clinically relevant quality attributes and also the extent of any such differences (e.g. close to a predefined range or if the differences are large). Given the size and complexity of biological drugs, no single method is sufficient to establish biosimilarity. Instead, manufacturers employ a battery of orthogonal analytical techniques to establish the "sameness" of the biosimilar and reference medicine (Lamanna et al. 2018).

The analytical and functional testing are commonly followed by non-clinical (animal) studies in relevant models and then by assessment of the human pharmacological properties in one or more clinical pharmacokinetic (PK) and if appropriate, pharmacodynamic (PD) studies (FDA 2015a; European Medicines Agency 2014; World Health Organisation 2009). For final confirmation that the same safety and efficacy can be expected from a biosimilar, a confirmatory study is commonly conducted, although, FDA and EMA biosimilar guidelines describe conditions under which such safety and efficacy studies can be waived (FDA 2015a; European Medicines Agency 2014). Comparative human immunogenicity data must also be collected. Once a biosimilar manufacturer compiles data from all sources and submits a license application, health authorities "will evaluate the applicant's integration of various types of information to provide an overall assessment whether a biological product is biosimilar" to the reference medicine (Christl 2018).

A key term used by many health authorities is demonstration of "high similarity" of the biosimilar and the reference medicine. The term "high similarity" has been the source of substantial confusion with regard to biosimilars with some suggesting that the term "similar" implies that the product is not the same and may even exhibit sub-standard quality. In fact, the term "similar" is used due to the inherent variability of biopharmaceuticals (Christl and Lim 2018). Indeed, due to their cell based production, biopharmaceuticals and their respective biosimilars contain micro-heterogeneous variants, which will always vary to some degree from one batch to the next. Thus, a biosimilar cannot, by definition, be identical to the reference medicine the same way a batch of the reference medicine cannot be identical with the next. Instead, a biosimilar must be "highly similar" to the reference medicine, notwithstanding minor differences in clinically inactive components (FDA 2015a; European Medicines Agency 2014).

It is important to emphasize that although the clinical development program for a biosimilar is smaller in scope than the reference medicine, the biosimilar is held to the same standards of safety, purity and potency as are applied to any other biopharmaceutical. There is only a single set of product quality and manufacturing quality standards, and those are applied to all biological drugs, irrespective of whether it is an originator licensed as a new molecular entity (and suitable for use as a reference medicine) or a biosimilar (European Medicines Agency 2014). Overall, the Totality of Evidence must demonstrate that there is no clinically meaningful difference between the biosimilar and its reference medicine in terms of safety, efficacy and quality (FDA 2015a; European Medicines Agency 2014).

After having read the chapters in previous sections of this book and understood the role of comparative analytical studies and an abbreviated but targeted clinical program conducted to support the approval of biosimilars, we now will take a deeper dive into the nature of the data used to establish the "totality of evidence" that supported the decisions of health authorities to approve specific biosimilars.

Totality of Evidence is Product-Specific

Since biosimilar regulation typically applies to protein based drugs, it follows that the analysis of these drugs will include methods designed to evaluate properties common to all proteins, such as sequence, higher order structure, mass, purity and charge. The nature of the molecule and manufacturing process will also help guide selection of additional physico-chemical assays that are specific to fully characterize a particular molecule, such as assays to detect and quantify glycosylation, product related variants and process impurities. The nature of the bioassays will depend on the functionalities the molecule can elicit.

On occasion, differences have been detected when comparing reference product and a proposed biosimilar. In developing Inflectra® (infliximab-dyyb, also known as Remsima® in the EU) analyses revealed a different glycosylation pattern that impacted the antibody dependent cell cytotoxicity (ADCC) properties of the molecule (Pisupati et al. 2017). But while ADCC is not thought to be linked to the mechanism of action (MoA) of this molecule, the disparity in glycosylation and resulting ADCC led the manufacturer to conduct a more extensive series of in vitro functional assays (see Sect. "Analytical Data" of this chapter). This example illustrates that if differences are observed that are thought to be not clinically relevant, there is still an element of residual uncertainty that must be addressed by the manufacturer to establish the Totality of Evidence necessary to support approval of a proposed biosimilar.

Some details of the methods and studies conducted to support drug approval can be obtained from health authority websites. The identities of the analytical test methods and the general nature of nonclinical and clinical studies conducted to support a product approval are provided for most products on the FDA website when

the FDA's review summaries are posted, often with a top-line presentation of the results. Another source of information available on the FDA website are the briefing documents provided by the FDA to support U.S. Advisory Committee reviews of products; however not all products are reviewed by Advisory Committees. Using these sources of information, we were able to gather information on the evidence provided to support the totality of evidence for six biosimilars that were approved in the U.S. as of the end of 2017, which are compared in Tables 22.1, 22.2, 22.3, and 22.4 and are discussed in the sections below (FDA 2018a, b, c, d, e).

Table 22.1 US biosimilar approvals as of December 31, 2017

Date of biosimilar FDA approval	Biosimilar product [manufacturer]	Reference medicine [manufacturer]
March 6, 2015	Zarxio® (filgrastim-sndz) [Sandoz]	Neupogen® (filgrastim) [Amgen]
April 5, 2016	Inflectra® (infliximab-dyyb) [Celltrion]	Remicade® (infliximab) [Janssen]
August 30, 2016	Erelzi® (etanercept-szzs) [Sandoz]	Enbrel® (etanercept) [Amgen]
September 23, 2016	Amjevita® (adalimumab-atto) [Amgen]	Humira® (adalimumab) [AbbVie]
April 21, 2017	Renflexis® (infliximab-abda) [Samsung Bioepis]	Remicade® (infliximab) [Janssen]
August 25. 2017	Cyltezo™ (adalimumab-abdm) [Boehringer Ingelheim]	Humira® (adalimumab) [AbbVie]
September 14, 2017	Mvasi™ (bevacizumab-awwb) [Amgen]	Avastin® (bevacizumab) [Genentech]
December 1, 2017	Ogivri™ (trastuzumab-dkst) [Mylan]	Herceptin® (trastuzumab) [Genentech]
December 13, 2017	Ixifi™ (Infliximab-qbtx) [Pfizer]	Remicade® (infliximab) [Janssen]

Table 22.2 Molecular characteristics of biosimilars approved in US

Molecule	Molecular weight	# amino acids	# chains	# disulfide bonds	Isoelectric point	Glycosy-lated?	Elimination half-life
Adalimumab	~148,000	1330	4	16	8.25	Yes	4–7.8 days
Bevacizumab	~149,000	1330	4	16	8.3	Yes	~20 days
Etanercept	~150,000	934	2	29	7.89	Yes	70–132 h
Filgrastim	18,803	175	1	2	5.65	No	3–4 h
Infliximab	140,190	1328	4	16	8.25	Yes	7.7–9.5 days
Trastuzumab	~145,500	1328	4	16	8.45	Yes	1.7–28 days

Table 22.3 Quality attributes and methods used in evaluation of Zarxio (filgrastim-sndz) and Erelzi (etanercept-szzs)

Quality attribute	Methods used for Zarxio® (filgrastim-sndz)	Methods used for Erelzi® (etanercept-szzs)
Primary structure • Amino acid sequence disulfide mapping • Site of chemical modification or glycosylation • Free thiols	• N-terminal sequencing • Peptide mapping with UV and MS detection • Protein molecule mass by ESI MS • Protein molecule mass MALDI-TOF MS • DNA sequencing of construct cassette • Peptide mapping coupled with MS/MS	• LysC peptide mapping coupled with RP-UPLC with fluorescence and MS detection • Tandem MS/MS
Bioactivity	• Proliferation if murine myelogenous leukemia cells (NFS-60 cell line)	• TNF-alpha neutralization, TGF-beta neutralization and TNF-alpha binding as measured by reporter gene assay • FcRn binding as measured by SPR • CDC activity, ADCC, FCγRIIIa binding • FCγRIIa binding • Other FCγ receptors, as measured by cell-based SPR
Receptor binding	• SPR	Method used is not listed in FDA documents
Protein content	• RP-HPLC	Method used is not listed in FDA documents
Clarity	• Nephelometry	Method used is not listed in FDA documents
Sub-visible particles	• Micro flow imaging	Method used is not listed in FDA documents
Higher order structure	• Far and near UV circular dichroism • ^{1}H nuclear magnetic resonance • ^{1}H-^{15}N heteronuclear single quantum coherence spectroscopy • LC-MS to detect disulfide bonds	• Far and near UV circular dichroism • Differential scanning calorimetry • FTIR • Hydrogen deuterium exchange mass spectrometry • NMR • X-ray crystalography • Non-reducing peptide mapping to detect wrongly bridged disulfide bonds

(continued)

Table 22.3 (continued)

Quality attribute	Methods used for Zarxio® (filgrastim-sndz)	Methods used for Erelzi® (etanercept-szzs)
Higher molecular weight variants/aggregates	• Size exclusion chromatography (validated using MALLS) • Reduced and non-reduced PAGE • 90° light scattering • Nile red steady-state fluorescence anisotrophy	• SEC • SEC-MALLS • SDS-PAGE • AUC • FFF-MALLS
Product related substances and impurities • Oxidized variants • Acidic variants • Acetylated variants • Covalent dimers • Partially reduced variants • Sequence variants • Formyl-Met 1 • Succinimide • Phosphoglucunoylation	• RP-HPLC • LC/MS • LC-MS/MS	• Capillary zone electrophoresis (CZE) • Peptide mapping
Product related substances and impurities • N-terminal truncated variants	• LC-MS/MS	• CZE • Peptide mapping
Product related substances and impurities • Norleucine • Deamidation	• RP-HPLC • LC/MS • IEX • CZE	• CZE • Peptide mappring

Glycosylation • Terminal GlcNAc-variants	Not product relevant	• NP-HPLC
Glycosylation • Triatennary glycan strucutres • Non-fucosylated glycan variants • Overall sialylation • Sialylation N-glycans • Sialylation O-glycans • Sialic acids (NGNA)	Not product relevant	• DMB-labeling (NANA, NGNA) • AEX (relative retention time) • WAX
Glycosylation • Beta-galactosylation	Not relevant	• DMB-labeling (NANA, NGNA) • AEX (relative retention time) • WAX

ADCC antibody dependent cell cytotoxiicty, *AEX* anion exchange chromatography, *AUC* analytical ultracentrifugation, *CDC* complement dependent cytotoxicity, *CZE* capillary zone electrophoresis, *DMB* 1,2-diamono-4,5-methylenedioxibezene, *ESI* electrospray ionization, *GlcNAc* N-acetylglucosamine, *FDA* Food and Drug Adminsitration, *FFF* field flow fractionation, *FTIR* Fourier transform infrared spectroscopy, *IEX* ion exchange chromatography, *LC/MS* liquid chromatography with mass spectrometry detection, *LC-MS/MS*, liquid chromatography with tandem mass spectrometry detection, *LysC* endoproteinase LysC, *MALDI-TOF* matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *MALLS* multi-angle laser light scattering, *MS* mass spectrometry, *NANA* N-acetylneuraminic acid, *NGNA* N-glycoylneuraminic acids, *NMR* nuclear magnetic resonance, *NP-HPLC* norrmal phase high performance liquid chromatography, *PAGE* polyacrylamide gel electrophoresis, *RP-HPLC* reverse phase high performance liquid chromatography, *SEC* size exclusion chromatography, *SPR* surface plasmon resonance, *TGF* transforming growth factor, *TNF* tumor necrosis factor, *UV* ultraviolet, *WAX* weak anion exchange chromatography

Table 22.4 Quality attributes and methods used in evaluation of four biosimilar monoclonal antibodies approved in the U.S

Quality Attribute	Methods used for analysis of Inflectra® (infliximab-dyyb)	Methods used for analysis of Amjevita® (adalimumab-atto)	Methods used for analysis of Mvasi® (bevacizumab-awwb)	Methods used for analysis of Ogivri® (trastuzumab-dkst)
Primary structure • Amino acid sequence • Disulfide mapping • Site of chemical modification or glycosylation • Free thiols	• Peptide mapping with UV and MS detection • Amino acid analysis • MS/MS • Intact mass reduced (LC-MS) • Peptide mapping coupled MS/MS	• Peptide mapping with UV and MS detection (reduced and non-reduced) • Amino acid analysis • Intact molecular mass (LC-MS) • Reduced and deglycosylated molecular mass (LC-MS)	• ESI-TOF-MS • Reduced and deglycosylated Molecular Mass (LC-MS) • Amino acid analysis by reduced peptide mapping with UC and LC-MS-MS • Disulfide structure using non-reduced and reduced peptide mapping LC-MS	• Peptide mass fingerprinting by LC-ESI-MS detection • Intact molecular mass (LC-ESI-MS) • Reduced molecular mass (LC-ESI-MS)
Protein content	• UV280	• UV280	• UV280	• UV280
Higher order structure	• Far and near UV CD • FTIR • Free thiols • Antibody array • LC-MS to characterize disulfide bonds • DSC	• Near UV CD • FTIR • LC-MS • LC-MS (disulfide bond characterization) • DSC	• Near UV CD • FTIR • DSC	• Near and far UV CD • FTIR • IF • LC-ESI-MS (disulfide bond characterization) • DSC • Ellman's reagent (free cysteines)

Molecular weight variants/aggregates	• SEC-HPLC • SEC-MALS • CE-SDS (reduced and non-reduced) • AUC	• SEC (UV Detection) • SEC-LSD • FFF • AUC • CE-SDS (Reduced and Non-Reduced)	• SEC (UV Detection) • SEC with Light Scattering • FFF • Dynamic light scattering • AUC • CE-SDS (Reduced and Non-Reduced)	• Size exclusion high performance liquid chromatography (SEC-HPLC) with UV detection • SEC with MALS • AUC • CE-SDS (Reduced and Non-Reduced)
Charge	• IEF • IEF-HPLC	• Capillary IEF • CEX-HPLC	• Capillary IEF • CEX-HPLC	• Capillary isoelectric focusing (cIEF) • CEX-HPLC • HIC-HPLC • LC-ESI-MS (oxidation)
Glycosylation	• Oligosaccharide profiling • N-linked glycan analysis • Sialic acid analysis • Monosaccharide analysis	• Afucosyaltion • Galactosylation • High mannose • Sialic acid	• "Glycan Mapping"	• Normal phase (NP)-HPLC (glycan mapping) • RP-HPLC (afucosylation and sialic acid) • CE-SDS reduced (Non-glycosylated heavy chain) • Boronate affinity chromatography (glycation)
Potency	• In vitro TNF-α neutralization assay	• Apoptosis inhibition bioassay	• Inhibition of human umbilical vein endothelia cell proliferation bioassay	• HER2 binding assay by flow cytometry • Inhibition of proliferation bioassay • ADCC bioassay (PBMC as effector cells)
Binding assay	• ELISA • Cell based binding affinity	• ELISA	• ELISA • SPR	Not listed

(continued)

Table 22.4 (continued)

Quality Attribute	Methods used for analysis of Inflectra® (infliximab-dyyb)	Methods used for analysis of Amjevita® (adalimumab-atto)	Methods used for Analysis of Mvasi® (bevacizumab-awwb)	Methods used for Analysis of Ogivri® (trastuzumab-dkst)
Binding assay—Fc	• NK cell binding affinity via Fc receptors (in presence of 50% serum or 1% BSA) • FcγRIIIa V and F typebinding affinity (SPR) • FcγRIIIb binding affinity (SPR) • FcγRIIa binding affinity (SPR) • FcγRIIb binding affinity (SPR) • FcγRI binding affinity (ELISA) • FcRn binding affinity (SPR)	• FcγRIIIa V type binding affinity • FcγRIIIa F type binding affinity • FcγRIIa binding affinity • FcγRIa binding affinity • FcRn binding affinity (cell-based)	• FcRn binding affinity • FcγRIa binding affinity • FcγRIIa (SPR) • FcγRIIb (SPR) • FcγRIIIa V type binding affinity • FcγRIIIa F type binding affinity • FcγRIIIb (SPR)	• FcγRIIIa V type binding affinity (SPR) • FcγRIIIb binding affinity (SPR) • FcγRIIa binding affinity (SPR) • FcγRIIb binding affinity (SPR) • FcγRIa binding affinity (SPR) • FcRn binding affinity (SPR) • CDC bioassay
Binding assay—complement	• C1q binding assay (ELISA)	• C1q binding assay (ELISA)	• C1q binding assay (ELISA)	• C1q binding assay (ELISA)
Bioassay/mechanism of action exploration	• ADCC (PBMC as effectors) • ADCC (NK cells as effectors) • ADCC (LPS-stimulated monocytes as targets) • CDC • Induction of apoptosis by reverse signaling • Inhibition of pro-inflammatory cytokine release by reverse signaling (Caco-2 cells)	• ADCC (NK cells as effectors) • CDC • Inhibition of sTNFα-induced IL-8 in HUVEC • Specificity against LTα in HUVEC assay • Inhibition of sTNFα-induced cell death in L929 cells	Not listed	Not listed

	• Wound healing (closure %) • Inhibition of T Cell proliferation • Induction of regulatory macrophages	• Inhibition of sTNFα-induced chemokines in whole blood • Inhibition of T-Cell proliferation • Induction of regulatory macrophages		
Clarity/sub-visible particles	• Nephelometry	• Micro fluid imaging • Light obscuration	• Micro fluid imaging • Light obscuration	Not listed
General properties	Not listed	• Deliverable volume • Osmolality • pH • Appearance • Polysorbate	• Deliverable volume • Osmolality • pH • Appearance • Color	Not listed
Product related substances and impurities	Not listed	Not listed	Not listed	Not lsited
Thermal degradation	Not listed	Not listed	• SEC-HPLC • CE-SDS • CEX-HPLC • Potency	Not listed

ADCC antibody dependent cell cytotoxiicty, *AUC* analytical ultracentrifugation, *CD* circular dichroism, *CDC* complement dependent cytotoxicity, *CE* capillary electrophoresis, *CEX* cation exchange, *cIEF* capillary isoelectric focusing, *DSC* differential scanning calorimetry, *ELISA* enzyme-linked immunosorbent assay, *ESI* electrospray ionization, *ESI-TOF* electrospray time-of-flight, *FFF* field flow fractionation, *FTIR* Fourier transform infra-red spectroscopy, *HIC* hydrophobic interaction chromatography, *HPLC* high performance liquid chromatography, *HUVEC* human umbilical vein endothelial cells, *IEF* isoelectric focusing, *IF* intrinsic fluorescence, *LPS* lipopolysaccharide, *MALS* multi-angle light scattering, *MS* mass spectrometry, *MS/MS* tandem mass spectrometry, *NK* natural killer, *PBMC* peripheral blood mononuclear cells, *RP-HPLC* reverse phase high performance liquid chromatography, *SDS* sodiumdodecylsulfate, *SEC* size exclusion chromatography, *SEC-LSD* size exclusion chromatography with light scattering detection, *SPR* surface plasmon resonance, *TNF* tumor necrosis factor, *UV* ultraviolet, *UV280* ultraviolet detection at 280 nm

Analytical Data

The very first biosimilar to be approved in any jurisdiction was Omnitrope® (somatropin), approved in the European Union (EU) by the European Medicines Agency (EMA) in 2006. In the decade since, the number of approved biosimilars in the EU has expanded dramatically, with 41 biosimilar products approved to date to 13 unique reference medicines (Generics and Biosimilars Initiative (GABI) 2018). In the US, the biosimilar pathway is still comparatively new, with a formal pathway first established in 2010 and the first biosimilar, Zarxio® (filgrastim-sndz), approved in 2015. As of the end 2017, there are nine biosimilars approved in the US towards six unique reference medicines (Table 22.2). These approved biosimilars include molecules with a broad array of molecular characteristics including one non-glycosylated, single chain protein (filgrastim), one complex fusion protein and 5 monoclonal antibodies.

Compared to the other biosimilars, Zarxio® is a relatively small and simple protein composed of a single amino acid chain, containing only two disulfide bonds and devoid of glycosylation. The demonstration of biosimilarity for this product utilized approximately 22 methods for the evaluation of 19 physicochemical quality attributes (Table 22.3). At the other end of the size spectrum, the etanercept biosimilar Erelzi® (etanercept-szzs) represents a highly complex protein. This 150 kDa fusion protein is composed of two amino acid chains, contains 29 disulfide bonds, six N-glycans and can be further modified with up to 14 O-glycans. This increased molecular complexity is reflected in the analytical program used for demonstrating similarity, involving more than 50 analytical methods for over 100 physicochemical quality attributes, a selection of which are shown in Table 22.3.

Many of the methods used in the physico-chemical characterization of Zarxio® were also used in product-specific adaptations for characterization of Erelzi®. Additional techniques were used to characterize Erelzi® which reflects the need to evaluate glycosylation, dilsulfide binding features, and a more complex three-dimensional structure. Use of state of the art analytical tools as were applied to Zarxio® and Erelzi® is a regulatory expectation to ensure thorough product understanding and improve manufacturing control. One of the most important advancements in the analytical characterization of protein therapeutics in recent years has come from mass spectrometry, allowing highly sensitive characterization of variants and impurities in the femtomolar and attomolar range (Beck et al. 2012, 2015). Additionally, advancements in methods for the assessment of high order structure of large proteins such as hydrogen—deuterium exchange, mass spectrometry or new applications of 2D-NMR, have improved the ability to assess three dimensional structure similarity (Strand et al. 2017).

The state of the art analytical characterization required for biosimilar development and approval can provide new physicochemical insights into the structure and function of biopharmaceuticals as well as increased knowledge of the reference medicine. An example of this was recently published for the biopharmaceutical etanercept, where the development of a novel mass spectrometry-based method

allowed characterization of incorrect disulfide bridge structures in the reference medicine Enbrel® whose abundance directly correlate with reduced in vitro potency (Lamanna et al. 2017). The research also demonstrated that incorrect disulfide bridge structures in the etanercept molecule are corrected under physiological conditions, restoring normal potency. These findings were important for understanding the clinical relevance of this impurity and provided new insight into the structure—function relationships in etanercept which can be applied to produce consistently safe and potent biosimilar medicines.

The concept of totality of evidence permits minor differences in structure as long as they are not clinically meaningful. Such differences can include post-translational modifications that do not influence safety or efficacy, such as small differences in the levels of C-terminal lysine on monoclonal antibodies. Indeed, the abundance of C-terminal lysine has been shown to be functionally inert, most likely due to the fact that it is quickly removed by enzymes in the blood stream once a drug is administered to patients. Differences in levels of minor variants may also be acceptable as long as these are present at levels clearly below biologically relevant thresholds. This was recently illustrated by the comparative characterization of the filgrastim reference medicine Neupogen® and commercially available biosimilars using mass spectrometry, revealing differences in the abundance of minor variants such as N-terminal acetylation and rare sequence variants at levels below 1% relative abundance (Hausberger et al. 2016). Notably, biosimilars can be developed to contain lower levels of potentially harmful variants, such as aggregates or immunogenic glycans, to help ensure patient safety (FDA 2015a; European Medicines Agency 2014).

Since seven of the first nine biosimilars to be approved in the US are monoclonal antibodies, it is interesting to compare the methods used by different companies to characterize these products. As of the end of 2017 the FDA has only posted the summary reviews for two of these biosimilars Inflectra® (infliximab-dyyb) (FDA 2018b) and Amjevita® (adalimumab-atto) (FDA 2018d) but the briefing documents for the U.S. Oncology Advisory Committee reviews of Mvasi® (bevacizumab-awwb) and Ogivri® (trastuzumab-dkst) (FDA 2018e), both held on July 13, 2017, also provide summaries of the analytical programs used to support approval of those products. But even given these limitations, it is possible to compare the data packages that supported approval of the four different biosimilar monoclonal antibodies (Table 22.4). It is remarkable how similar the analytical methodology is, although there are some product-to-product differences in methodology. The similarities likely reflect the scientific consensus on the methods most suited to analyze these products, with the few differences (e.g. use of fast-flow fractionation (FFF), far ultra-violet (UV) circular dichroism) likely reflecting the preferences of scientists in the different companies.

In examining the publically available analytical data packages provided to support different biosimilars approved in the US, we observe several overarching themes that are discussed below. In general, they reflect a recent FDA draft guidance on chemistry, manufacturing, and controls changes (FDA 2017b):

- There is an expectation that the most modern techniques will be utilized to compare a proposed biosimilar and the corresponding reference medicine. It is therefore common to utilize methods that were not available at the time that the reference medicine was first approved. As an example, the sensitivity of mass spectroscopy of proteins has increased 10 million-fold in the past two decades (Mire-Sluis 2012).
- Orthogonal methods are common and are expected, consistent with the FDA's CMC guidance for biosimilar development and characterization (FDA 2015b).
- Many of the same methods are employed by different companies, although the FDA permits use of different methods if preferred by a company and suitable for the purpose.
- Functional assays will reflect the nature of the molecule and its functions in biological systems. For molecules with multiple functions, it is a health authority expectation that the bioassays will assess all those functionalities (FDA 2015a; European Medicines Agency 2014).

Non-clinical Animal Data

Non-clinical animal testing can be a useful in development of therapeutic proteins. For new molecular entities, extensive nonclinical testing establishes that the proposed drug is sufficiently safe (within the limits of the animal models) for further exploration in humans. Invasive testing and necropsies are possible in animal models to give a clear picture of the impact of the drug on different organ systems. Non-clinical animal testing is also utilized to help select a starting dose for human studies. Finally, PK and if possible, PD studies testing of new molecular entities in relevant animal models provides a baseline of what might be expected when the drug is administered to humans. For biosimilar medicines, non-clinical animal models can be valuable for evaluating similarity when residual uncertainty remains following analytical and functional evaluation as to whether any differences in variants or formulation are clinically meaningful.

There is a disparity in the expectations and requirements for non-clinical comparative animal studies of biosimilars and their reference medicines. Whereas non-clinical animal studies are necessary for the development of novel new drugs, the EU and WHO require only limited or no non-clinical animal data to support registration of biosimilars (EMA 2005). This reflects a position that in vitro analyses, as discussed in section "Analytical Data" of this chapter, are more sensitive and specific than animal testing. Similarly, the FDA does have the ability to determine that a specific type of testing is not needed for a given biosimilar, but in general the FDA recommends that

> if the structural and functional data are limited in scope or there are concerns about the proposed product quality, a general toxicology study may be needed that includes full animal pathology, histopathology, PD, PK, and immunogenicity assessments (FDA 2015a).

Until recently the WHO guidance for development of biosimilars specified the need to conduct animal toxicology studies when developing biosimilars, calling for a head-to-head repeat dose toxicity study (World Health Organisation 2009). However, the recently released draft WHO Questions and Answers for Similar Biotherapeutic Products instead emphasizes the utility of in vitro non-clinical studies, noting that they are often more specific or sensitive for detecting differences than studies in animals (WHO 2017). The draft guidance explains that

> in some jurisdictions, legislation requires the application of the 3R (Reduction, Refinement and Replacement of animal experiments) principle in the product development in order to reduce suffering of animals. In particular, studies with non-human primates should be avoided if possible. In vivo animal studies should only be considered when it is expected that such studies would provide relevant additional information.

The WHO draft guideline identifies five factors that reduce the need for in vivo animal studies of biosimilars (WHO 2017):

1. "The risk of the first-in-man use of a biosimilar can usually be estimated on basis of the knowledge about the clinical safety profile of the reference product and the outcome of the physico-chemical, structural and in vitro functional tests with the biosimilar.
2. Most toxic effects of a therapeutic proteins are often related to an exaggeration of their known pharmacological effects.
3. The functional activity of a biotherapeutic drug substance is often species-specific making it difficult to identify a suitable animal species.
4. Being foreign, human drug substances are often immunogenic in the conventional animal models which prevents or hampers the interpretation of repeat dose animal studies.
5. Conventional animal models are often not sensitive enough to detect small differences."

Chapman et al. (2016) have suggested a paradigm to eliminate or minimize non-clinical animal testing of biosimilar monoclonal antibodies, focusing on addressing specific scientific questions as opposed to conducting non-clinical animal studies as a "check-box exercise." The suggestion is especially relevant to monoclonal antibodies because the value of non-clinical animal testing to predict safety in humans has been questioned (Martin and Bugelski 2012; van Meer et al. 2013). Further, the concept of only conducting testing when it can answer a scientific question is consistent with the concept of Totality of Evidence and as such is valid for all biosimilars. The added value of nonclinical animal studies for assessing biosimilarity should be product specific, depending on the degree of molecular or functional complexity, the extent of physicochemical similarity and the availability of a relevant and sufficiently sensitive animal model. The latter point is especially important and has led to a great reduction of animal testing in pharmaceutical development overall (National Institute of Environmental Health Science 2017).

Full summaries of nonclinical animal data are publically available for only four FDA licensed biosimilar drugs as of the end of December 2017, all from FDA review

Table 22.5 Nonclinical studies conducted to support US approval of four biosimilar drugs

Product	Non-clinical studies conducted
Zarxio® (filgrastim-sndz)	1. 12-day PD study in rats 2. 28-day repeat dose toxicity study in rats, EU buffer 3. 28-day repeat dose toxicity study in rats, US buffer 4. Local tolerance in rabbits 5. 14-day toxicokinetic study in rats
Erelzi® (etanercept-szzs)	1. Comparative efficacy study in transgenic mice 2. Single dose, subcutaneous PK studies in rabbits, evaluating different formulations 3. Single dose, subcutaneous PK studies in rabbits, final formulation 4. 28-day, repeat-dose toxicology study in Cynomolgus monkeys
Inflectra® (infliximab-dyyb)	1. Comparative TNF-α binding affinity in different species 2. Cross-reactivity with a panel of different human tissues 3. Single-dose toxicity study in rats 4. Two-week, repeat-dose toxicity in mice to evaluate tumorigenicity
Amjevita® (adalimumab-atto)	1. Toxicokinetic study in Cynomolgus monkeys 2. Four-week toxicology study in Cynomolgus monkeys

EU European Union, *PD* pharmacodynamics, *PK* pharmacokinetics, *TNF* tumor necrosis factor, *US* United Sates of America

summaries that are posted on the FDA website (Table 22.5). An examination of this information reveals that all studies were conducted to evaluate safety-related parameters, and were not conducted to address residual uncertainty regarding a specific issue or a question that may have arisen after the comparative analytical testing program.

Clinical Studies

Unlike development of new biopharmaceutical drugs where clinical studies provide the key safety and efficacy information supporting a product approval, the foundation of the totality of evidence for a biosimilar is the analytical and functional data which are used to establish molecular similarity. As noted above, the goal of the clinical program is not to reestablish efficacy and safety de novo, but to confirm similarity in a sensitive in vivo setting, providing additional confidence that a biosimilar will have the same safety and clinical performance as that of the reference medicine. Requirements include demonstrating similar PK and, if possible, PD as well as comparable immunogenicity profiles, which at present requires confirmation in humans.

Human PK studies are used as the first clinical tool to compare a proposed biosimilar and reference medicine. Human PD studies are also conducted if suitable

and meaningful markers are available. These PK/PD studies are conducted ideally in healthy individuals because they do not have confounding factors such as variation in the underlying disease, or concomitant medications that could impact PK/PD responses. However, there may be situations where it is necessary to conduct PK/PD studies in patients due to concerns regarding safety.

Immunogenicity of therapeutic proteins varies widely, with some proteins eliciting clinically relevant antibodies in a large proportion of recipients (e.g. infliximab and adalimumab) while other therapeutic medicines are much less immunogenic (e.g. filgrastim and etanercept). Highly similar structure, (e.g. identical amino acid sequence) helps ensure that the immunogenic epitopes are the same. However, there are also risk factors, such as aggregates, non-human glycans or process impurities which can increase immunogenicity (Jahn and Schneider 2009; Singh 2011). Biosimilar products must demonstrate comparable immunogenicity with the reference medicine, but it is acceptable to develop biosimilars with lower levels of variants or impurities as long as potency is not impacted relative to the reference medicine. To allow adequate control of immunogenicity in a biopharmaceutical it is important to compare the risk factors related to both the proposed biosimilar and the reference product and to control for them in the manufacturing process.

Health authorities typically require additional human clinical data to compare the immunogenicity profile of a proposed biosimilar to that of the reference medicine. However, human clinical immunogenicity data is not always a requirement as is illustrated by follow on versions of enoxaparin, a complex glycosaminoglycan, which has the potential to elicit a rare but potentially fatal adverse events caused by anti-drug immunogenicity (Martel et al. 2005). Enoxaparin is considered as a biosimilar in Europe but in the U.S. as a complex generic. The FDA approved generic enoxaparin without a requirement for human immunogenicity data based on a very detailed impurity assessment provided by the manufacturer to ensure comparable immunogenicity (Center for Drug Evaluation and Research 2010).

Valuable immunogenicity data from immunocompetent individuals can be already collected during the PK/PD studies, however, immunogenicity data is also collected from the confirmatory safety and efficacy studies to confirm that the immunogenic profile is the same even in the face of concomitant medications or a disease state. The analysis of immunogenicity is commonly evaluated in two-stage process whereby clinical study samples are initially screened for presence of anti-drug-antibodies, and then positive samples are evaluated to assess whether these anti-drug-antibodies are neutralizing, meaning whether they directly prevent the function of the drug (e.g. by clearing the drug from the body compartments or by direct inhibition of the drug-target interaction).

Confirmatory Safety and Efficacy Studies

Confirmatory phase III-type safety and efficacy clinical trials in a single sensitive indication have historically been an integral part of the totality of evidence for evaluating biosimilarity. The goal of confirmatory safety and efficacy clinical trials

for a biosimilar is not to re-establish safety and efficacy but to verify similar safety and efficacy between the biosimilar and its respective reference medicine. This approach is scientifically sound as safety and efficacy for a given biopharmaceutical has already been established by the reference medicine in each indication for each target patient population. A biosimilar leverages this knowledge by demonstrating molecular sameness and subsequently verifying that the biosimilar elicits the same physiological response as the respective reference medicine when used to treat patients in a sensitive indication.

To stringently assess similar clinical performance between a biosimilar and its reference medicine, the chosen indication and endpoints for a comparative trial will be based on the ability to effectively detect potential differences in safety or efficacy. Thus, the indication selected for assessing clinical similarity may not be the most common indication and the chosen clinical endpoints may differ from those previously evaluated by the reference medicine. Further, while originator medicines subsequently used as reference medicines must perform broad clinical trials that enroll representative patient diversity, comparative clinical trials for biosimilars often utilize a more homogenous patient population to lessen patient-to-patient variability. For these reasons, despite the fact that the confirmatory biosimilar studies must be appropriately powered, their more focused approach and tailored design allows these trials to be comparatively smaller than today's phase III studies that would be conducted for originator biopharmaceuticals that are licensed as new molecular entities.

The FDA has the authority to forego phase III confirmatory clinical testing, if such studies are deemed unnecessary (FDA 2015a). Likewise, the EMA provides guidance that "in specific circumstances, a confirmatory clinical trial may not be necessary (European Medicines Agency 2014)." Nonetheless, in the decade since the advent of biosimilars, most biosimilars have been supported by confirmatory phase III clinical studies. An exception to this approach was accepted for the approval of Zarxio® in the EU (where it is known as "Zarzio®), in which case the EMA took the position that PK, PD and open label immunogenicity studies were sufficient to confirm clinical similarity of Zarxio® and its reference medicine (Martel et al. 2005). In contrast, the FDA required the same manufacturer to conduct a separate comparative safety and efficacy study in patients in order to obtain approval in the U.S. (FDA 2018a). While the FDA may be amenable to this approach for other biosimilar products in the future, there is no doubt that such an approach would need to be carefully justified on a product by product basis.

Non-US Sources of Information Supporting the "Totality of Evidence" for Approved Biosimilars

The concept of "Totality of Evidence" is applied by all health authorities worldwide, including the US, EU, WHO, Japan, Canada and Australia and elsewhere. However,

dossiers and health authority reviews are considered proprietary. Individual companies have at time published results of comparative analytical or clinical studies in peer-reviewed journals. While this can provide useful information, it does not provide a complete picture of the full data package supporting a proposed biosimilar, nor does it undergo the level of scrutiny that would be applied by a stringent health authority.

After a product approval, the European Medicines Agency (EMA) posts a European Product Assessment Report (EPAR) on their website. The EPAR contains a summary of the evaluation of human clinical studies, including human PK/PD and safety and efficacy studies. Some level of detail is often also provided for animal studies. But detailed results of analytical and functional assays are not posted on the website.

A review of the EMA website (EMA 2018) shows that with few exceptions, the clinical studies used to support biosimilar approvals in the US also were used to support approval of the same products in the EU. One noteworthy exception was that in place of a confirmatory safety and efficacy study for their filgrastim biosimilar, Sandoz submitted an open label human immunogenicity study in the EU (EMA 2008).

Extrapolation

Extrapolation is the approval of a product for indications for which it was not directly tested and is used to reduce unnecessary clinical testing. The utilization of this concept is not specific to biosimilars but has long been employed for originator biopharmaceuticals licensed as new molecular entities to assess product comparability following major manufacturing changes in the limited number of circumstances when clinical trials were required (Weise et al. 2014; Krendyukov and Schiestl 2018). As this topic is discussed in detail in Chap. 22 of this book, we will not discuss extrapolation in detail. However it is important to note that the concept of extrapolation is also based on the Totality of Evidence. Specifically, extrapolation to additional indications is determined based on a molecule to molecule comparison that calls upon the totality of evidence for a specific biosimilar.

To date, all biosimilar clinical development programs have relied on extrapolation to obtain approval of all available indications of the reference medicine. Four of the six clinical development programs evaluated relied on a single confirmatory efficacy and safety study with the remaining two having conducted confirmatory safety and efficacy studies in two distinct indications (Table 22.6). The assessment of more than one indication may be justified should one indication be most sensitive to assess efficacy while another may be more sensitive to assess safety or immunogenicity. However, during the U.S. Advisory Committee review of Amjevita®, the FDA pointedly commented that conducting multiple confirmatory efficacy and safety studies when only a single study would suffice is contrary to the concept of a streamlined clinical development program for biosimilars (FDA 2016).

Table 22.6 Clinical studies conducted to support approval in the U.S

Biosimilar	PK/PD	Confirmatory safety and efficacy
Zarxio® (filgrastim-sndz) (source: FDA medical review posted on FDA website)	6 studies Randomized, double-blind, multiple-dose, 2-way crossover study comparing Zarxio to EU-approved Neupogen in N = 40 healthy volunteers. Randomized, double-blind, single-dose, 2-way crossover study comparing Zarxio to EU-approved Neupogen in N = 26 healthy volunteers Randomized, double-blind, 2 dose-level, multiple-dose, 2-way crossover study comparing Zarxio to EU-approved Neupogen in N = 56 healthy volunteers Randomized, double-blind, single-dose, 3-way crossover study comparing two different formulations of Zarxio drug substance and EU-approved Neupogen in n = 30 healthy volunteers Randomized, double-blind, single-dose, 2-way crossover study comparing Zarxio to EU-approved Neupogen in N = 24 healthy volunteers Randomized, double-blind, single-dose, 2-way crossover study comparing Zarxio to US-licensed Neupogen in N = 28 healthy volunteers	1 study Randomized, double-blind, comparison of Zarxio and US-licensed Neupogen for prevention of severe neutropenia in N = 218 patients with breast cancer being treated with up to 6 cycles of combination chemotherapy using docetaxel, doxorubicin, and cyclophosphamide

Inflectra® (infliximab-dyyb) (source: FDA medical review posted on FDA website)	3 studies Randomized, double-blind, single-dose study of 5 mg/kg of Inflectra, US-licensed Remicade, or EU-approved Remicade in healthy volunteers (n = 71/arm) 54-week randomized, double-blind study of Inflectra vs. EU approved Remicade in 250 patients with AS 54-week comparative clinical study of Inflectra vs. EU-approved Remicade in approximately 600 patients with RA who were on background methotrexate.	2 studies 54-week, randomized, double-blind, parallel group study conducted outside the US in 606 patients with moderate to severely active RA on background methotrexate 54-week randomized, double-blinded, parallel group in 250 AS patients
Erelzi® (etanercept-szzs) (source: FDA medical review posted on FDA website)	5 studies Randomized, double-blind, 2-way crossover in 54 healthy subjects, single dose, comparing Erelzi and US-licensed reference product Randomized, double-blind, 2-way crossover in 57 healthy subjects, single dose comparing Erelzi and EU-licensed reference product Randomized, double-blind, 2-way crossover in 54 healthy subjects, single dose, comparing Erelzi and EU-licensed reference product Randomized, double-blind, 2-way crossover in 54 healthy subjects, comparing Erelzi prefilled syringe to Erelzi autoinjector Cross-study comparison of two PK studies that utilized US-comparator vs EU-comparator	1 study with 2 treatment periods Randomized, double-blind, 52 weeks (treatment period 1 = 0–12 weeks, treatment period 2 = 12–30 weeks). N = 531 PsO patients. Treatment period 1 was parallel group design, followed thereafter by switchover design with 3 switching periods

(continued)

Table 22.6 (continued)

Biosimilar	PK/PD	Confirmatory safety and efficacy
Amjevita® (adalimumab-atto) (source: FDA medical review posted on FDA website)	3 studies Randomized, 3-parallel groups, single dose, N = 203 healthy subjects. Amjevita vs US reference product vs EU reference product PK data collected from RA efficacy study PK data collected from PsO efficacy study	2 studies Randomized, double-blinded, 2 parallel groups. 26 weeks. N = 526 RA patients. Randomized, double-blinded, 2 parallel groups. Weeks 1–16 of a 48 week study. N = 350 PsO patients
Mvasi® (bevacizumab-awwb) (source: FDA briefing document provided for product review at an Oncology Advisory Committee meeting)	2 studies Randomized, single-blind, single-dose, 3-arm, parallel group study in 202 healthy male subjects. Mvasi vs US-licensed and vs EU-approved Avastin Randomized, double-blind, single-dose comparing Mvasi to EU-approved Avastin in 642 patients with advanced NSCLC	1 study Randomized, double-blinded, parallel group in 642 patients with non-squamous NSCLC receiving first-line therapy with carboplatin and paclitaxel. Mvasi vs. EU-approved Avastin
Ogivri® (trastuzumab-dkst) (source: FDA briefing document provided for product review at an Oncology Advisory Committee meeting)	1 study Single-dose, randomized, double-blind, 3-arm, parallel group study in 120 healthy male subjects comparing Ogivri vs US-Herceptin and EU-Herceptin	1 study Two-part, multicenter, double-blind, randomized, parallel-group study. 493 HER-2 positive metastatic breast cancer patients Part 1: Ogivri or EU-Herceptin plus a taxane were administered for a minimum of 8 cycles Part 2: Patients with at least stable disease in Part 1 were allowed to continue with Ogivri or EU-Herceptin

AS ankylosing spondylitis, *EU* European Union, *FDA* Food and Drug Administration, *HER* human epidermal growth factor receptor, *NSCLC* non-small cell lung cancer, *PD* pharmacodynamics, *PK* pharmacokinetics, *PsO* psoriasis, *RA* rheumatoid arthritis

Future Outlook

Analytical methods have progressed enormously in the past several decades and will continue to improve in sensitivity, accuracy and resolution. Due to this rapid evolution in methodology, many techniques currently accepted as routine had not been invented or were not broadly available at the time when the reference products were first approved. As discussed in this chapter, advancements in analytics have resulted in the detection of new quality attributes and have revealed batch-to-batch differences in the reference medicine (Schiestl et al. 2011; Kim et al. 2017). Thus as our ability to detect and characterize quality attributes improves, it will become increasingly important to assess the clinical relevance of any such differences.

Improved product understanding and more detailed demonstrations of similarity using physicochemical and functional assays will result in diminished residual uncertainty and are likely to reduce the requirement for clinical confirmation of similarity, in particular for confirmatory safety and efficacy studies. However, while it is possible that health authorities may be willing to approve products based on a totality of evidence data package that do not include human clinical safety and efficacy studies, it will be important that patients and health care providers are educated to understand the scientific basis of such an approach. Failure to do so would undermine the acceptance of biosimilar products and abrogate their intended purposes of increasing access to biopharmaceutical treatments.

It is hoped that the current country-specific development paradigm for biosimilars will evolve in the coming years to accommodate a more global approach. At present, many regulatory jurisdictions require that a proposed biosimilar be compared against a locally approved reference medicine. Use of a reference medicine approved outside a given country for comparative clinical trials is acceptable in US, Europe, Japan and elsewhere, however only in cases where detailed evaluations have clearly demonstrated that the "foreign reference medicine" is essentially the same as the locally approved reference medicine. As a scientific matter, such bridging assessments are in fact unnecessary given that the reference medicines are almost always licensed based on a single data set of clinical studies, irrespective of where they are approved. Any subsequent change in the quality attributes of a reference medicine in a market regulated by an advanced health authority requires justification. Once this is understood and appreciated, it may be possible to further streamline biosimilar development programs using a "global reference product" as the comparator (Webster and Woollett 2017).

Overall, the totality of evidence describes a science based approach for the tailored development and approval of follow on biopharmaceuticals, termed biosimilars, with matching safety and efficacy. The concept is based on the analytical demonstration of molecular sameness followed by clinical evaluation to confirm similarity and address any residual uncertainty. Although biosimilars are relatively new in some countries, these products are established health care products in some regions such as the European Union, where over a decade of experience support their safe and effective use, as well as their ability to improve patient access and standard of care. While evolution in analytical capability and product understanding may

allow changes in the development paradigm of these medicines, it will be important to maintain a high level of education and understanding for biosimilars to create high levels of confidence in these products.

Declaration of Interest The authors are employees of Sandoz, a division of Novartis, which develops, manufactures and markets biopharmaceuticals, including biosimilar medicines.

References

Beck A, Sanglier-Cianferani S, Van Dorsselaer A. Biosimilar, biobetter and next generation antibody characterization by mass spectrometry. Anal Chem. 2012;84:4637–46.

Beck A, Debaene F, Diemer H, Wagner-Rousset E, Colas O, Van Dorsselaer A, Cianferani S. Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. J Mass Spectrom. 2015;50:285–97.

Center for Drug Evaluation and Research. Approval package for Application Number ANDA 77-857, Enoxaparin Sodium Injection. 2010. https://www.accessdata.fda.gov/drugsatfda_docs/anda/2010/077857Orig1s000.pdf. Accessed Jan 18, 2018.

Chapman K, Adjei A, Baldrick P, da Silva A, De Smet K, DiCocco R, et al. Waiving in vivo studies for monoclonal antibody biosimilar development: National and Global Challenges. MAbs. 2016;8(3):427–35. https://doi.org/10.1080/19420862.2016.1145331.

Christl L. FDA's overview of the regulatory guidance for the development and approval of biosimilar products in the US. 2018. https://www.fda.gov/downloads/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/therapeuticbiologicapplications/biosimilars/ucm428732.pdf. Accessed March 1, 2018.

Christl L, Lim S. Overview of the regulatory framework and FDA's guidance for the development and approval of biosimilar and interchangeable products in the US. 2018. https://www.fda.gov/downloads/AboutFDA/WorkingatFDA/FellowshipInternshipGraduateFacultyPrograms/PharmacyStudentExperientialProgramCDER/UCM587522.pdf. Accessed March 1, 2018.

EMA. Guideline on similar biological medicinal products. CHMP/437/04. 2005. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003517.pdf.

EMA. CHMP assessment report for Zarzio. Procedure No. EMEA/H/C/000917. 2008. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/000917/WC500046528.pdf.

EMA. European public assessment reports. 2018. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&murl=menus/medicines/medicines.jsp&mid=WC0b01ac058001d125&jsenabled=true.

European Medicines Agency. Guideline on similar biological medicinal products. CHMP/437/04 Rev 1. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176768.pdf. Accessed Jan 16, 2018.

FDA. Guidance for industry - scientific considerations in demonstrating biosimilarity to a reference product. 2015a. https://www.fda.gov/downloads/drugs/guidances/ucm291128.pdf. Accessed Jan 16, 2018.

FDA. Guidance for industry. Quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. 2015b. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf. Accessed March 12, 2018.

FDA. Transcript of the U.S. Arthritis Advisory Committee meeting held on Tuesday. 2016. https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM520028.pdf. Accessed January 18, 2018.

FDA. Guidance for industry- statistical approaches to evaluate analytical similarity. 2017a. https://www.fda.gov/downloads/drugs.guidancecomplianceregulatoryinformation/guidances/ucm576786.pdf. Accessed March 1, 2018.

FDA. Draft guidance-chemistry, manufacturing, and controls changes to an approved application: certain biological products. 2017b. https://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/General/UCM590118.pdf.

FDA. Zarxio (filgrastim-sndz) FDA drug approval package. 2018a. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/125553Orig1s000TOC.cfm. Accessed Jan 18, 2018.

FDA. Inflectra (infliximab-dyyb) FDA drug approval package. 2018b. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/125544Orig1s000TOC.cfm. Accessed Jan 18, 2018.

FDA. Erelzi (etanercept-szzs) FDA drug approval package. 2018c. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761042Orig1_toc.cfm. Accessed Jan 18, 2018.

FDA. Amjevita (adalimumab-atto) FDA drug approval package. 2018d. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761024_toc.cfm. Accessed Jan 18, 2018.

FDA. Briefing information for the July 13, 2017 meeting of the Oncologic Drugs Advisory Committee. (Morning session: Mvasi; Afternoon session: Ogivri). 2018e. https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/ucm566364.htm. Accessed Jan 18, 2018.

Generics and Biosimilars Initiative (GABI). Biosimilars approved in Europe. 2018. http://www.gabionline.net/Biosimilars/General/Biosimilars-approved-in-Europe. Accessed Jan 17, 2018.

Hausberger A, Lamanna WC, Hartinger M, Seidl A, Toll H, Holzmann J. Identification of low-level protein-related variants in filgrastim products presently available in highly regulated markets. BioDrugs. 2016;30:233–42.

Jahn EM, Schneider CK. How to systematically evaluate immunogenicity of therapeutic proteins – regulatory considerations. Nat Biotechnol. 2009;25:280–6.

Kim S, Song J, Park S, Ham S, Paek K, Kang M, et al. Drifts in ADCC-related quality attributes of Herceptin®: impact on development of a trastuzumab biosimilar. MAbs. 2017;9:704–14. https://doi.org/10.1080/19420862.2017.1305530.

Krendyukov A, Schiestl M. Extrapolation concept at work with biosimilar: a decade of experience in oncology. ESMO Open. 2018;3:e000319. https://doi.org/10.1136/esmoopen-2017-000319.

Lamanna WC, Mayer RE, Rupprechter A, Fuchs M, Higel F, Fristch C, et al. The structure-function relationship of disulfide bonds in etanercept. Sci Rep. 2017;7:3951–7.

Lamanna WC, Holzmann J, Cohen HP, Guo X, Schwiegler M, Stangler T, Seidl A, Schiestl M. Maintaining consistent quality and clinical performance of biopharmaceuticals. Exp Opin Biol Ther. 2018;18(4):369–79. https://doi.org/10.1080/14712598.2018.1421169.

Lim S, Christl L. Overview of the regulatory framework and FDA's guidance for the development and approval of biosimilar and interchangeable products in the US. 2017. https://www.fda.gov/downloads/AboutFDA/WorkingatFDA/FellowshipInternshipGraduateFacultyPrograms/PharmacyStudentExperientialProgramCDER/UCM587522.pdf. Accessed Jan 16, 2018.

Markus R, Liu J, Ramchandani M, Landa D, Born T, Kaur P. Developing the totality of evidence for biosimilars: regulatory considerations and building confidence for the healthcare community. BioDrugs. 2017;31(3):175–87. https://doi.org/10.1007/s40259-017-0218-5.

Martel N, Lee J, Wells PS. Risk for heparin-induced thrombocytopenia with unfractionated and low-molecular-weight heparin thromboprophylaxis: a meta-analysis. Blood. 2005;106:2710–5.

Martin PL, Bugelski PJ. Concordance of preclinical and clinical pharmacology and toxicology of monoclonal antibodies and fusion proteins: soluble targets. Br J Pharmacol. 2012;166(3):806–22. https://doi.org/10.1111/j.1476-5381.2011.01812.x.

van Meer PJ, Kooijman M, van der Laan JW, Moors EH, Schellekens H. The value of non-human primates in the development of monoclonal antibodies. Nat Biotechnol. 2013;31(10):882–3. https://doi.org/10.1038/nbt.2709.

Presentation given by Anthony Mire-Sluis, 9th Symposium on the Practical Applications of Mass Spectrometry in the Biotechnology Industry (Mass Spec 2012); September 11–14, 2012, San Diego, CA USA.

National Institute of Environmental Health Science. Alternatives to animal testing. 2017. https://www.niehs.nih.gov/health/topics/science/sya-iccvam/index.cfm. Accessed Jan 22, 2017.

Pisupati K, Tian Y, Okbazghi S, Benet A, Ackermann R, Ford M, Schwendeman A. A multidimensional analytical comparison of Remicade and the biosimilar Remsima. Anal Chem. 2017;89(9):4838–46. https://doi.org/10.1021/acs.analchem.6b04436.

Regulation.gov. Statistical approaches to evaluate analytical similarity. 2018. https://www.regulations.gov/docketBrowser?rpp=50&so=DESC&sb=postedDate&po=0&dct=PS&D=FDA-2017-D-5525. Accessed March 23, 2018.

Schiestl M, Stangler T, Torella C, Cepeljnik T, Toll H, Grau R. Acceptable changes in quality attributes of glycosylated biopharmaceuticals. Nat Biotechnol. 2011;29:310–2.

Singh SK. Impact of product-related factors on immunogenicity of biotherapeutics. J Pharm Sci. 2011;100:354–87.

Strand V, Girolomoni G, Schiestl M, Mayer RE, Friccus-Quecke H, McCamish M. The totality-of-the-evidence approach to the development and assessment of GP2015, a proposed etanercept biosimilar. Curr Med Res Opin. 2017;33(6):993–1003. https://doi.org/10.1080/03007995.2017.1288612.

Webster CJ, Woollett GR. A “global reference” comparator for biosimilar development. BioDrugs. 2017;31(4):279–86. https://doi.org/10.1007/s40259-017-0227-4.

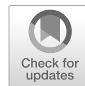
Weise M, Kurki P, Wolff-Holz E, Bielsky MC, Schneider CK. Biosimilars: the science of extrapolation. Blood. 2014;124:3191–6.

WHO. WHO Questions And Answers: Similar Biotherapeutic products. WHO/SBP_Q&A/DRAFT/DEC 2017. 2017. http://www.who.int/biologicals/QA_for_SBPs_HK_12_Dec_2017_(2).pdf. Accessed March 20, 2018.

World Health Organisation. Guidelines on evaluation of similar biotherapeutic products. 2009. http://www.who.int/biologicals/areas/biological_therapeutics/BIOTHERAPEUTICS_FOR_WEB_22APRIL2010.pdf. Accessed Jan 16, 2018.

Part VI
Biosimilars: Global Development and Clinical Experience

Chapter 23
Pharmacovigilance of Biosimilars: Global Experience and Perspective



Thomas Felix, Binakumari Patel, Brian D. Bradbury, and Gustavo Grampp

Abstract This chapter reviews important aspects of pharmacovigilance of biosimilars. Biologics are structurally complex molecules that are more difficult to characterize, produce, and reproduce than most small-molecule compounds. Ongoing robust pharmacovigilance is critical in the monitoring, detection, and assessment of safety signals over the life cycle of every biologic. The availability of multisource biologics, including biosimilars, warrants rigorous pharmacovigilance to accurately detect and disaggregate safety signals. Although biosimilars are highly similar to their reference biologics, they are not required or expected to be identical, and regulatory pathways permit slight variations in structural and pharmaceutical attributes and clinical development approaches. During development, candidate biosimilars are evaluated in a stepwise manner against their reference product for similarity in structure, function, clinical efficacy, and safety. However, clinical studies to evaluate biosimilarity may not detect rare adverse events, and potential differences in safety resulting from minor differences in manufacturing procedures between a biosimilar and its reference product (or other biosimilars) may not be detected before approval. Risk management plans, particularly during the early postmarketing period, are also an important component of pharmacovigilance planning for biosimilars. Important components of pharmacovigilance programs include ongoing and rigorous data collection, adverse event reporting, and analysis of causal relationships resulting in

T. Felix (✉)
Global Regulatory and R&D Policy, Amgen Inc., Washington, DC, USA
e-mail: thfelix@amgen.com

B. Patel
Global Patient Safety & Labelling, Amgen Inc., Uxbridge, UK
e-mail: bipatel@amgen.com

B. D. Bradbury
Center for Observational Research, Amgen Inc., Thousand Oaks, CA, USA
e-mail: bradbury@amgen.com

G. Grampp
Corporate Product Quality, Amgen Inc., Thousand Oaks, CA, USA
e-mail: ginog@amgen.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_23

accurate attribution of an adverse event to the correct product. Methods to improve product-specific monitoring, like the assignment of distinguishable nonproprietary names for all biologics and the use of additional product identifiers (e.g., batch number, trade name, manufacturer) for adverse event reporting, are vital to ensure accurate surveillance and traceability of all biologics, including biosimilars.

Keywords Biosimilars · Pharmacovigilance · Product-level traceability

Introduction

The patents for a large number of biologic drugs will expire in the next decade, creating the possibility of an influx of biosimilars into clinical practice. In the US, the Biologics Price Competition and Innovation Act provides an abbreviated pathway for the approval of biosimilars by the US Food and Drug Administration (FDA) (US Food and Drug Administration 2009); analogous pathways exist for different regions of the world (Biswas 2013; European Medicines Agency 2017a; Tsai 2017). Biosimilars are highly similar to their reference product; however, minor differences are expected and allowed because of the structural complexity of biologics and expected/anticipated differences in manufacturing procedures (Felix et al. 2014). Biosimilars are approved on the basis that any differences in manufacturing procedures and resulting analytical differences do not have clinical consequences (European Medicines Agency 2017a; US Food and Drug Administration 2015). However, it is difficult to predict how small structural changes may affect the safety or efficacy of a biosimilar compared with its reference product. Furthermore, clinical studies to evaluate biosimilarity for regulatory approval may not detect rare or delayed adverse events (AEs) resulting from such differences owing to relatively short duration and often small study population (Casadevall et al. 2013). Over the course of a biologic product's lifecycle, its structural integrity remains susceptible to changes in manufacturing and handling that could impact its efficacy, safety, or quality. This adds to the need for collection and analysis of postapproval safety data through effective surveillance systems that accurately track and trace all biologics and biosimilars (Casadevall et al. 2013). Pharmacovigilance is an integral part of the postmarketing phase for all biologic products, whether they are originator biologics or biosimilars; the goal is to effectively manage any safety risks by identifying and evaluating safety signals in a timely manner. This chapter will provide an overview of the importance of pharmacovigilance as it pertains to biologics and biosimilars and of the systems used in different regions, highlighting their strengths and limitations.

Background on Pharmacovigilance

The World Health Organization (WHO) defines pharmacovigilance as “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other possible drug-related problems” (World Health Organization 2002). Pharmacovigilance can generally be divided into 3 main areas: product quality, adverse drug reactions (ADRs), and medication errors (Management Sciences for Health Inc. 2012). Proper vigilance in each of these areas is important for maintaining stringent quality standards and optimizing product safety for patients.

Product quality generally pertains to the physiochemical product attributes that may vary based on manufacturing differences and changes during distribution. Manufacturing differences (e.g., expression systems, culture conditions, purification processes, formulation, drug product container) between a biosimilar and its reference product can lead to subtle differences in quality attributes (e.g., glycosylation, impurities) that may affect the safety profile, requiring ongoing safety monitoring (Grampp and Ramanan 2015). Other quality concerns may arise from physical changes that can occur because of improper storage or product tampering (Management Sciences for Health Inc. 2012). ADRs resulting from poor product quality can impose unnecessary burden on patients and healthcare providers (Pharmaceutical Technology Editors 2012). Such ADRs may lead to unnecessary emergency department visits, prolongation of hospital stays, or loss of product efficacy (Praditpornsilpa et al. 2011; Sultana et al. 2013). Poor product quality, whether or not it results in product recalls, can directly and negatively affect biopharmaceutical companies from a reputational and revenue perspective (Pharmaceutical Technology Editors 2012).

The International Conference on Harmonisation defines ADRs as negative reactions caused by the drug itself when administered as recommended (i.e., with the proper dose, frequency, route, and administration technique) and include allergic reactions, withdrawal effects, or interactions with other drugs, whereas AEs are defined as “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” (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 1994). Adverse drug events are defined as “harm caused by the use of a drug” (Nebeker et al. 2004) and therefore are synonymous with ADRs.

Adverse drug reactions have multiple health and economic effects. Studies indicate that ADRs represent one of the leading causes of morbidity and mortality in healthcare (Brvar et al. 2009; Bundy et al. 2012; Hug et al. 2012; International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 1994; Levinson 2010; Nebeker et al. 2004; Sultana et al. 2013; US Food and Drug Administration 2014d; Weiss et al. 2013) and can pose a significant burden to the healthcare system and budgets. It has been

estimated that nearly 4% of hospitalizations in France in 2006 and 2007 were due to ADRs (Benard-Laribiere et al. 2015). Similarly, in England between 1999 and 2008, hospitalizations associated with ADRs increased by 77%, from 42,453 to 75,076; in-hospital mortality associated with ADR admissions increased from 4.3% to 4.7% over this same time period (Wu et al. 2010). The economic effect of ADRs in the US may be as high as $30.1 billion annually (Sultana et al. 2013).

Medication errors (which may be classified as preventable ADRs) are defined by the European Medicines Agency (EMA) as “an unintended failure in the drug treatment process that leads to, or has the potential to lead to, harm to the patient” and are the most common preventable cause of ADRs (European Medicines Agency 2015). It has been estimated that medication errors account for 1 in 131 outpatient and 1 in 854 inpatient deaths (Wittich et al. 2014). There are numerous potential causes for medication errors, including medication factors such as similar sounding drug names, low therapeutic index, novel formulations, new packaging or similar packaging to another drug, and unclear labeling; patient factors such as polypharmacy or interaction with nonprescribed drugs or supplements, poor patient awareness or education, and comorbidities such as dementia; and factors related to healthcare professionals such as communication issues, poor handwriting on prescriptions, inexperience with administration, and administering to the incorrect patient (National Academies of Sciences Engineering and Medicine 2017; Wittich et al. 2014).

Importance of Pharmacovigilance

It was not until the thalidomide tragedy in 1961 during which children born to some women treated for morning sickness with thalidomide developed life-altering congenital deformities that systematic, international efforts to address drug safety issues were initiated. These efforts focused on developing a system “for detecting previously unknown or poorly understood adverse effects of medicines” (World Health Organization 2002). Registrational clinical trials provide data on potential safety issues associated with new drugs but are often considered insufficient owing to their relatively short duration, small sample sizes, and stringent inclusion/exclusion criteria that remove patients with significant comorbidities or who are receiving multiple concomitant medications (Mann and Andrews 2014; Reynolds et al. 2005). Consequently, product safety issues—often serious but rare—are not identified until sufficient exposure has accumulated in the postmarketing setting (Sharrar and Dieck 2013; US Food and Drug Administration 2005). A 2016 review by Onakpoya et al. found that between 1953 and 2013, of all the medicinal products approved worldwide, 462 were subsequently withdrawn from the market. In the 196 countries and states throughout the world, the most common reasons included death (25%), hepatotoxicity (18%), immune-related reactions (17%), neurotoxicity (16%), and cardiotoxicity (14%) (Onakpoya et al. 2016). Case

reports were used as evidence in 71% of all withdrawals, affirming the importance of effective pharmacovigilance systems.

Importance of Pharmacovigilance in a Multisource Biologics Market

Biosimilars and their reference products are complex molecules produced by biotechnology in living cells of humans, animals, or microorganisms. Approved biosimilars have highly similar physiochemical and functional characteristics to the reference biologic (European Medicines Agency 2014; US Food and Drug Administration 2009). However, unlike chemically synthesized drugs that contain the same active ingredient as the reference product, biosimilars are not exact replicas of the reference biologic because of differences in proprietary manufacturing processes (European Medicines Agency 2014; US Food and Drug Administration 2009). Regulatory pathways permit minor structural differences between biosimilars and their reference product because of these process differences (Felix et al. 2014; US Food and Drug Administration 2009). Although a biosimilar is required to be highly similar to its reference product, multiple biosimilars of the same reference product should not be expected to be a biosimilar to each other (i.e., biosimilarity is not transitive across all biosimilars to a common reference product) (Grampp and Ramanan 2015). Consequently, differences across biosimilars of the same reference product may be greater than differences between each biosimilar and the reference drug (Grampp and Ramanan 2015). It is important to recognize that there are numerous decisions made during the development of a biosimilar medicine and these are more complex than simply matching all critical quality attributes within the allowed reference product range. Sponsors of biosimilars may accept differences in certain attributes under the assumption that these will have no impact on function or clinical performance (Grampp and Ramanan 2015). For example, 2 Eprex® (epoetin alfa)/Erypro® (epoetin alfa) biosimilars—Retacrit® (epoetin zeta; SB309) and Binocrit® (epoetin alfa; HX-575)—have demonstrable differences in glycation attributes; compared with the reference epoetin alfa, variants of sialic acid (N-glycolyl neuraminic acid and O-acetyl neuraminic acid) were lower in the biosimilar products compared with the reference product (European Medicines Agency 2007b; Grampp and Ramanan 2015). Additional structural differences in terms of higher lactosamine repeats and lower levels of sialylation were found for Retacrit relative to Eprex/Erypro (Harazono et al. 2013). Differences in potency have also been reported, with Retacrit demonstrating 8% lower bioactivity than the reference product, likely resulting from a difference in protein concentration (European Medicines Agency 2007b). However, these products were deemed to be biosimilar on the basis of having no clinically relevant differences (European Medicines Agency 2007a, b).

The implication of these findings is that biosimilars cannot be administered in the same manner as multisource generic drugs without special considerations for reporting of ADRs (Grampp and Ramanan 2015). Switching among generics is commonly practiced by pharmacists without awareness by prescribers, a practice that poses further challenges for AE tracing because the AEs associated with a generic drug are frequently incorrectly attributed to the branded originator drug. For example, using publicly available data from the FDA's Adverse Event Reporting System (FAERS) between 2004 and 2012, Lietzan et al. reported that misattribution of AEs from generics to the branded originator occurred frequently for commonly used drugs that became subject to generic competition from 2005 to 2011, including Ambien® (zolpidem tartrate), Zocor® (simvastatin), and Zoloft® (sertraline hydrochloride) (Lietzan et al. 2013). Similarly, misattribution of heparin-induced thrombocytopenia from enoxaparin (Lovenox®) generics to the branded originator product in FAERS has also been observed. The manufacturer of the branded originator, Sanofi, processed approximately 73% of the FAERS reports before loss of exclusivity and 63% of reports after loss of exclusivity, despite losing 50% of market share; only 5% of AE reports were processed by generic manufacturers (Grampp et al. 2015).

Similarly, data from the Danish Medicines Agency showed a rapid market shift in 2015 from the branded originator infliximab (Remicade®) to biosimilar infliximab (Remsima™), with the biosimilar accounting for nearly half of the infliximab doses consumed in the second quarter and nearly all doses by the fourth quarter (Danish Medicines Agency 2016). During 2015, 55 suspected ADR reports were filed with the Danish Medicines Agency using the brand name Remicade, 64 using the brand name Remsima, and 6 with no brand name given. Whereas Remsima represented nearly all of the doses consumed in the fourth quarter of 2015, only about 75% of the suspected ADR reports during that period were attributed to Remsima. Although some of the discrepancy between market share and ADR reports during that time may reflect delayed events, it is possible that some of the events were misattributed to the originator brand.

Misattribution or ambiguous attribution has also been observed for biologics. In pharmacovigilance systems in the US, EU, and Australia, trade name attribution for somatropins, insulins, epoetins, filgrastims, and monoclonal antibodies ranged from 58% for filgrastim in Australia to 99% for epoetins in the EU (Grampp et al. 2016). Although the EMA best practices and the European Parliament call for the use of brand names in health records and AE reports (European Medicines Agency 2016; The European Parliament and the Council of the European Union 2010), 21% of filgrastim-related spontaneous reports between April 2012 and December 2014 received by Amgen were not attributed to a particular brand name (Amgen Inc. 2015). A similar occurrence was observed in Australia, where 42% of reports received by the Therapeutic Goods Administration (TGA) public database between March 2011 and November 2014 were coded as "filgrastim (not specified)" (Amgen Inc. 2015). ADRs specific to biologics and biosimilars include immune-related events; sometimes these can be delayed and manifest months or years after initial

exposure (Casadevall 2009). This was demonstrated by the pure red cell aplasia outbreak in Europe in patients with chronic kidney disease treated with recombinant human erythropoietin. In this scenario, a change in formulation resulted in a higher frequency of pure red cell aplasia than had previously been reported (Casadevall 2009). Similarly, a nonbiosimilar interferon-β follow-on version marketed in Latin America varied from the originator product in terms of structure and bioactivity (Meager et al. 2011). The same nonoriginator interferon-β was assessed in an observational study and was associated with lowered pharmacodynamic response and a lower frequency of pharmacology-associated ADRs (flu-like symptoms) relative to the originator product, suggesting a lower bioavailability (Cuevas et al. 2015). These results underscore the need for longitudinal patient medical records (Casadevall et al. 2013).

Vermeer et al. examined over 13,000 ADR reports submitted to EudraVigilance from 2004–2010 suspected to be related to biologic drugs and found that 96% had identifiable product names and 5% had both identifiable product names and batch numbers traceable to the manufacturer (Vermeer et al. 2013). More recently, Klein et al. found that 76% of biologics for which a suspected ADR report had been submitted to The Netherlands Pharmacovigilance Centre Lareb from 2009–2014 had an identifiable brand name, whereas 5% had a batch number (Klein et al. 2016). Implications of misattribution include a delay in detection of safety signals. Based on a sensitivity analysis of lag time in detecting disproportionate safety signals for a given product in scenarios with underreporting or misattribution of AEs among multisource biologics, a tool to calculate the impact of misclassified ADR reports was recently published (Vermeer et al. 2016). Using this tool, the authors found that misclassification likely led to a delay in identifying risks associated with a specific product; with a 34% misclassification rate (i.e., AE attributed to the wrong product), the number of cases and time to detection of the risk doubled (Vermeer et al. 2016).

Components of Pharmacovigilance

Although every region has its own specific methods for detecting and handling AEs associated with drugs, the basic components across each system are generally similar. These include signal detection, causality analysis and risk determination, reporting, decision making, and appropriate action, with interactions at the local, regional, national, and supranational levels (Management Sciences for Health Inc. 2012; World Health Organization 2015b).

Signal Detection

Data on safety are collected by pharmacovigilance systems through various methods, including voluntary case reporting (ie, spontaneous reporting), mandatory

reporting (varies by country), and active data collection, which includes trigger tools, patient chart audits, and direct observation methods, including prospective exposure registries or the use of healthcare databases, each with its own strengths and limitations (Management Sciences for Health Inc. 2012).

Spontaneous Reporting

All biopharmaceutical manufacturers are required to routinely report ADRs collected through voluntary reporting by patients and healthcare providers. These data are compiled in databases such as the FAERS in the US and EudraVigilance in the EU; however, underreporting is a recognized limitation (Jones and Kingery 2014). With the rare exception of newly launched drugs that have limited or restricted distribution, observing all patients exposed to a particular drug is simply not feasible across all jurisdictions, and as a result, companies provide estimates of postmarketing exposure based on sales data (Cook 2006; Telfair et al. 2006) to serve as denominators for ADR reporting rates.

Active Surveillance

Pharmacoepidemiologic studies using healthcare databases or disease/product exposure registries offer alternative approaches for conducting safety signal detection and evaluation (Mann and Andrews 2014; Reynolds et al. 2005). Postmarketing drug safety systems that rely on healthcare databases such as FDA's Sentinel system and the Asian Pharmacoepidemiology Network (AsPEN) leverage data collected as part of routine care, either through administrative claims or electronic health records, to identify and evaluate potential safety signals (AsPEN Collaborators et al. 2013; Lai et al. 2015; Platt et al. 2009; Robb et al. 2012). These systems by their nature rely on international coding systems for identifying medications (National Drug Codes [NDCs], Anatomical Therapeutic Chemical classification, Healthcare Common Procedure Coding System [HCPCS] and Current Procedural Terminology, 4th Edition [CPT-4] codes, or specific national coding systems) and disease events that occur in the inpatient or outpatient setting through the use of International Classification of Diseases, 9th/10th revision, codes. By employing rigorous study design and analytic techniques, regulators, biopharmaceutical manufacturers, and others in the healthcare ecosystem can conduct noninterventional studies to estimate risks related to drugs in thousands or millions of patients. Similar approaches can be taken using data collected as part of disease or exposure registries, as is commonly required by the EMA. Registries are sometimes deemed preferable owing to the standardized data collection approaches typically employed. Among the many recognized key design principles that are required for high-quality drug safety studies using data collected as part of registries or healthcare databases, accurate exposure classification is paramount for obtaining valid estimates of risk.

Causality Analysis and Risk Determination

After AE data are collected and collated, disproportionality analyses can be conducted to determine if there is evidence of a possible elevation in the reporting of a specific AE for a given medication relative to other similar medicines or to historical trends. If the results of these analyses suggest a potential signal, biopharmaceutical companies and regulatory bodies will review the evidence, determine plausibility, and if deemed of sufficient concern, additional (pharmacoepidemiologic) studies designed to evaluate the signal may be required. Much of the information on postmarketing AEs is obtained from case reports or case series; to date, no standards exist for determining causality in such cases. Typically the categories of probable, possible, or unlikely are used (US Food and Drug Administration 2005). Risk determination takes into account all available information. Per the FDA, factors considered in assessing the potential safety risk include the strength of the association, temporal relationship of product use and the event, consistency of findings across available data sources, evidence of a dose response for the effect, biological plausibility, seriousness of the event relative to the disease being treated, potential to mitigate the risk in the population, feasibility of further study using observational or controlled clinical study design, and degree of benefit the drug provides, including availability of other therapies (US Food and Drug Administration 2005).

Communication of Risk

Once a causal relationship is determined, it is necessary to share this information with prescribers, manufacturers, and the public. This may occur at the clinic, national, or international level (Management Sciences for Health Inc. 2012). Depending on the risks associated with the drug, it may be necessary to adjust the product labeling, institute a Risk Evaluation and Mitigation Strategy, issue a recall of a batch or lot, or withdraw the product completely from the market (e.g., Omontys® [peginesatide]; Affymax Inc., Cupertino, CA, USA, and Takeda Pharmaceutical Company Limited, Osaka, Japan) (Takeda 2014). Increasingly, changes in product labeling resulting from safety signals can also result in restrictions on reimbursement (Centers for Medicare and Medicaid Services 2007, 2017b). At the clinic level, the safety signal may result in a change in the medical formulary, implementation of new prescribing or dispensing procedures, increased patient monitoring procedures, or further education of professionals or patients (Management Sciences for Health Inc. 2012).

Importance of Unique Identifiers for All Biologics

Because of the complex systems for reporting AEs and the precedents set by generics and first-generation biosimilars, the potential for misattribution of AEs for biosimilars is high. Thus, the use of unique identifiers for all biologics, including biosimilars, is important for appropriate tracking and tracing of AEs if substitution occurs, with or without the prescriber's intervention (Portela et al. 2017). Definitions and rules regarding which biosimilars are interchangeable with the originator biologic and if substitution can occur vary by country and state. Regardless of such rules, the ability to trace an AE back to a specific biologic at a manufacturer and batch level is essential. The experience with pure red cell aplasia in patients treated with recombinant epoetin biosimilars underscores this (Casadevall 2009).

Several regulatory authorities have provided guidance on naming. FDA guidance indicates that each biosimilar should use a shared non-proprietary "core" name, followed by a unique 4-letter suffix designated by the FDA (the US Food and Drug Administration 2017). In the 2016 Physician Fee Schedule Final Rule, the Centers for Medicare and Medicaid Services (CMS) proposed that biosimilars based on a common reference product were to be grouped into the same payment calculation and share a common payment limit and HCPCS code (Centers for Medicare and Medicaid Services 2017c). The CMS recently issued new guidance on reimbursement for biosimilars stating that beginning January 1, 2018, all approved biosimilars will receive their own HCPCS reimbursement code (Brennan 2017; Centers for Medicare and Medicaid Services 2017a). The EMA has issued a directive to "ensure, through the methods for collecting information and where necessary through the follow-up of suspected adverse reaction reports, that all appropriate measures are taken to identify clearly any biological prescribed, dispensed or sold in their territory which is the subject of a suspected adverse reaction report, with due regard to the name of the medicinal product . . . and the batch number" (European Medicines Agency 2016). "To fulfil this obligation, national competent authorities should agree with marketing authorisation holders, where applicable, a system to ensure the traceability of the biologicals that are prescribed, dispensed or sold, inform healthcare professionals and patients of the need to provide the product name (i.e., brand/invented name or, as appropriate, INN accompanied by the name of the marketing authorisation holder) and batch number/code when reporting a suspected adverse reaction and make this information available to assessors for signal detection and evaluation of individual case reports" (European Medicines Agency 2016).

Although the initial recommendation provided by the WHO INN Expert Group stated that biosimilars did not require unique names, they later proposed the addition of a biological qualifier to the INN (World Health Organization 2006, 2015a).

In addition to the use of unique names, additional steps can be taken to ensure accuracy of ADR reporting. Manufacturers of reference products can apply additional screening at points where reports are received to clarify whether the report is truly applicable to the reference product or a biosimilar from another company. This can be accomplished by preparing call centers to ask reporters additional

questions or sending back queries for ADR reports to clarify the implicated product. For example, when a reporter is in a region where biosimilars are approved, this could trigger additional screening measures. Furthermore, inclusion of 2- and 3-dimension barcodes on all biologic drugs should improve traceability, with the full implementation of the US Drug Quality and Security Act/Drug Supply Chain Security Act (DQSA/DSCSA) that outlines an interoperable electronic system to identify and trace prescription drugs in the US (Drug Quality and Security Act, Public Law 113–54 2013). A similar measure has been implemented by the EU through the Falsified Medicines Directive, which will require medicines to have a unique randomized number contained in a 2-dimensional barcode (Robinson 2016). These measures should enable improved traceability of ADRs while ensuring authenticity of dispensed medicines.

US: Current Methods for Monitoring and Reporting Adverse Events or Risks Associated with Drugs

In the US, the Center for Drug Evaluation and Research within the FDA coordinates the premarketing and postmarketing safety monitoring of drugs, including biologics and biosimilars, and has issued guidance for industry for good pharmacovigilance practice (US Food and Drug Administration 2005, 2014a). Both spontaneous reporting and active surveillance systems are used for monitoring AEs in the postmarketing phase (Grampp and Felix 2015). Spontaneous reporting occurs through reports to the manufacturer and the FDA MedWatch programs. Reports are stored in the FAERS database, which is a computerized repository for reports of AEs and medication errors (US Food and Drug Administration 2016). These reports are voluntarily submitted by physicians, pharmacists, or other healthcare providers; patients can also submit through a MedWatch reporting form (US Food and Drug Administration 2014c). Manufacturers are required to periodically submit reports of AEs (US Food and Drug Administration 2012). Serious reactions, product quality problems, or incidences of therapeutic failure can be reported. Of course, this has the limitation of depending on healthcare providers and/or patients to recognize the event and then follow through and report it; consequently, although spontaneous reporting can be valuable to identify emergent safety signals, this system cannot quantify the incidence of AEs or risks to patients (Grampp and Felix 2015). In a systematic literature review, Onakpoya et al. noted that although the interval between launch date and ADR reports had shortened over the past few decades, product withdrawals following reports of suspected ADRs that were sufficiently severe to warrant product removal had not consistently improved (Onakpoya et al. 2016). A potential reason for the delay in withdrawal decisions is underreporting of ADRs. Despite being a known public health problem, the underreporting of ADRs among clinicians is well documented (Brvar et al. 2009; Gonzalez-Rubio

et al. 2011; Martin et al. 1998; Perez Garcia and Figueras 2011) and is one of the main challenges facing pharmacovigilance systems.

Furthermore, reports can be submitted to the FDA with little product identification; typically the forms include an identifiable patient, an identifiable reporter, a suspect drug, and the AE or fatal outcome (US Food and Drug Administration 2001). The form includes a field for drug name and strength. The form does not include a separate field for the brand name or nonproprietary name; reports with the same nonproprietary name (regardless of manufacturer) are grouped together, further limiting accurate product and manufacturer traceability (US Department of Health and Human Services 2015). There have been recent efforts at the state level to improve transparency of patient prescription drug records with respect to substitution of biosimilars, including legislation requiring pharmacists to notify prescribers when interchangeable biosimilars are substituted and state efforts to ensure that product identifiers are transmitted by pharmacists to the prescriber upon such substitution either directly or via an accessible database (Cauchi and National Conference of State Legislatures 2017; Singh and Bagnato 2015).

Active surveillance occurs through a number of mechanisms, including retrospective database studies, registries, open-label long-term studies, and prospective observational studies (Grampp and Felix 2015). The most common is retrospective analysis of existing databases. The FDA's Sentinel Initiative, which began in 2008 as Mini-Sentinel, uses electronic health information available from claims data and registries, inpatient and outpatient healthcare records, and patient registries to match a specific medicine (provided that billing codes are specific) to clinically reported outcomes (Behrman et al. 2011; US Food and Drug Administration 2014b). Advantages of this system include (1) the use of NDCs and product-specific HCPCS codes, where possible, to identify patients receiving specific medicines; (2) the ability to follow patients longitudinally over time to more fully capture patient exposure to medicines including duration; and (3) the ability to identify adverse clinical events recorded during inpatient hospitalizations and outpatient visits for exposed patients. Together, this information in context of a well-designed pharmacoepidemiologic study can provide more interpretable estimates of patient risk. There are, however, important limitations to these systems, including (1) inconsistent identification or coding of AEs, and in particular, immunologic reactions, resulting in the potential for outcome misclassification; and (2) no coding of lot numbers.

EU: Current Methods for Monitoring and Reporting Adverse Events

In many ways, the EMA pharmacovigilance system is similar to that of the FDA. The EMA is responsible for coordinating the existing scientific resources for the pharmacovigilance of biologic drugs and coordinating the assessment of the risk analysis submitted by the marketing authorization holder. It also coordinates peri-

odic safety updates for products and outlines best practices for pharmacovigilance for all member states in the EU (European Medicines Agency 2016).

EudraVigilance

EudraVigilance is a database maintained by the EMA that brings together all serious and nonserious adverse reactions reported within the EU as well as serious ADR reports from outside the EU during the postauthorization period submitted by marketing authorization holders in accordance with EU legislation (European Medicines Agency 2017b). Reports of suspected ADRs can be submitted by the marketing authorization holder, healthcare professionals, patients, and EU national competent authorities (Directive 2004/27/EC of the European Parliament and the Council of 31 March 2004 amending directive 2001/83/EC on the community code relating to medicinal products for human use 2004; European Medicines Agency 2017b). In contrast to the FDA system, as indicated previously, the EU pharmaceutical directive states that reports of AEs are required to include the batch number for traceability. Despite this requirement, evidence suggests this rule is not always followed. For example, in a case study conducted by Vermeer et al., more than 2 million unique ADR reports were examined from the EudraVigilance and FAERS databases and the level of information available in these reports was assessed (Vermeer et al. 2013). In EudraVigilance, batch numbers were included in 21.1% of reports on suspected events with biopharmaceuticals and in 3.6% of reports with small-molecule drugs. For FAERS, the respective percentages were 24.0% and 7.4% (Vermeer et al. 2013). Interestingly, consumers were more likely than healthcare providers to provide this information (Vermeer et al. 2013). Similarly, Klein et al. found that batch numbers were infrequently recorded in ADR reports (5%) on biologics received by The Netherlands Pharmacovigilance Centre Lareb, whereas brand name recording was 76% for ADR reports of biologics overall and slightly lower (67%) for monoclonal antibodies (Klein et al. 2016). These examples highlight insufficiencies in spontaneous reporting systems in both the EU and US systems relevant to biologics and biosimilars.

Other Regions: Current Methods for Monitoring and Reporting Adverse Events

Japan

Japan is the second-largest pharmaceutical market (Biswas 2013; Tsai 2017). Over the past 2 decades, Japan has taken a more conservative approach toward biosimilar development compared with other Asian countries (Tsai 2017). The Pharmaceuticals and Medical Devices Agency, a subsidiary of the Ministry of Health, Labour and Welfare, is responsible for the scientific assessment and marketing

authorization of all pharmaceuticals, including biosimilars. Japan's guidance on biosimilars is based on EU principles (Tsai 2017). All medical institutions and pharmacies participate in pharmacovigilance (Biswas 2013). The Pharmaceutical Affairs Act stipulates that all healthcare providers are to report ADRs or suspected hazards (Biswas 2013). Automatic substitution of biosimilars is not permitted (Tsai 2017), preventing associated challenges with ADR attribution when substitution is practiced.

South Korea

South Korea has an active biosimilars market, with biosimilars originating from South Korea approved in Japan, the US, and the EU (Tsai 2017). Pharmacovigilance occurs primarily through monitoring of spontaneous reporting of suspected ADRs through the Decentralized Pharmacovigilance System, with Regional Pharmacovigilance Centers (such as local teaching hospitals) reporting to the central center using an electronic reporting system (Biswas 2013). Similar to Japan, challenges related to ADR attribution are lessened because automatic substitution of biosimilars is not permitted (Tsai 2017).

Taiwan

In Taiwan, biologic drugs used for serious diseases such as rheumatoid arthritis are provided at no cost to patients, creating a need for lower-cost biosimilars (Tsai 2017). The Taiwan Food and Drug Administration has established guidelines for pharmacovigilance for biosimilars stating that it "is important to verify the product name and batch number of biological product to define related adverse reactions and such information shall be collected through appropriate methods" (Taiwan Food and Drug Administration 2015).

Thailand

Since 2010, Thailand has had a national Health Product Vigilance Center, which is overseen by the Thai Food and Drug Administration. Suspected ADRs are reported by the community and hospitals (Biswas 2013). During the outbreak of pure red cell aplasia from erythropoietin antibodies, the incidence was higher in Thailand than in other affected countries (Casadevall 2009). At the time of this outbreak, storage and cold chain were not guaranteed at out-of-hospital pharmacies, there was often substitution with no traceability, and counterfeit products existed in Thailand (Casadevall 2009). Furthermore, biosimilar recombinant erythropoietin products were licensed in Thailand following the conventional generic paradigm that focused on bioequivalence, without clinical data or a set of data similar to that required for biosimilar approval by the US FDA (Praditpornsilpa et al. 2011; US Food and

Drug Administration 2009). Research examining biosimilar erythropoietin licensed through the generics pathway demonstrated increased neutralizing antibodies leading to loss of efficacy and increased incidence of pure red cell aplasia, indicating that such an approval pathway was not appropriate for these drugs (Praditpornsilpa et al. 2011). Subsequent analyses of the quality attributes of marketed noncomparable erythropoietins from Thailand confirmed the existence of several products with inadequate quality, likely correlating with elevated immunogenicity (Halim et al. 2014).

World Health Organization

The thalidomide disaster in 1961 prompted WHO to establish the Programme for International Drug Monitoring, which promotes pharmacovigilance at the country level (World Health Organization 2002). To participate in the program, a country must have a system for collecting individual case safety reports, relevant funding to maintain this system, and a national center for drug monitoring designated by the country's Ministry of Health or equivalent (World Health Organization 2010). As of the end of 2010, 134 countries were participating in this program (World Health Organization 2010, 2017).

Pharmacovigilance in Other Areas of Medicine

Pharmacovigilance of drugs that treat HIV necessitates long-term safety monitoring because many patients with HIV have taken antiretroviral therapies for decades (Marcus et al. 2012). Thus, the current pharmacovigilance infrastructure needs to adapt to accommodate the millions of patients infected worldwide. Changes to FAERS with the Sentinel Initiative and the Pharmacovigilance Risk Assessment Committee within EudraVigilance are expected to assist with this endeavor (Marcus et al. 2012).

The Vaccine Safety Datalink program, sponsored by the Centers for Disease Control and Prevention, represents a population of 9.2 million people per year. Surveillance is conducted in near real-time with sophisticated statistical methods analyzing the dataset for true safety signals versus noise or false positives (VAERS 2014; Yih et al. 2011). The Sentinel system is based in part on this database (Behrman et al. 2011), and it is hoped that in the future such a system will exist for biologic drugs and biosimilars.

Conclusions

Because of their large size and greater complexity compared with small-molecule generics, biosimilars are more susceptible to subtle structural differences from their reference product due to variations in manufacturing procedures and sensitivity to handling. Although biosimilars are highly similar to their reference biologic, it is not possible for them to be identical to their reference product owing to the nature of biologic medicines. Accordingly, they are not required or expected to be identical, and regulatory pathways permit these expected slight variations in structural and pharmaceutical attributes and clinical development approaches. In some cases, despite extensive testing, it can be difficult to predict how small structural changes may affect the safety and/or efficacy of a biosimilar compared with its reference product. Moreover, clinical trials designed to assess biosimilarity may not detect rare or delayed safety events, thereby requiring robust pharmacovigilance. Continuous assessment of the risk-benefit profile of biologics, including biosimilars, is critical and relies on global, national, and local infrastructure to promote accurate identification and reporting of suspected ADRs, especially in a multimanufacturer environment created when multiple biosimilars for a single reference product are available.

Robust pharmacovigilance systems require accurate recording of product information for safety monitoring of all medicines, including biologics and biosimilars. This necessitates the consistent use of a brand or distinguishable nonproprietary name, and fields to include manufacturer name, batch, and lot number. Inaccurate or ambiguous product identification in reporting suspected ADRs hinders effective, accurate tracing of immunogenic AEs and may result in a loss of efficacy or the occurrence of other AEs (Casadevall 2009; Grampp et al. 2015). As more biosimilars are approved worldwide, it will be important to improve pharmacovigilance systems to ensure accurate recording and reporting of suspected ADRs (e.g., through increased education, use of batch codes and technologies such as 2D barcodes) to monitor safety of this new drug class.

Acknowledgments Medical writing support was provided by Miranda Tradewell, PhD, Rick Davis, MS, and James Balwit, MS, CMPP (Complete Healthcare Communications, LLC, Wales Chester, PA, an ICON plc company), whose work was funded by Amgen Inc.

Conflicts of Interest All authors are employees of and own stock and/or hold stock options in Amgen Inc.

References

Amgen Inc. Docket Nos. FDA-2013-P-1153; FDA-2013-P-1398, respectively, nonproprietary naming of biosimilars. 2015. Available at http://www.regulations.gov/#!documentDetail;D=FDA-2013-P-1398-0022. Accessed October 19, 2017.

AsPEN Collaborators, Andersen M, Bergman U, Choi NK, Gerhard T, Huang C, Jalbert J, Kimura M, Kimura T, Kubota K, Lai EC, Ooba N, Park BJ, Pratt N, Roughead EE, Sato T, Setoguchi S, Shin JY, Sundstrom A, Yang YH. The Asian Pharmacoepidemiology Network (AsPEN): promoting multi-national collaboration for pharmacoepidemiologic research in Asia. Pharmacoepidemiol Drug Saf. 2013;22:700–4.

Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the Sentinel System–a national resource for evidence development. N Engl J Med. 2011;364:498–9.

Benard-Laribiere A, Miremont-Salame G, Perault-Pochat MC, Noize P, Haramburu F. Incidence of hospital admissions due to adverse drug reactions in France: the EMIR study. Fundam Clin Pharmacol. 2015;29:106–11.

Biswas P. Pharmacovigilance in Asia. J Pharmacol Pharmacother. 2013;4:S7–S19.

Brennan Z. CMS shift on biosimilar reimbursement under part B: what to expect. 2017. Available at http://www.raps.org/Regulatory-Focus/News/2017/11/06/28829/CMS-Shift-on-Biosimilar-Reimbursement-Under-Part-B-What-to-Expect/. Accessed November 14, 2017.

Brvar M, Fokter N, Bunc M, Mozina M. The frequency of adverse drug reaction related admissions according to method of detection, admission urgency and medical department specialty. BMC Clin Pharmacol. 2009;9:8.

Bundy DG, Marsteller JA, Wu AW, Engineer LD, Berenholtz SM, Caughey AH, Silver D, Tian J, Thompson RE, Miller MR, Lehmann CU. Electronic health record-based monitoring of primary care patients at risk of medication-related toxicity. Jt Comm J Qual Patient Saf. 2012;38:216–23.

Casadevall N. Immune-response and adverse reactions: PRCA case example. 2009. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2009/11/WC500011064.pdf. Accessed October 18, 2017.

Casadevall N, Edwards IR, Felix T, Graze PR, Litten JB, Strober BE, Warnock DG. Pharmacovigilance and biosimilars: considerations, needs, and challenges. Expert Opin Biol Ther. 2013;13:1039–47.

Cauchi R, National Conference of State Legislatures. State laws and legislation related to biologic medications and substitution of biosimilars. 2017. Available at http://www.ncsl.org/research/health/state-laws-and-legislation-related-to-biologic-medications-and-substitution-of-biosimilars.aspx. Accessed October 20, 2017.

Centers for Medicare and Medicaid Services. Part B biosimilar biological product payment and required modifiers. 2007. Available at https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Part-B-Drugs/McrPartBDrugAvgSalesPrice/Part-B-Biosimilar-Biological-Product-Payment.html. Accessed August 16, 2017.

Centers for Medicare and Medicaid Services. Decision memo for erythropoiesis stimulating agents (ESAs) for non-renal disease indications (CAG-00383N). 2017a. Available at https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=203&NcaName=Erythropoiesis+Stimulating+Agents+. Accessed May 1, 2018.

Centers for Medicare and Medicaid Services. Final policy, payment, and quality provisions in the Medicare physician fee schedule for calendar year 2018. 2017b. Available at: https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2017-Fact-Sheet-items/2017-11-02.html. Accessed November 16, 2017.

Centers for Medicare and Medicaid Services. Medicare national coverage determinations manual. Chapter 1, part 2 (sections 90-160.26). Coverage determinations (Rev. 200, 07-25-17). 2017c. Available at https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/ncd103c1_Part2.pdf. Accessed May 1, 2018

Cook MN. Estimating national drug consumption using data at different points in the pharmaceutical supply chain. Pharmacoepidemiol Drug Saf. 2006;15:754–7.

Cuevas C, Deisenhammer F, You X, Scolnik M, Buffels R, Sperling B, Flores-Ramirez F, Macias-Islas M, Sauri-Suarez S. Low immunogenicity but reduced bioavailability of an interferon beta-1a biosimilar compared with its biological parent: results of MATRI X, a cross-sectional, multicenter phase 4 study. Biosimilars. 2015;5:75–81.

Danish Medicines Agency. Danish pharmacovigilance update. 2016. Available at https://laegemiddelstyrelsen.dk/en/news/nyhedsbrevet-nyt-om-bivirkninger-arkiv/danish-pharmacovigilance-update,-february-2016/~/media/FF5865ABAC6D492AB650739216D65DCB.ashx. Accessed June 13, 2018.

Directive 2004/27/EC of the European Parliament and the Council of 31 March 2004 amending directive 2001/83/EC on the community code relating to medicinal products for human use. Off J Eur Union. 2004;136:34–57.

Drug Quality and Security Act, Public Law 113–54. 2013. Available at http://www.gpo.gov/fdsys/pkg/PLAW-113publ54/content-detail.html. Accessed July 6, 2015.

European Medicines Agency. Biosimilars in the EU. Information guide for healthcare professionals. 2007a. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2017/05/WC500226648.pdf. Accessed August 16, 2017.

European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VI - collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 2007b. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2017/08/WC500232767.pdf. Accessed December 12, 2017.

European Medicines Agency. Guideline on similar biological medicinal products. 2014. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176768.pdf. Accessed October 16, 2017.

European Medicines Agency. Good practice guide on recording, coding, reporting and assessment of medication errors. 2015. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2015/11/WC500196979.pdf. Accessed December 12, 2017.

European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Product- or population-specific considerations II: biological medicinal products. 2016. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/08/WC500211728.pdf. Accessed August 17, 2017.

European Medicines Agency. Binocrit: European public assessment report: scientific discussion. 2017a. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000725/WC500053615.pdf. Accessed March 6, 2017.

European Medicines Agency. Retacrit: European public assessment report: scientific discussion. 2017b. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000872/WC500054374.pdf. Accessed March 6, 2017.

Felix T, Johansson TT, Colliatie JA, Goldberg MR, Fox AR. Biologic product identification and US pharmacovigilance in the biosimilars era. Nat Biotechnol. 2014;32:128–30.

Gonzalez-Rubio F, Calderon-Larranaga A, Poblador-Plou B, Navarro-Peman C, Lopez-Cabanas A, Prados-Torres A. Underreporting of recognized adverse drug reactions by primary care physicians: an exploratory study. Pharmacoepidemiol Drug Saf. 2011;20:1287–94.

Grampp G, Felix T. Pharmacovigilance considerations for biosimilars in the USA. BioDrugs. 2015;29:309–21.

Grampp G, Ramanan S. The diversity of biosimilar design and development: implications for policies and stakeholders. BioDrugs. 2015;29:365–72.

Grampp G, Bonafede M, Felix T, Li E, Malecki M, Sprafka JM. Active and passive surveillance of enoxaparin generics: a case study relevant to biosimilars. Expert Opin Drug Saf. 2015;14:349–60.

Grampp G, Kozak RW, Schreitmueller T. Policy considerations for originator and similar biotherapeutic products. Pharm Policy Law. 2016;18:121–39.

Halim LA, Brinks V, Jiskoot W, Romeijn S, Praditpornsilpa K, Assawamakin A, Schellekens H. How bio-questionable are the different recombinant human erythropoietin copy products in Thailand? Pharm Res. 2014;31:1210–8.

Harazono A, Hashii N, Kuribayashi R, Nakazawa S, Kawasaki N. Mass spectrometric glycoform profiling of the innovator and biosimilar erythropoietin and darepoetin by LC/ESI-MS. J Pharm Biomed Anal. 2013;83:65–74.

Hug BL, Keohane C, Seger DL, Yoon C, Bates DW. The costs of adverse drug events in community hospitals. Jt Comm J Qual Patient Saf. 2012;38:120–6.

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Clinical safety data management: definitions and standards for expedited reporting E2A. 1994. Available at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf. Accessed October 18, 2017.

Jones JK, Kingery EE. History of pharmacovigilance. In: Andrews EB, Moore N, editors. Mann's pharmacovigilance. New York: Wiley Blackwell; 2014. p. 11–24.

Klein K, Scholl JH, Vermeer NS, Broekmans AW, Van Puijenbroek EP, De Bruin ML, Stolk P. Traceability of biologics in the Netherlands: an analysis of information-recording systems in clinical practice and spontaneous ADR reports. Drug Saf. 2016;39:185–92.

Lai EC, Man KK, Chaiyakunapruk N, Cheng CL, Chien HC, Chui CS, Dilokthornsakul P, Hardy NC, Hsieh CY, Hsu CY, Kubota K, Lin TC, Liu Y, Park BJ, Pratt N, Roughead EE, Shin JY, Watcharathanakij S, Wen J, Wong IC, Yang YH, Zhang Y, Setoguchi S. Brief report: databases in the Asia-Pacific region: the potential for a distributed network approach. Epidemiology. 2015;26:815–20.

Levinson DR. Adverse events in hospitals: national incidence among Medicare beneficiaries. 2010. Available at https://oig.hhs.gov/oei/reports/oei-06-09-00090.pdf. Accessed October 18, 2017.

Lietzan EF, Sim LE, Alexander EA. Biosimilar naming: how do adverse event reporting data support the need for distinct nonproprietary names for biosimilars? FDLIs Food Drug Policy Forum. 2013;3:1–24.

Management Sciences for Health Inc. Chapter 35. Pharmacovigilance. In: MDS-3: managing access to medicines and health technologies. Arlington: Management Sciences for Health; 2012. p. 35.1–35.18.

Mann RD, Andrews EB. Introduction: updated from second edition. In: Andrews EB, Moore N, editors. Mann's pharmacovigilance. New York: Wiley Blackwell; 2014. p. 1–9.

Marcus KA, Sorbello A, Truffa M, Williams J, Raine JM, Powderly WG. Current advances in pharmacovigilance in the USA and Europe: meeting the challenges of safety monitoring in HIV. Curr Opin HIV AIDS. 2012;7:292–8.

Martin RM, Kapoor KV, Wilton LV, Mann RD. Underreporting of suspected adverse drug reactions to newly marketed ("black triangle") drugs in general practice: observational study. BMJ. 1998;317:119–20.

Meager A, Dolman C, Dilger P, Bird C, Giovannoni G, Schellekens H, Thorpe R, Wadhwa M. An assessment of biological potency and molecular characteristics of different innovator and noninnovator interferon-beta products. J Interf Cytokine Res. 2011;31:383–92.

National Academies of Sciences Engineering and Medicine. Communicating clearly about medicines: proceedings of a workshop. Washington, DC: The National Academies Press; 2017. p. 1–124.

Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. Ann Intern Med. 2004;140:795–801.

Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC Med. 2016;14:10.

Perez Garcia M, Figueras A. The lack of knowledge about the voluntary reporting system of adverse drug reactions as a major cause of underreporting: direct survey among health professionals. Pharmacoepidemiol Drug Saf. 2011;20:1295–302.

Pharmaceutical Technology Editors. Costs of failure in product quality. 2012. Available at http://www.pharmtech.com/costs-failure-product-quality. Accessed October 6, 2017.

Platt R, Wilson M, Chan KA, Benner JS, Marchibroda J, McClellan M. The new Sentinel Network–improving the evidence of medical-product safety. N Engl J Med. 2009;361:645–7.
Portela M, Sinogas C, Albuquerque de Almeida F, Baptista-Leite R, Castro-Caldas A. Biologicals and biosimilars: safety issues in Europe. Expert Opin Biol Ther. 2017;17:871–7.
Praditpornsilpa K, Tiranathanagul K, Kupatawintu P, Jootar S, Intragumtornchai T, Tungsanga K, Teerapornlertratt T, Lumlertkul D, Townamchai N, Susantitaphong P, Katavetin P, Kanjanabuch T, Avihingsanon Y, Eiam-Ong S. Biosimilar recombinant human erythropoietin induces the production of neutralizing antibodies. Kidney Int. 2011;80:88–92.
Reynolds RF, Glasser DB, Dieck GS. A view from industry. In: Strom BL, editor. Pharmacoepidemiology. New York: Wiley; 2005. p. 77–101.
Robb MA, Racoosin JA, Sherman RE, Gross TP, Ball R, Reichman ME, Midthun K, Woodcock J. The US Food and Drug Administration's Sentinel Initiative: expanding the horizons of medical product safety. Pharmacoepidemiol Drug Saf. 2012;21:9–11.
Robinson J. What is the falsified medicines directive? 2016. Available at https://www.pharmaceutical-journal.com/your-rps/what-is-the-falsified-medicines-directive/20201820.article. Accessed February 21, 2018.
Sharrar RG, Dieck GS. Monitoring product safety in the postmarketing environment. Ther Adv Drug Saf. 2013;4:211–9.
Singh SC, Bagnato KM. The economic implications of biosimilars. Am J Manag Care. 2015;21:s331–40.
Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. J Pharmacol Pharmacother. 2013;4:S73.
Taiwan Food and Drug Administration. Guidelines on drug review and approval–guidelines on biosimilar products. 2015.
Takeda. Affymax and Takeda announce termination of Omontys® (peginesatide) product collaboration and license agreement. 2014. Available at https://www.takeda.com/newsroom/newsreleases/2014/affymax-and-takeda-announce-termination-of-omontys-peginesatide%2D%2Dproduct-collaboration-and-license-agreement/. Accessed February 6, 2018.
Telfair T, Mohan AK, Shahani S, Klincewicz S, Atsma WJ, Thomas A, Fife D. Estimating post-marketing exposure to pharmaceutical products using ex-factory distribution data. Pharmacoepidemiol Drug Saf. 2006;15:749–53.
The European Parliament and the Council of the European Union. Directive 2010/84/EU of the European Parliament and of the Council of 15 December 2010 amending, as regards pharmacovigilance, Directive 2001/83/EC on the Community code relating to medicinal products for human use. 2010. pp. 74–99.
Tsai WC. Update on biosimilars in Asia. Curr Rheumatol Rep. 2017;19:47.
US Department of Health and Human Services. MedWatch. The FDA Safety Information and Adverse Event Reporting System. 2015. Accessed October 20, 2017.
US Food and Drug Administration. Guidance for industry: postmarketing safety reporting for human drug and biological products including vaccines. 2001. Available at http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm092257.pdf. Accessed October 27, 2014.
US Food and Drug Administration. Guidance for industry: good pharmacovigilance practices and pharmacoepidemiologic assessment. 2005. Available at https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071696.pdf. Accessed October 19, 2017.
US Food and Drug Administration. Biologics price competition and innovation act of 2009. 2009. Available at https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/ucm216146.pdf. Accessed October 18, 2017.
US Food and Drug Administration. Guidance for industry: nonproprietary naming of biological products. 2012. Available at https://www.fda.gov/downloads/drugs/guidances/ucm459987.pdf. Accessed February 20, 2017.

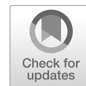
US Food and Drug Administration. About the Center for Drug Evaluation and Research. 2014a. Available at https://www.fda.gov/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/. Accessed August 17, 2017.
US Food and Drug Administration. FDA's Sentinel Initiative. 2014b Available at http://www.fda.gov/Safety/FDAsSentinelinitiative/ucm2007250.htm. Accessed October 18, 2017.
US Food and Drug Administration. How consumers can report an adverse event or serious problem to the FDA. 2014c. Available at http://www.fda.gov/Safety/MedWatch/HowToReport/ucm053074.htm. Accessed October 18, 2017.
US Food and Drug Administration. Preventable adverse drug reactions: a focus on drug interactions. 2014d. Available at http://www.fda.gov/drugs/developmentapprovalprocess/developmentresources/druginteractionslabeling/ucm110632.htm#ADRs. Accessed October 18, 2017.
US Food and Drug Administration. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. 2015. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf. Accessed March 30, 2016.
US Food and Drug Administration. Questions and answers on FDA's Adverse Event Reporting System (FAERS). 2016. Available at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInforma tion/Surveillance/AdverseDrugEffects/default.htm. Accessed October 18, 2017.
US Food and Drug Administration. Code of federal regulations title 21. 2017. Food and Drug Administration. Available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=201.2. Accessed January 11, 2013.
VAERS. Vaccine adverse event reporting system. Centers for Disease Control and Prevention, US Food and Drug Administration. 2014. Available at https://vaers.hhs.gov/index. Accessed October 18, 2017.
Vermeer NS, Straus SM, Mantel-Teeuwisse AK, Domergue F, Egberts TC, Leufkens HG, De Bruin ML. Traceability of biopharmaceuticals in spontaneous reporting systems: a cross-sectional study in the FDA Adverse Event Reporting System (FAERS) and EudraVigilance databases. Drug Saf. 2013;36:617–25.
Vermeer NS, Ebbers HC, Straus SM, Leufkens HG, Egberts TC, De Bruin ML. The effect of exposure misclassification in spontaneous ADR reports on the time to detection of product-specific risks for biologicals: a simulation study. Pharmacoepidemiol Drug Saf. 2016;25:297–306.
Weiss AJ, Elixhauser A, Bae J, Encinosa W. Statistical Brief #158: origin of adverse drug events in U.S. hospitals, 2011. Agency for Healthcare Research and Quality. 2013. Available at www.hcup-us.ahrq.gov/reports/statbriefs/sb158.jsp. Accessed October 18, 2017.
Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014;89:1116–25.
World Health Organization. The importance of pharmacovigilance: safety monitoring of medicinal products. 2002. Available at http://apps.who.int/medicinedocs/pdf/s4893e/s4893e.pdf. Accessed August 24, 2017.
World Health Organization. 2006. INN for biosimilars. Available at http://www.who.int/medicines/services/inn/inn_bio_sim/en/. Accessed August 16, 2017.
World Health Organization. Joining the WHO Programme for International Drug Monitoring. 2010. Available at http://www.who.int/medicines/areas/quality_safety/safety_efficacy/Joining_the_WHO_Programme.pdf?ua=1. Accessed October 18, 2017.
World Health Organization. Biological qualifier: an INN proposal. 2015a. Available at http://www.who.int/medicines/services/inn/WHO_INN_BQ_proposal_2015.pdf?ua=1. Accessed October 18, 2017.
World Health Organization. WHO pharmacovigilance indicators. A practical manual for the assessment of pharmacovigilance systems. 2015b Available at http://www.who.int/medicines/areas/quality_safety/safety_efficacy/EMP_PV_Indicators_web_ready_v2.pdf. Accessed October 13, 2017.

World Health Organization. Pharmacovigilance. 2017. Available at http://www.who.int/medicines/areas/quality_safety/safety_efficacy/pharmvigi/en/. Accessed August 17, 2017.

Wu TY, Jen MH, Bottle A, Molokhia M, Aylin P, Bell D, Majeed A. Ten-year trends in hospital admissions for adverse drug reactions in England 1999-2009. J R Soc Med. 2010;103:239–50.

Yih WK, Kulldorff M, Fireman BH, Shui IM, Lewis EM, Klein NP, Baggs J, Weintraub ES, Belongia EA, Naleway A, Gee J, Platt R, Lieu TA. Active surveillance for adverse events: the experience of the vaccine safety datalink project. Pediatrics. 2011;127:S54–64.

Chapter 24
Development and Commercialization of Biosimilars in India: Current Regulatory and Clinical Experience



Anurag S. Rathore, Srishti Joshi, Ankita Bhargava, and Neh Nupur

Abstract The ever-increasing cost of healthcare together with our improved understanding of biotech therapeutic drugs has fueled the rise of biosimilars. A step towards achieving the successful development of a biosimilar is to establish analytical similarity with the innovator drug. This is necessary so as to avail the significant reduction in clinical data required for achieving regulatory approval. A key concern is the limited understanding of how the different quality attributes (QA) affect its safety and efficacy profile. India has successfully demonstrated its ability to make affordable, high-quality pharmaceutical products for the world, particularly the small molecule generics. This fact is validated by the trend that the share of Indian made pharmaceutical products in the US market has been constantly increasing and is presently more than 30%. The question is if India can successfully replicate its success, in manufacturing complex biotherapeutic products. This chapter explores India's journey in the field of biosimilar manufacturing with an emphasis on the regulatory aspect. Followed by a concise overview of the evolution of global regulatory guidelines, the Indian framework has been discussed in detail. Major changes introduced in the latest guidelines for similar biologics (2016) have been highlighted. Insight into the key developments related to clinical experiences and thereby addition of more sophisticated platforms to the analytical armory in the past decade for characterization of biosimilars has been given. Two recently published case studies on analytical platform approach used to establish similarity for microbial (GCSF) and mammalian product (Rituximab), in the Indian market-place, using an array of advanced, orthogonal, high-resolution analytical methods, have been discussed. Finally, the importance post-approval pharmacovigilance as a feedback mechanism to update and improve existing regulatory framework has been outlined.

A. S. Rathore (✉) · S. Joshi · A. Bhargava · N. Nupur
Department of Chemical Engineering, Indian Institute of Technology, New Delhi, India
e-mail: asrathore@biotechcmz.com; http://www.biotechcmz.com

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_24

Keywords Affordable biopharmaceutical · Analytical comparability · Biosimilar market · Clinical experience · Regulatory framework · Revised guidelines · Similar biologics

Introduction

Biotechnology has had a significant influence on all spheres of economic development and currently India's biotech sector is valued at over US$10 billion, with a compound annual growth rate (CAGR) of 20.0% over the last decade (Shaw 2016). India has long been recognized as a market leader in development and production of pharmaceutical products and now it is looking to replicate its success in manufacturing of biotech therapeutics. The biopharmaceuticals segment, which accounts for two-thirds of the Indian biotech sector's revenue, offers significant prospects for growth in the coming years. One estimate predicts that the Indian pharma market will grow at the CAGR of 22.4% to touch US$ 100 billion by 2025 (Swamilingiah 2014).

Biopharmaceuticals have been in the market for more than 20 years. A diverse set of products have been commercialized including hormones, immuno-modulators, monoclonal antibodies (mAbs), blood coagulation factors, enzymes and vaccines. Biosimilars are an important and relatively new category of biopharmaceuticals and are made using, or derived from, living organisms. It is estimated that from US$ 1.3 billion in 2013, the biosimilars market will reach US$ 24 billion by 2019 (Shaw 2016).

Indian patients have been using biosimilars of products such as insulin analogs and Filgrastim since early 2000s. More recently, biosimilars to complex products such as monoclonal antibodies such as Trastuzumab, Rituximab, and Adalimumab have also entered the market. The success with biosimilars is expected to continue and have a significant impact towards accessibility to affordable biopharmaceuticals to Indian as well as global patients (Shaw 2016).

A key objective during biosimilar development is to establish product comparability and to establish a 'scientific bridge' to the clinical experience of the reference product. This is achieved through an elaborate set of activities targeted to confirm that the active ingredient in the biosimilar matches the reference product (Berkowitz 2017). An extended introduction to biosimilar development has been covered in Chap. 1 of this book. Briefly, the biosimilar confirmatory steps include:

1. *Analytical characterization*: Since the basic premise of a biosimilar is that it is a copy of an existing biotech therapeutic, an extensive analytical characterization is necessary to establish comparability with respect to molecular structure and functionality. This is a key activity as it serves as a basis for the significantly reduced clinical program that is typical for biosimilars.
2. *Pre-clinical studies*: These studies are targeted to confirm that any differences between the reference product and biosimilar do not significantly impact its safety or efficacy.

3. *Pharmacokinetic (PK)/Pharmacodynamic (PD) studies*: PK and PD studies in humans are performed to establish bioequivalence.
4. *Clinical studies*: Clinical trials are conducted to confirm that the clinical safety and efficacy of the biosimilar is similar to that of the reference product (Tank Chintankumar et al. 2009; Singh Sekhon Vikrant Saluja and Singh Sekhon 2011; Kanase et al. 2013).

To appreciate the challenges associated with development, manufacturing and characterization of biosimilar products, it is important to understand how biosimilars differ from "generic" drugs (Table 24.1). Generic drugs are chemically synthesized small molecules with a single, stable, well-defined structure. In contrast, biosimilars are large biomolecules with higher order structures (Rathore 2009). These structural differences lead to distinct manufacturing practices and issues for each class of molecules. These have been covered in detail in Chap. 1.

Regulations governing development and approval of "biosimilars" have evolved with time, in part influenced by learnings from the clinical debacles with biotherapeutics like Eprex, where a minor change in formulation from human serum albumin to detergent polysorbate 80 led to a remarkable increase in cases of pure blood aplasia in patients receiving the drug (Schellekens and Jiskoot 2006). This incident highlighted the complex nature of biotherapeutics and their sensitivity towards seemingly slight alterations in the manufacturing process (Konara et al. 2016). Moreover, the last two decades have seen a continuous array of major advancements towards development of new analytical approaches and techniques that are capable of delivering an in-depth understanding of the physicochemical characteristics of the product, thereby offering an unprecedented insight into the structural comparability between the biosimilar and the innovator molecules (Schellekens and Moors 2010).

Table 24.1 Comparison between the small molecule generics and biosimilars

	Small molecule generic	Biosimilar
Structure	• Low molecular weight • Single, defined structure with minimal structural variability	• High molecular weight, complex molecules • Higher order structures present with considerable variability
Synthesis	• Well controlled, fewer chemical synthesis steps	• Biological systems (bacteria, yeast) dependent synthesis with limited control over the numerous enzymatic and chemical reactions occurring inside the cells
Stability	• Rigid structures leading to higher product stability	• Stability strongly dependent on preservation of higher order structure (pH, temperature, buffer composition)
Immunogenicity	• Low likelihood of degradation and/ or aggregation causing immunogenicity	• Sensitivity to degradation and aggregation of monomers can lead to immunogenic responses

Table 24.2 Prices of similar biologics compared to originator biologicals in India (Domestic Biologicals Cost Less in India 2016)

Biological	Discount to reference biological (%)	Average discount to brand-name biologicals (%)
Epoetin	12–40	25
Filgrastim	16–29	23
Insulin glargine	38–43	40
Interferon alfa-2a	16–56	36
Interferon alfa-2b	44–55	49
Pegfilgrastim	36–74	55
Rituximab	38–47	42

Global Biosimilar Market Overview

The biosimilar market is entering a golden era in terms of manufacturing and a major contributing factor to this is the upcoming patent cliffs to some of the blockbuster biotherapeutics (GaBI Journal Editor 2015). The lower cost of the biosimilars as compared to innovator products (Table 24.2) is expected to result in a much needed improvement with respect to affordability of these products (Domestic Biologicals Cost Less in India 2016).

The global market of biosimilars can be broadly classified as follows (Mordor Intelligence 2017):

1. *Type of products*: This category consists of recombinant non-glycosylated proteins, recombinant glycosylated proteins, and recombinant peptides. Of these, recombinant glycosylated products have the highest CAGR at present.
2. *Manufacturing*: Segmented into in-house manufacturing and contract manufacturing organizations, the latter is projected to grow at the highest CAGR due to the higher efficiency they offer with respect to facilities, equipment, and labour costs.
3. *Applications*: These can be segmented into oncology, blood disorders, chronic and auto-immune diseases, growth hormone deficiency, infectious diseases, and other applications. The oncology segment is projected to show the highest CAGR in near future.
4. *Region*: Geographically, the market for biosimilars is dominated by Europe, followed by Asia Pacific, North America, South America, Middle East, and Africa. However, growth in the European market is expected to have the highest CAGR driven by certain socio-economic and political factors that include growing pressure to curtail healthcare costs, patent expiry of biologic products, favourable view of the regulators, and an increasing comfort with the biosimilars on the market.

While US has recently proposed the regulatory pathway for approval of biosimilars, it lags behind Europe with respect to the number of products on the market (Christl 2016). Within Europe, Germany has the largest share, followed by other

European countries (UK, France, Spain, and Italy). (Biosimilar Development 2016; James 2016; Unnikrishnan 2016).

To remain competitive, manufacturers are developing global biosimilars strategies that factor in 'where to play' and 'how to win'. To achieve optimal growth, it is imperative that the biosimilar manufactures understand the future growth trajectory in terms of geographic location. At present, developed markets include United States (US), European Union (EU), and Japan and emerging markets including BRICS (Brazil, Russia, India, China and South Africa) and MIST (Mexico, Indonesia, South Korea and Turkey) countries (Biosimilars in Global Markets CPhI China 2017).

Indian Biosimilar Market Overview

In India, biosimilars offers a $240 billion global opportunity for the pharmaceutical industry. The domestic market itself is anticipated to grow from $300 million to $40 billion by 2030, thus capturing 20% share in the global market. Currently, the domestic sales are close to $250 million growing at CAGR of 14%, while exports contribute close to $51 million.(Indian Biosimilars Market May Reach $40 bn by 2030 2016).

Since Central Drug Standard Control Organization (CDSCO) introduced the revised biosimilar guidelines in 2016, several Indian pharma companies are making significant investments into biosimilar development and production into global markets. In an effort to align itself more with global standards of quality, safety and efficacy, one of the major changes in the new guideline is the requirement for post-marketing studies on approved biosimilars, with safety as the primary endpoint. Efforts have been made to bring Indian regulations at par with the global standards and pave the way for transition of the industry from a developing to a highly regulated market (Lo 2016).

In order to achieve global acceptance, many Indian pharmaceutical firms are investing and establishing biosimilar production facilities outside India. For instance, Dr. Reddy's Laboratories and Zydus Cadila have signed separate agreements with Turkish biopharma companies to establish partnerships that facilitate their entry into the local market. On the other hand, companies such as Cipla have made significant investments in India and outside India to acquire manufacturing facilities and potential product pipelines in the biosimilar segment in other countries (such as China) (Rao 2016) (Table 24.3).

As the experience of Indian pharmaceutical industry has shown, biosimilars pose a different set of challenges as compared to those presented by small molecule generics. Key differences include a more demanding clinical development program, complex manufacturing, and in general a greater set of risks (Kumar and Singh 2014; Singh Sekhon Vikrant Saluja and Singh Sekhon 2011).

Considering the challenges posed, India has already shown some remarkable efforts in meeting international biosimilar standards and the evolving regulatory system is expected to further streamline the process.

Table 24.3 Key Indian players with global partnerships in biosimilar manufacturing (Partnerships Driving Similar Biologics Development in India 2016)

Indian company	Partner	Product(s)	Type of deal
Strides Arcolab	Oncobiologics Inc. (USA)	Follow-on biologicals	Partnership
Cipla Ltd.	Mabpharm Pvt. Ltd (India)	Monoclonal antibodies	Majority acquisition
Serum Institute of India Pvt. Ltd.	NeuClone (Australia)	Monoclonal antibodies	Licensing agreement
IPCA Laboratories	Oncobiologics Inc. (USA)	Monoclonal antibodies	Partnership
Agila Biotech Pvt. Ltd.	Pfenex (USA)	Follow-on biologicals	Partnership
Biocon	Mylan (USA)	Insulin analogues	Partnership
Ranbaxy Laboratories	Epirus (USA)	Infliximab follow-on biologicals	Licensing agreement
Biocon	CCM Pharmaceuticals (Malaysia)	Insulin follow-on biologicals	Licensing agreement
Lupin Ltd.	NeuClone (Australia)	Cell line technology	Partnership
Intas Biopharmaceuticals	APOTEX (Canada)	GCSF and PEG-GCSF	Partnership
Inbiopro Solutions Pvt. Ltd.	Strides Arcolab (India)	Follow-on biologicals	Majority acquisition
Biocon	Pfizer Inc. (USA)	Insulin and analogues	Partnership
DM Corporation	ORF Genetics (Iceland)	Human recombinant therapeutic proteins and follow-on biologicals	Joint venture
Cipla Ltd.	BioMab (Hong Kong)	Monoclonal antibodies	Minority acquisition
Ranbaxy Laboratories Ltd.	Pfenex (USA)	Expression technology	Development agreement
Ranbaxy Laboratories Ltd.	Biovel Life Sciences (India)	Follow-on biologicals	Majority acquisition

Overview of Biosimilars Regulations

Evolution of Guidelines for Approval of Biosimilars

The European Medical Agency (EMA) was the first regulatory authority to propose biosimilar guidelines in 2005. The first approval came soon after in 2006 (European Medicines Agency 2006) (refer to Chap. 7). This was followed by the guidelines formulated by the Ministry for Health Labour and Welfare (MHLW) in Japan and issued in 2009 (Bennett et al. 2014).

Likewise, Korea's Ministry of Food and Drug Safety (MFDS), formerly the Korean Food and Drug Administration (KFDA), also issued their guideline to regulate biosimilar products in Korea in 2009. In the same year, Health Canada also finalized the draft subsequent entry biologics (SEBs) guidelines and these have been recently revised in 2016 (Welch 2016) (refer to Chap. 6).

In an effort to create international reference standards for biotherapeutics, the World Health Organisation (WHO) has collaborated with the international industry and science experts to design international regulatory standards in 2009. These have been approved worldwide to evaluate the quality, safety, and efficacy profile of the biosimilar drugs (WHO 2016).

The US Food and Drug Administration (US-FDA) proposed a bill that was signed on the 23rd March, 2010 for governing the regulation of biosimilars. The BPCI Act, 2009 permits the licensing of biological products that have been shown to be biosimilar to previously licensed reference products (Shuster and Farmer-Koppenol 2013) (refer to Chap. 5).

Indian biosimilar guidelines, "Guidelines on Similar Biologics (SB)" were put together by Department CDSCO. They were first introduced in 2012 to regulate the pathway of development, manufacturing and marketing of biosimilars in India. The revised guidelines were recently published in 2016 (DBT and CDSCO 2016).

Biosimilar Regulatory Framework in India

India first announced the release of draft regulatory guidelines for similar biologics at the BIO conference in Boston, USA, on 19 June 2012. These were then finalized and implemented on 15 September 2012. The guidelines outline a simple abridged procedure for evaluation of 'similar biologics' which have been approved and marketed in India, Europe or USA for more than four years (Nagarjunareddy and Brahmaiah 2017).

The CDSCO is the apex and national regulatory authority under the Government of India (GoI) and addresses issues and challenges for ensuring the safety and efficacy of similar biologics and establishing the appropriate regulatory pathway. Apart from CDSCO, there are other competent authorities which are involved in the approval process of biosimilars (Table 24.4) (Konangi et al. 2013; Malhotra et al. 2015).

Table 24.4 Key stakeholders in the Indian regulatory framework (DBT and CDSCO 2016)

S. no	Committee	Department	Responsibility
1.	Review Committee on Genetic Manipulation (RCGM)	Department of Biotechnology (DBT)—Ministry of Science and Technology	It regulates import, export, transfer, receive of genetically modified organisms (GMOs) and activities related to research, production, preclinical and other studies involving GMOs, as per the DBT guidelines
2.	Genetic Engineering Approval Committee (GEAC)	Ministry of Environment, Forest and Climate Change	It acts as a statutory body for review and approval of activities involving large-scale use of genetically engineered organisms and their products in research and development, industrial production, environmental release and field applications
3.	Institutional Biosafety Committee (IBSC)	Institutions/Organization	Reviews all recombinant research carried out by an organization, recommends emergency plan in case of large-scale operations, as and when required, which would be then approved by competent authorities and training of personnel on biosafety and instituting health monitoring program for laboratory personnel
4.	Drug Controller General of India (DCGI)	Ministry of Health and Family Welfare	Ensures product safety and efficacy as well as grants approval for clinical trials and marketing of biotech therapeutics
5.	Food and Drugs Control Administration (FDCA)	Ministry of Health and Family Welfare	Approves manufacturing plants and ensures adherence to current Good Manufacturing Practices (c-GMP)
6.	State Biotechnology Coordination Board (SBCB)	Ministry of Environment, Forest and Climate Change	Reviews periodically the safety and control measures in various institutions handling GMOs and act as nodal agency at state level to assess the damage, if any, due to release of GMOs and takes appropriate on site control measures
7.	District Level Committee (DLC)	Ministry of Environment, Forest and Climate Change	Monitors the safety regulations in installations and has power to inspect, investigate and report to the SBCB or the GEAC about compliance or non compliance of r-DNA guidelines or violations under Environment Protection Act (EPA). Acts as nodal agency at District level to assess the damage, if any, due to release of GMOs and to take on site control measures
8.	Recombinant DNA Advisory Committee (RDAC)	Department of Biotechnology (DBT)—Ministry of Science and Technology	Reviews developments in Biotechnology at national and international levels and recommend suitable and appropriate safety regulations for India in r-DNA research, use and applications
9.	Institutional Animal Ethics Committee (IAEC)	Institutions/Organization	Supervises, reviews, and evaluates all aspects of the animal care and facility periodically, approve research involving animals which conform to the various guidelines for use of animals in research and development

Guidelines on Similar Biologics 2016

Similar to the European and the US guidelines, India adopted the "sequential approach" to market biosimilars. The initial guidelines addressed the requirements related to manufacturing process and quality control (Kumar and Singh 2014). These were revised in 2016 jointly by CDSCO and DBT in consultation with stakeholder experts from industry associations, leading scientific institutions, and laboratories. The revision was performed with the objective to align with the ever-changing global standards and to streamline the regulatory process for the authorization of SBs in India. The revised guideline is based upon recent guidelines from EMA on biosimilar products (2014), as well as the current WHO guidelines on the quality, safety, and efficacy of biotherapeutic protein products prepared by r-DNA technology so as to ensure that the product meets acceptable levels of safety, efficacy, and quality (George 2016).

The focal changes introduced in the 2016 guidelines are highlighted below (DBT and CDSCO 2016):

(1) *Principles for development of SB*: Under section 6, the 2012 guidelines mentioned that abbreviated comparability will no longer be acceptable for quality components but now abbreviated data requirements are acceptable for preclinical and/or clinical components of the development program but not for the quality components (George 2016).

(2) *Selection of reference biologics (RB)*: Under section 6.1, in the revised guidelines, biologics approved/licensed and marketed in a member country of The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human use (ICH) countries like EU, USA, Japan, Canada, and Switzerland can be used as reference biologic in India. In the older guidelines, a biologic had to be either licensed in India or licensed outside India for at least 4 years, before it could be selected as a Reference Biologic (Kashyap 2016)

(3) *Manufacturing process*

Upstream process development: Along with the requirements outlined in the 2012 guidelines, the revised guidelines call for a detailed description of the upstream process, including media components that were utilized for cell growth.

(4) *Product characterization*

Biological Activity: Under section 6.3.2, the revised guidelines state that the biological product should be validated against an international or national reference standard where available and appropriate instead of calibrated as mentioned in older guidelines.

Specifications: Under section 6.3.3, the revised guidelines require acceptance limits based on RB data and data from sufficient number of batches from preclinical or clinical batches, which must be in line with international norms. In the 2012 guideline, the specifications were not mentioned, leaving scope for variability in the approaches used for setting these limits among manufacturers (George 2016).

(5) *Quality comparability study*: As per current guidelines, under section 6.4, it is important to demonstrate that the molecular structure of the active Drug Substance (DS) present in the SB is comparable to active DS present in RB. The drug attributes have been clearly defined. As part of biosimilar characterization, companies have to show a comparison between the innovator's product and the biosimilar with respect to the defined drug attributes using state of the art analytical techniques. Quality attributes (QA) and the techniques that can be employed are listed in Table 24.5 (George 2016).

From the point of view of establishing similarity, quality attributes of a SB have been considered in two categories; Critical Quality Attributes (CQAs) and Key Quality Attributes (KQAs) in the current guidelines. In the 2012 guidelines, there was no mention of KQA, and there was no segregation of the attributes. Those quality attributes which have direct impact on the clinical safety or efficacy have been covered under CQAs which must be controlled within limits that need to be established based on the RB.

(6) *Data requirements for preclinical studies*: Under section 7, as per current guidelines, an applicant may submit parallel application to RCGM and CDSCO (DCGI office) seeking approval to conduct clinical trial. This is more efficient as earlier applicants had to wait for the RCGM to review and approve the toxicity study reports, before seeking clinical trial approval from the DCGI.

    Toxicology studies: Under section 7.2.2 of the revised guidelines, one can do the study only in pharmacologically relevant animal species (i.e. rabbits). If animal species is not available for doing repeat dose toxicity studies then proper justification is required. If justified, then the toxicity studies need to be undertaken either in rodent or non-rodent species, as per requirements of Schedule Y with RCGM permission. In previous guidelines, studies were allowed to be undertaken in two species: one rodent and another non-rodent.

(7) *Data requirements for clinical trial (CT) applications*: Under section 8, as per current guidelines, for CT application, apart from other requirements, CQAs and KQAs of the product are to be included with respect to 'physicochemical and biological characterization' of the product. The quality data submitted should indicate that there are no differences in CQAs and that all KQAs are well controlled in order to allow the initiation of clinical evaluation. In older guidelines, it was mentioned that the quality data submitted should establish comparability of SB manufactured at clinical stage against RB.

    Confirmatory safety and efficacy study: Under section 8.3 of the revised guidelines, in phase 3, Clinical Trials (CTs) intended for seeking marketing approval of SB shall be conducted as single arm study in at least 100 evaluable subjects that may be carried out in most sensitive indication to address any residual uncertainty. The 2012 guidelines did not give an indication of the number of patients for a Phase 3 study and just mentioned that sample size should have statistical rationale.

    Waiver of confirmatory clinical safety and efficacy study: Under section 8.3.1, as per the proposed guidelines, a SB may be approved for all the

Table 24.5 Tools and techniques used in physiochemical characterization of biosimilars as per the Indian guideline for approval of biosimilar products (DBT and CDSCO 2016)

Critical quality attributes	Recommended analytical tools
Primary structure	LC-ESI-MS, MALDI-TOF-MS
Amino acid sequence	
Intact mass	
Reduced mass	
Secondary structure	UV-CD, FTIR, LC-MS, tryptic-peptide mapping
Higher order structure	Fluroscence, Near UV-CD, UV-VIS spectroscopy
Disulphide bond	
Charged variants	IEF/CE-IEF, IEX
Deamidated form	
Oxidized form	
Sialylated form	
Mass variants	SEC, RP-HPLC, IEX, Western Blot, SDS-PAGE, CE-SDS PAGE
Aggregates	
Truncated form	
Pegylation	
Post translational modification	HPLC-MALDI-TOF, IEX, N/C-terminal sequencing
Glycosylation	
Oligosaccharide patterns	
Biological activities	Cell based assay, Apoptosis assay, in-vivo bioassay, in-gel activity assay (for enzymes), Receptor Binding Bioassay (for MAbs)
ADCC	
ADCP	
CDC	
Apoptosis	
Antigen/antibody binding	
FcγR binding	
C1q binding	
FcRn binding	

CD circular dichroism, *FTIR* Fourier transform infra-red spectrometry, *LC* liquid chromatography, *MS* mass spectrometry, *IEX* ion exchange chromatography, *CE* capillary electrophoresis, *IEF* iso-electric focusing, *RP-HPLC* reverse phase high performance liquid chromatography, *SEC* size exclusion chromatography, *UV* ultra-violet, *SDS-PAGE* sodium dodecyl sulfate poly acrylamide gel electrophoresis, *TEM-* transmission electron microscopy, *ELISA* enzyme linked immuno sorbant assay

indications of approved reference product, if the safety and efficacy study has been waived for such SB based on comparable quality, non-clinical as well as convincing PK/PD data. If the phase III trial has to be waived for the SB, it would be mandatory to gather immunogenicity in the PK/PD study, and immunogenicity would be required to be generated during the post-approval Phase IV study.

Extrapolation of efficacy and safety data to other indications: Under section 8.5, both the proposed and older guidelines have the provision for approval of

SB for all the indications as that of the RB. However, the introduction of detailed description of "quality" and "clinical" similarity is an additional feature in the revised guidelines.

(8) *Post-marketing data for similar biologics (Phase IV study)*: Under section 10.3 of the new guidelines, in post marketing stage, additional safety data may be collected by conducting a pre-defined single arm study of more than 200 evaluable patients and compared with historical data of the RB product. The timeline to complete said study is also defined to be preferably within 2 years of marketing permission/manufacturing license. Companies will now have to make safety surveillance a priority and submit the Phase IV study protocol, meeting the study design and sample size requirements as part of the new Drug Application (NDA) dossier in order to be considered for approval. In previous guideline, phase IV studies did not require follow up by the manufacturers of SB. As part of post marketing, Periodic Safety Update Reports (PSURs) were being submitted as per the defined periodicity in the Schedule Y of the *Drugs and Cosmetics Act*, 1940.

Analytical Similarity of Biosimilar Products

The current practice in granting regulatory approval for biosimilars presides on the demonstration of similarity, starting with an extensive physicochemical and biological characterization and clinical studies (FDA 2015). In the past, regulatory demands for clinical studies have been challenged as being too cautious and hindering the development of subsequent generations of biosimilars (Schellekens and Moors 2010). The EMA has recently released a concept paper to revise the clinical requirements for granulocyte colony stimulating factor (G-CSF), thereby suggesting waiver of clinical studies (European Medicines Agency 2015). In line with EMA, the US FDA has released guidance of clinical pharmacological data to support demonstration of biosimilarity to a reference product, indicating the possibilities to perform only selected clinical studies when comparative analytical characterization indicates "highly similar proposed biosimilar with fingerprint-like similarity" (US Department of Health and Human Services 2016). During the 67th World Health Assembly, WHO agreed that the next similar biotherapeutic product (SBP) guideline should also include affordability as a major consideration for biosimilars, while still ensuring their quality, safety and efficacy (WHO 2016).

The Indian regulators have been quite supportive of biosimilars and have approved over 50 biosimilars to date, but the regulatory expectations do not include mandatory comparative testing for clinical trials and PK/PD studies and requires an in-depth physicochemical and biological characterization (Similar Biologics Approved and Marketed in India 2016). ICH Q5E and Q6B guidelines provide guidance on those physicochemical and biological features that should be considered in assessment of comparability (ICH Guidelines 2013; ICH Guidelines 2000).

Independent in-depth similarity studies are being conducted to demonstrate the capability of extensive analytical characterization in helping pick up subtle differences between the biosimilar and reference product and their potential impact on safety and efficacy of the biosimilar. Two such case studies have recently been published for biotherapeutic products that have a large number of commercial manufacturers. The first study presents data from analysis of biosimilar products for Filgrastim, a microbial product (non-glycosylated). The second study targets analysis of biosimilars of Rituximab, a mammalian product (glycosylated). Both the studies were designed to determine any disparity in protein structure and/or other physicochemical properties of the product as well as its bioactivity. As an outcome, a meaningful insight was gained into the quality of biotherapeutics present today in the Indian market, adding value to such biosimilar similarity exercises. Backed with scientific data, such studies also allow to explore the possibility to establish criteria for waiving confirmatory clinical trials thus making biosimilars more affordable.

Analytical Similarity of Filgrastim Biosimilars

Filgrastim or recombinant methionyl-granulocyte colony stimulating factor (rGCSF), significantly impacts proliferation and differentiation of cells of hematopoietic lineage. It reduces the incidence and duration of post-chemotherapy neutropenia in patients with non-myeloid malignancies and mobilization of hematopoietic progenitor cells in transplantation patients. It is an 18.8 kDa protein and consists of 175 amino acids with one free Cys residue (C17) and two disulphide bonds (Skrlin et al. 2010). Unlike natural GCSF molecule which is O-glycosylated at Thr-133, Filgrastim is manufactured in *E. coli* and has been reported to have similar potency as the former, demonstrating that glycosylation is not essential for its bioactivity (Carlsson et al. 2004). About eight Filgrastim biosimilars have received marketing authorization in India (GaBI Journal Editor 2015) and Fig. 24.1 summarizes the physicochemical and biological characterization results of five such biosimilar products along with the reference standard. A detailed study listing the analytical methods used and the results obtained has recently been published (Nupur et al. 2016).

Each sample of Filgrastim biosimilar and its reference standard was independently measured using reverse-phase chromatography electrospray ionization mass spectrometry (RPC-ESI-MS) for intact mass and peptide mapping confirming primary structure. Deconvoluted ESI mass spectra revealed that the molecular mass of filgrastim biosimilars ranged from 18798.6–18798.9 Da. The peptide mapping analysis using chymotrypsin alpha provided an absolute coverage with predicted modifications such as carbamylation and oxidation at M1, M122, M127 and M138. Far ultraviolet circular dichroism spectroscopy (far UV CD) for determining secondary structure showed presence of α-helix rich protein as indicated by two dips at 208 and 222 nm. Size-exclusion chromatography (SEC) assessed degradation profile and revealed that the biosimilars were highly pure with <2% aggregates.

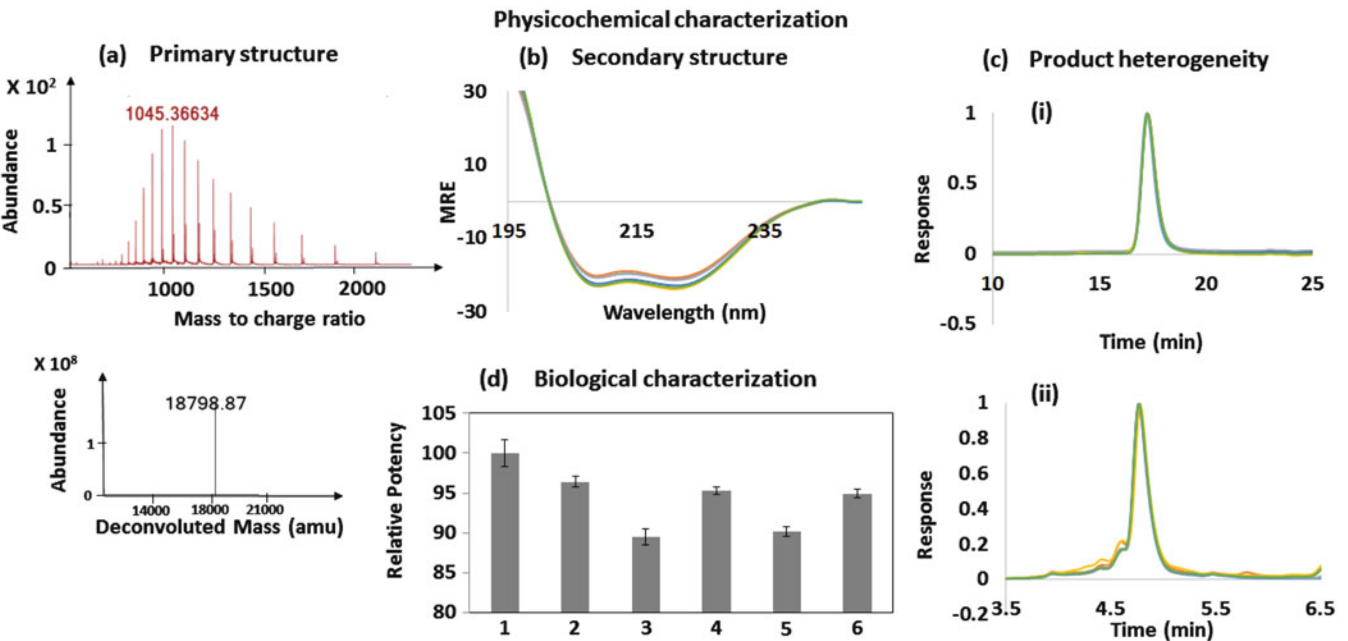


Fig. 24.1 Physicochemical and biological characterization of Filgrastim biosimilars and its reference standard (**a**) primary structure by mass spectrometry, (**b**) secondary structure by far Ultra-Violet Circular Dichroism, (**c**) product heterogeneity by (1) size exclusion chromatography and (2) reverse phase chromatography, (4) biological activity by in-vitro proliferation assay (Nupur et al. 2016)

Reverse-phase chromatography (RPC) assessed product related impurities profile and found low levels of oxidized/reduced variants. In-vitro cell proliferation assay assessed relative potency and induced similar degree of cell proliferation with relative potencies ranging between 85 and 100%.

The results offer robust evidence for the structural and biological similarity of Filgrastim biosimilars with its reference standard through a comprehensive analytical similarity exercise. The only category showing significant difference with respect to reference was charge variant, where 3 of the biosimilars showed presence of a variant previously reported as an oxidized variant (Holzmann et al. 2013; Joshi et al. 2015) Further investigation using ESI-MS revealed the occurrence of double charged envelopes. Moreover, RPC with fluorescence detection (RPC-FLD) showed multiple fluorescence peaks with a shift in the maximal wavelength (λ_{max}). This indicated that the observed anomaly was likely a conformational variant of GCSF and not an oxidized variant. Such type of assessment is required to determine whether a follow-on biologic is highly similar to reference product, and if a 'clinically meaningful' difference exists between the two.

Analytical Similarity of Rituximab Biosimilars

Rituximab is an anti-CD20 monoclonal antibody (mAb) that destroys B cells and is therefore approved to treat diseases which are characterized by excessive, overactive or dysfunctional B cells i.e. rheumatoid arthritis, leukaemia and non-Hodgkin's lymphoma. An IgG1k, Rituximab is a chimeric mAb produced in CHO cells, considered blockbuster drugs and a worldwide sales of USD 8.6 billion (Philippidis 2017). Rituximab comprises of 1328 amino acid (Flores-Ortiz et al. 2014), with mass of 144,244 Da for deglycosylated protein. Heavy chain (HC) comprise of 451 amino acids and light chain (LC) of 214 amino acids. HCs and LCs are linked by a single disulphide (S-S) bond and the HCs by two S-S bridges located in the hinge domain. Twelve additional cysteine bridges are intramolecular and delimit the six globular domains: one variable (V_L) and one constant for the LC (C_L); and, one variable (V_H) and three constant for the HCs (C_H1, C_H2 and C_H3) (Beck et al. 2013). It has several modification; i.e. N-glycosylation site within its Fc (Asn^{297}) region, N-terminal Gln to pGlu cyclization, and C-terminal Lys variant. The main mechanism of action of Rituximab comprises of binding of its Fab domain to CD20+ B-lymphocytes for induction of apoptosis by either CDC or ADCC (Hernandez et al. 2015). About eight Rituximab biosimilars have received marketing authorization in India (GaBI Journal Editor 2015) and Fig. 24.2 summarizes the physicochemical and biological characterization results of five such biosimilar products along with the reference standard. A detailed study listing the analytical methods used and the results obtained has been submitted elsewhere (Nupur et al. 2017).

Each sample of Rituximab biosimilar that was independently measured using reverse-phase chromatography electrospray ionization mass spectrometry (RPC-ESI-MS) for intact mass and peptide mapping revealed occurrence of different

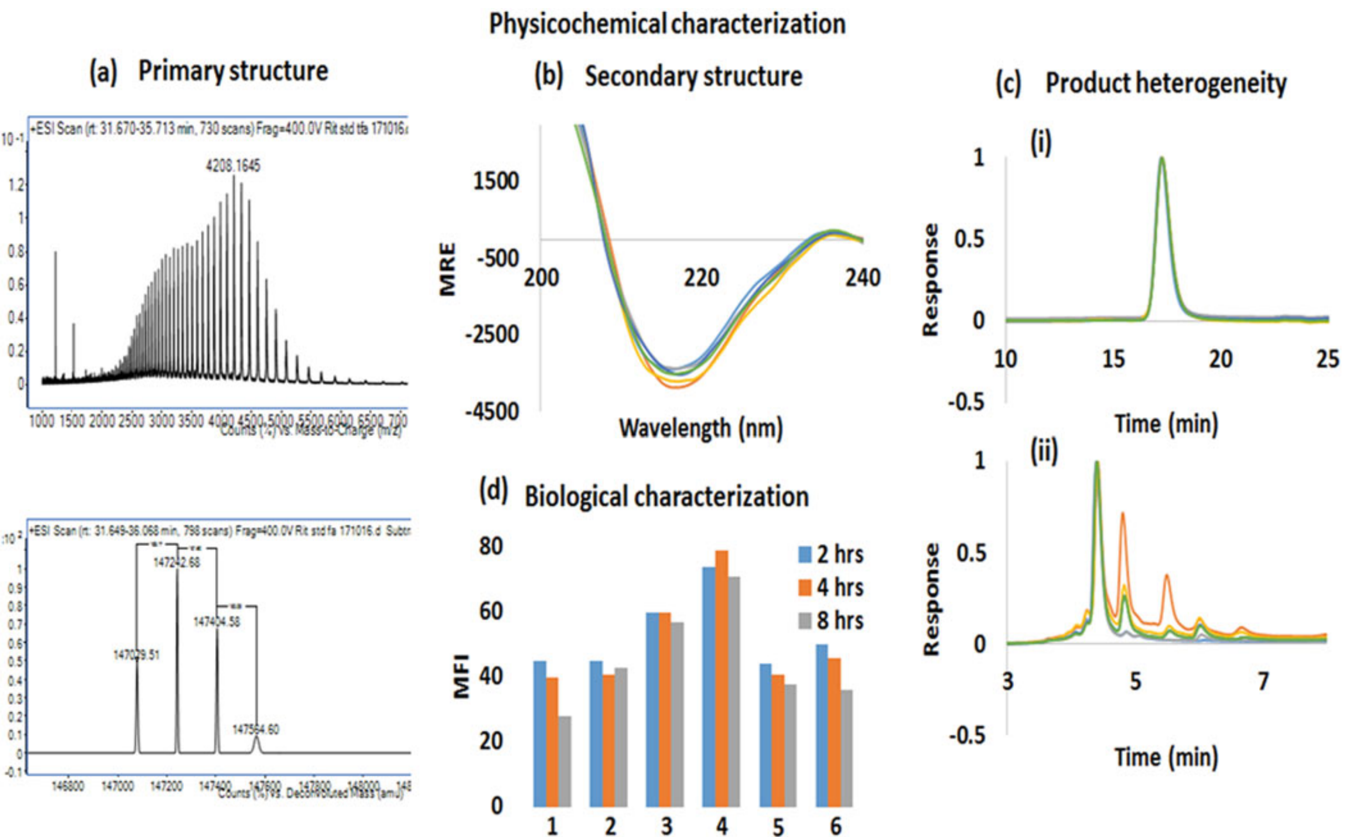


Fig. 24.2 Physicochemical and biological characterization of rituximab biosimilars and its reference standard (**a**) Primary structure by mass spectrometry, (**b**) Secondary structure by far ultraviolet circular dichroism, (**c**) Product heterogeneity by (1) size exclusion chromatography and (2) reverse phase chromatography, (**d**) Biological activity by in-vitro proliferation assay

molecular masses due to presence of different glycoforms. A sequence coverage of 98% and 96% was obtained for LC and HC respectively via trypsin digested peptide mapping. PTMs, such as N-terminal pyroglutamation, deamidation and oxidation at HC position M20, M34, M256, M432 and LC position B21 were observed. More than 10 glycan species were identified differing in the degree of galactosylation, afucosylation and sialylation among Rituximab biosimilars and reference standard. The far UV CD spectra determined the presence of antiparallel β-pleated sheet indicated by a characteristic negative dip at 216–218 nm. The second derivative of FTIR spectra also showed characteristic amide I and amide II bands resulting in quantification of α-helix, β-pleated sheet, β-turn and random coil with β-pleated sheet in majority. SEC showed degradation profile with monomer ranging from 98.1 to 99.6%. Cation-exchange chromatography (CEX) showed presence of varying N-terminal lysine residues amongst the biosimilars when compared to the reference product. Interestingly, the variability in ADCC assay was significantly higher than in CDC assay. Binding kinetics of Rituximab biosimilars to human FcRγIIIa was carried out using SPR and BLI. The equilibrium constant (Kd) was found within the same order of magnitude and thus, can be assumed to follow the same molecular basis for its modulatory function that leads to B-cell depletion.

The results offer robust evidence for the structural and biological similarity of Rituximab biosimilars with its reference standard through a comprehensive analytical similarity exercise. Few Rituximab biosimilars and reference standard have exhibited differences in their glycoforms, charge isoform and biological properties. The major difference that was observed was with respect to the N-glycan structure. Low galactosylation, high mannose, high sialylated and afucosylated glycans were present in few biosimilars. This might directly impact safety and efficacy of the product but such differences in quality attributes are unavoidable and in line with the "similar but not identical" paradigm. The regular physicochemical and biological assessment remain crucial to identifying whether deviation in these products should require additional data on the preclinical and clinical level to determine a follow-on biologic as highly similar to reference product, so that a 'clinically meaningful' difference between the two does not exist.

Clinical Experience

Clinical development of biologic products is characterized by its regulatory complexity as well as considerable challenges unseen in traditional product development (Declerck 2012).

Clinical trials for biological products require targeted experience and knowledge. Immunogenicity, rare adverse events, and efficacy are all factors to take into consideration when developing a biosimilar. Biosimilar trials require a customized approach based on the therapeutic indication and study specific goals (US Department of Health and Human Services 2016). Some of the approaches for development of an effective plan and executive strategy for these studies are listed in Table 24.6

Table 24.6 Summary of approaches for establishing clinical similarity (Bennett et al. 2014; Bui et al. 2015; Kanase et al. 2013)

S. No	Approach	Description
1.	Distinctive recruitment strategies	Success in the development of biosimilar relies on the ability to clearly demonstrate PK similarity in a single dose clinical trial. Biologics trials demand high vigilance for safety issues, as well as efficient recruitment of healthy volunteer patients, as it targets specific diseases
2.	Understanding regulatory complexities	Companies should be flexible in adapting to the continuously evolving international guidelines
3.	Dedicated research physicians	Researchers, principal investigators, and scientists should be involved in all aspects of clinical trials, ensuring that proper medical and technical procedures are completed according to the guidelines
4.	Specialized pharmacy	Facilities should be designed to support biosimilar development
5.	Bioanalytical support	It is important to use the appropriate bioanalytical methods to evaluate PK and PD of a proposed biosimilar product and its reference product while performing an evaluation of clinical pharmacology similarity

Even post attaining licence, biotherapeutics require continued pharmacovigilance to monitor any adverse reactions to the drug being administered. This not only increases the accountability of the pharmaceutical companies, but also helps in monitoring sustainable product quality. India's revised guidelines take into consideration the importance of post-marketing evaluation and this has helped mitigate any adverse immunogenicity issues in certain cases. One such recent case was with the product Razumab (Ranizumab biosimilar) from Intas Pharmaceuticals. The product was approved by CDSCO on 19 June 2015 (Similar Biologics Approved and Marketed in India 2012) In April 2017, patients complained of eye inflammation post administration of Razumab injections. Out of the total 182 vials used of a particular batch (Batch no: 18020020), 11 instances of eye inflammation were reported and this galvanized Intas to recall that particular batch from the market (Patel 2017). Thereafter the product has been withdrawn from the market all together.

In another case, CDSCO issued a warning in January 2016 against the use of Roche's blockbuster drug Avastin in treating ophthalmologic conditions. Avastin (Bevacizumab) is an angiogenesis inhibitor, approved for treating several classes of cancer but does not have the FDA approval for treating ophthalmologic conditions such as Age related Macular Degeneration (AMD). Fifteen cases of damaged vision were reported in a centre in Gujarat upon administration of Avastin for AMD (Brennan 2016). The continued success of India's evolving biosimilar regulatory system lies in employing mechanisms for early spotting of such rare cases and taking action before further damage can occur.

Conclusions

Development of a biosimilar product is a cumbersome process requiring strategic planning and monitoring of the whole enterprise from inception to its approval, as well as post approval in the form of pharmacovigilance. All of the aspects discussed in this chapter are significant and cannot be omitted in an effort to save the associated time and cost. However, investing in certain aspects in greater detail, such as physicochemical characterization and biological characterization provides useful and specific insights regarding not only the complexity of the product but also the interactions between the process and the product quality.

Due to the limited clinical database at the time of approval of a biosimilar, vigorous pharmacovigilance is required post approval. What is needed is continued advancements in our ability to perform an extensive comparison of structural and functional characteristics of biosimilars and reference products so as to reduce (not eliminate) our reliance on clinical studies. India is one of the leading contributors in the world biosimilar market with market authorization for over 50 biosimilars and follow a step-wise similarity approach, including the need to demonstrate similar quality during manufacturing and non-clinical testing (Similar Biologics Approved and Marketed in India 2016).

Although in the case studies discussed above, the physicochemical and biological characterization results confirm the analytical similarity of biosimilars on the market with respect to the reference standard, noticeable differences were observed amongst them for the two cases discussed. However, these analytical differences did not translate into meaningful changes in the formulation. The revised Indian regulatory guidelines emphasize on use of state of the art analytical techniques for this purpose.

Acknowledgements This work was funded by the DBT Centre of Excellence for Biopharmaceutical Technology grant under Department of Biotechnology, Government of India (BT/COE/34/SP15097/2015).

References

Beck A, Diemer H, Ayoub D, Debaene F, Wagner-Rousset E, Carapito C, Van Dorsselaer A, Sanglier-Cianférani S. Analytical characterization of biosimilar antibodies and Fc-fusion proteins. Trends Anal Chem. 2013;48:81–95. https://doi.org/10.1016/j.trac.2013.02.014.

Bennett CL, Chen B, Wyatt MD, Schulz RM, Georgantopoulos P, Kessler S, Qureshi ZP, Lu ZK, Love BL, Noxon V, Bian J, Ray P, Ablin RJ, Hruskesky WJ, Hermanson T, Raisch DW, Bobolts L, Armitage M, Hrushesky WJ, Macdougall IC, Sartor O, Armitage JO. Regulatory and clinical considerations for biosimilar oncology drugs. Lancet Oncol. 2014;15(13):e594–605. https://doi.org/10.1016/S1470-2045(14)70365-1.

Berkowitz SA. Biosimilar drug product development. Boca Raton: CRC Press; 2017. p. 15–82. https://doi.org/10.1201/9781315119878-3.

Biosimilar Development. Global biosimilars follow-on-biologics market 2016 industry demand insight profile and research to 2020. 2016.

Biosimilars in Global Markets CPhI China. 2017. Available at http://www.cphi.com/china/visit/news-and-updates/biosimilars-global-markets.

Brennan Z. India's CDSCO warns against using Roche's Avastin as eye treatment, regulatory affairs professional society. 2016. Available at http://206.79.211.135/Regulatory-Focus/News/2016/01/21/23941/India's-CDSCO-Warns-Against-Using-Roche's-Avastin-as-Eye-Treatment/#.

Bui LA, Hurst S, Finch GL, Ingram B, Jacobs IA, Kirchhoff CF, Ng CK, Ryan AM. Key considerations in the preclinical development of biosimilars. Drug discovery today. 2015; 20:3–15. http://dx.doi.org/10.1016/j.drudis.2015.03.011.

Carlsson G, Ahlin A, Dahllof G, Elinder G, Henter J-I, Palmblad J. Efficacy and safety of two different rG-CSF preparations in the treatment of patients with severe congenital neutropenia. Br J Haematol. 2004;126(1):127–32. https://doi.org/10.1111/j.1365-2141.2004.05008.x.

Christl L. FDA's Overview of the regulatory guidance for the development and approval of biosimilar products in the US. 2016. Available at http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM431118.pdf.

DBT and CDSCO. Guidelines on similar biologics 2016 guidelines on similar biologics: regulatory requirements for markeng authorizaon in India. 2016.

Declerck PJ. Biologicals and biosimilars: a review of the science and its implications. GaBI J. 2012;1(1):13–6. https://doi.org/10.5639/gabij.2012.0101.005.

Domestic Biologicals Cost Less in India. Generics and Biosimilars Initiative Journal. 2016. Available at http://www.gabionline.net/Biosimilars/Research/Domestic-biologicals-cost-less-in-India. Accessed 21 June 2017.

European Medicines Agency. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues. 2006. pp. 1–8. doi: https://doi.org/10.1136/bmj.333.7574.873-a.

European Medicines Agency. Guideline on non-clinical and clinical development of similar biological medicinal products containing recombinant human insulin and insulin analogues. EMEA. 2015;44:1–7.

FDA. Scientific considerations in demonstrating biosimilarity to a reference product guidance for industry scientific considerations in demonstrating biosimilarity to a reference product. Guidance for industry (February). 2015.

Flores-Ortiz LF, Campos-García VR, Perdomo-Abúndez FC, Pérez NO, Medina-Rivero E. Physicochemical properties of rituximab. J Liq Chromatogr Relat Technol. 2014;37(10):1438–52. https://doi.org/10.1080/10826076.2013.794738.

GaBI Journal Editor. Patent expiry dates for best-selling biologicals. GaBI J. 2015;4(4):178–9. https://doi.org/10.5639/gabij.2015.0404.040.

George B. Changes in regulatory requirements for biosimilar development in India. 2016. Available at https://www.biosimilardevelopment.com/doc/changes-in-regulatory-requirements-for-biosimilar-development-in-india-0001.

Hernandez MP, Lopez-Moralez CA, Ramirez-Ibanez ND, Pina-Laura N, Perez NO, Molina-Perez A, Revilla-Beltri J, Flores-Ortiz LF, Medina-Rivero E. Assessment of physicochemical properties of rituximab related to its immunomodulatory activity. J Immunol Res. 2015;2015:910763. https://doi.org/10.1155/2015/910763.

Holzmann J, Hausberger A, Rupprechter A, Toll H. Top-down MS for rapid methionine oxidation site assignment in filgrastim. Anal Bioanal Chem. 2013;405(21):6667–74. https://doi.org/10.1007/s00216-013-7138-0.

ICH Guidelines. Specifications: test procedures and acceptance critreia for biotechnological/biological products Q6B (March). 2000. pp. 1–20.

ICH Guidelines. Comparability of biotechnological/biological products subject to changes in their manufacturing process Q5E (March). 2013. pp. 1–20.

Indian Biosimilars Market May Reach $40 bn by 2030. The Economic Times. 2016. Available at http://economictimes.indiatimes.com/industry/healthcare/biotech/pharmaceuticals/indian-biosimilars-market-may-reach-40-bn-by-2030-report/articleshow/54717517.cms.

James S. Biosimilars market size is projected to reach $41.7 billion, demand in various end-use industries till 2024: Grand View Research, Inc; 2016. Available at https://globenewswire.com/news-release/2016/07/19/857002/0/en/Biosimilars-Market-Size-Is-Projected-To-Reach-41-7-Billion-Demand-In-Various-End-use-Industries-Till-2024-Grand-View-Research-Inc.html.

Joshi VS, Kumar V, Rathore AS. 'Role of organic modifier and gradient shape in rp-hplc separation: analysis of gcsf variants. J Chromatogr Sci. 2015;53(3):417–23. https://doi.org/10.1093/chromsci/bmu222.

Kanase SJ, Gavhane YN, Khandekar A, Gurav AS, Yadav AV. Biosimilar: an overview. Int J Pharm Sci Res. 2013;4(6):2132–44. https://doi.org/10.13040/IJPSR.0975-8232.4(6).2132-44.

Kashyap V. Highlights of proposed biosimilar guidelines, LexOrbis. 2016. Available at https://www.lexorbis.com/highlights-of-proposed-biosimilar-guidelines-2016/. Accessed 21 June 2017.

Konangi S, Raviteja MN, Vishal GN. Comparison of global regulatory approvals for biosimilar products. Int J PharmTech Res. 2013;5(3):924–35.

Konara CS, Barnard RT, Hine D, Siegel E, Ferro V. The tortoise and the hare: evolving regulatory landscapes for biosimilars. Trends Biotechnol. 2016;34(1):70–83. https://doi.org/10.1016/j.tibtech.2015.10.009.

Kumar R, Singh J. 'Biosimilar drugs: current status. J Appl Basic Med Res. 2014;4(2):63–5. https://doi.org/10.4103/2229-516x.136774.

Lo C. Going global: India's big biosimilar move - Pharmaceutical Technology. 2016.

Malhotra H, Krishnan A, Mody R. Global regulatory landscape of biosimilars: emerging and established market perspectives. Biosimilars. 2015;2015:19–32. https://doi.org/10.2147/BS.S44052.

Mordor Intelligence. Global biosimilar market|industry report|outlook|analysis|size and forecast. 2017.

Nagarjunareddy D, Brahmaiah B. Regulatory requirements of "similar biologics" for marketing authorization in India. Int J Drug Regul Aff. 2017;51:20–4. https://doi.org/10.22270/ijdra.v5i1.193.

Nupur N, Singh SK, Narula G, Rathore AS. Assessing analytical comparability of biosimilars: GCSF as a case study. J Chromatogr B Analyt Technol Biomed Life Sci. 2016;1032:165–71. https://doi.org/10.1016/j.jchromb.2016.05.027.

Nupur N, Dash R, Chhabra N, Rathore A. Assessing analytical comparability of biosimilars: rituximab as a case study. MAbs. 2017;10(1):143–58.

Partnerships Driving Similar Biologics Development in India. Generics and Biosimilars Initiative Journal. 2016. Available at http://www.gabionline.net/Biosimilars/Research/Partnerships-driving-similar-biologics-development-in-India. Accessed 21 June 2017.

Patel D. Patients report eye inflammation, Intas recalls its drug. The Indian Express. 2017. Available at https://www.pressreader.com/india/the-indian-express/20170403/282063391812417. Accessed 18 June 2017.

Philippidis A. The top 15 best-selling drugs of 2016|the lists|GEN, Genetic Engineering and Biotechnology News. 2017. Available at http://www.genengnews.com/the-lists/the-top-15-best-selling-drugs-of-2016/77900868. Accessed 21 June 2017.

Rao R. Indian companies gearing up to solve the biosimilars puzzle|Business Standard News. 2016.

Rathore AS. Follow-on protein products: scientific issues, developments and challenges. Trends Biotechnol. 2009;27(12):698–705. https://doi.org/10.1016/j.tibtech.2009.09.004.

Schellekens H, Jiskoot W. Eprex-associated pure red cell aplasia and leachates. Nat Biotechnol. 2006;24(6):613–4.

Schellekens H, Moors E. Clinical comparability and European biosimilar regulations. Nat Biotechnol. 2010;28(1):28–31. https://doi.org/10.1038/nbt0110-28.

Shaw KM. Biotechnology: helping India innovate for the future. Employment News, February. 2016. p. 1.

Shuster M, Farmer-Koppenol P. A Comparison of US and EU biosimilars regimes. Fenmick and West LLP; 2013. pp. 3–4.

Similar Biologics Approved and Marketed in India. Generics and Biosimilars Initiative Journal. 2012. Available at http://www.gabionline.net/Biosimilars/General/Similar-biologics-approved-and-marketed-in-India. Accessed 21 June 2017.

Similar Biologics Approved and Marketed in India. Gernerics and Biosimilars Initiative. 2016. Available at http://www.gabionline.net/Biosimilars/General/Similar-biologics-approved-and-marketed-in-India. Accessed 18 June 2017.

Singh Sekhon Vikrant Saluja B, Singh Sekhon B. Biosimilars: an overview. Biosimilars. 2011;1:1–11. https://doi.org/10.2147/BS.S16120.

Skrlin A, Radic I, Vuletic M, Schwinke D, Runac D, Kusalic T, Paskvan I, Krsic M, Bratos M, Marinc S. Comparison of the physicochemical properties of a biosimilar filgrastim with those of reference filgrastim. Biologicals. 2010;38(5):557–66. https://doi.org/10.1016/j.biologicals.2010.05.002.

Swamilingiah G. Insight: the biopharmaceuticals ecosystem in India and opportunities for European Collaboration, European Business and Technology Centre. 2014. Available at http://ebtc.eu/index.php/communication/ebtc-news/195-biotechnology-news/367-insight-the-biopharmaceuticals-ecosystem-in-india-and-opportunities-for-european-collaboration. Accessed 17 June 2017.

Tank Chintankumar J, Borkhataria Chetan H, Baria Ashok H, Patel Rakesh P, Tamizharasi S, Sureja DK, Patel SD, Parmar GR. Formulation and evaluation of aceclofenac loaded maltodextrin based proniosome. Int J ChemTech Res. 2009;1(3):567–73. https://doi.org/10.1056/NEJM197302222880814.

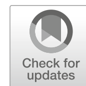
Unnikrishnan C. Biosimilar market to cross 40 billion in India by 2030 - BW Businessworld. 2016. Available at http://businessworld.in/article/Biosimilar-Market-To-Cross-40-Billion-In-India-By-2030/06-10-2016-106686/.

US Department of Health and Human Services. Clinical pharmacology data to support a demonstration of biosimilarity to a reference product. Guidance for Industry, (December). 2016. pp. 1–18.

Welch AR. Health Canada's unique approach to biosimilar regulation. 2016.

WHO. Guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (SBPs). World Health Organization (October). 2016. pp. 17–21.

Chapter 25
Immunogenicity and Adverse Reactions to Biosimilar Erythropoietin Products in Thailand: The Significance of Science and Quality Driven Process for Approval



Kearkiat Praditpornsilpa, Anunchai Assawamakin, and Kriang Tungsanga

Abstract The immunogenicity of recombinant human erythropoietin (r-HuEpo) by subcutaneous exposure has been well described. This adverse immunological effect causes anti-r-HuEpo-associated pure red cell aplasia (PRCA). There have been increasing cases of anti-r-HuEpo-associated PRCA after subcutaneous exposure in Thailand. The casual mechanism of this disease may relate to HLA immunogenic, protein aggregation, stability during storage and handling of drug products, formulation and drug product quality.

The r-HuEpos have been licensed for treatment of renal anemia in Thailand include innovator products and more than 20 biosimilar products. Lack of a scientific product characterization and quality driven process to approve such biosimilars may lead to different immunogenicity and safety profiles. The Prospective Immunogenicity Surveillance Registry of r-HuEpo with Subcutaneous Exposure in Thailand estimated the incidence of anti-r-HuEpo associated PRCA among subjects who had subcutaneous exposure to any r-HuEpo product currently available in Thailand, addressed the risk of anti-r-HuEpo-associated PRCA, and the association of product qualities towards adverse immunogenicity.

The experiences of biosimilar r-HuEpo in Thailand and anti-rHuEpo associated PRCA in Thailand illustrated the need to evaluate biosimilar product on a case-by case basis. Considering patients 'safety as the first priority, the approval process for biosimilar drug licensing should be designed to assess quality, characterization and impurity profile and comprehensive evaluation of the non-clinical and clinical

K. Praditpornsilpa (✉) · K. Tungsanga
Division of Nephrology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, King Chulalongkorn Memorial Hospital, Bangkok, Thailand
e-mail: kearkiat_p@md.chula.ac.th

A. Assawamakin
Department of Pharmacology, Faculty of Pharmaceutical Science, Mahidol University, Bangkok, Thailand
e-mail: anunchai.asa@mahidol.edu

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6_25

aspects. Pharmacovigilance could be the final step to narrow the gap between quality and safety and can be assessed in totality with cost-effectiveness and patient benefit that such drug offer.

Keywords Immunogenicity of recombinant human erythropoietin (r-HuEpo) · Renal anemia · Pure red cell aplasia (PRCA)

Introduction

Advances in biotherapeutic discovery and development have ultimately resulted in several breakthrough medicines and novel treatment options. Since the first recombinant-DNA-derived drug human insulin was approved for clinical use in 1982, biotherapeutics have been the most expanding drug class and have revolutionized the treatment of variety of diseases over the past decade. This breakthrough brings better clinical outcomes and quality of life for patients worldwide in many fields such as nephrology, rheumatology, organ transplantation, oncology, immunology, and orphan genetic diseases. There are currently over 200 approved biotherapeutic drugs on the market (Walsh 2014) and hundreds of potential products in clinical trials or awaiting regulatory approval. In addition, biotherapeutic drugs have a significantly higher likelihood of being a breakthrough, novel therapy compared with small molecule drugs since they are usually proteins, DNA, RNA or live cells whose mechanism of action in an underlying disease pathophysiology is usually better understood and characterized than small molecule drugs which typically are targeted inhibitors of certain physiological enzymes.

Development and manufacture of biotherapeutics has been a challenge for both the pharmaceutical industry and regulatory agencies (Dorey 2014). These challenges arise from the complexity of biotherapeutic molecules and the production processes that include production in a living cell, cell culture systems/plants, elaborate upstream and downstream manufacturing processes. Since biotherapeutics are made in cells and living organisms, they also present challenges in several other areas such as glycosylation and post-translational modifications, stabilization and formulation, biochemical modification and breakdown products of proteins (protein aggregation and fragmentation). Moreover, the manufacturing, packaging, shipping and handling procedures of biotherapeutic products also should be controlled and monitored since these products are often temperature sensitive and almost always require a cold chain. Breakdown products of biotherapeutics are known to present a safety and efficacy uncertainty and challenge. In this regards, biotherapeutics must be well characterized their potency, identity, quality, purity, and stability be thoroughly characterized and monitored throughout the process of licensing approval and post-licensing/marketing.

With the loss of biotherapeutics patent protection, there has been an upsurge of more affordable biosimilars, biocopy or follow-on-biologics agents, and increasing patient access to therapeutic agents. The affordable price of such biosimilar drugs

has been definitely favorable for health economics particuilarly in low- and middle-income countries (Blackstone and Fuhr 2013). However, the complexity of the manufacturing process for these recombinant proteins, if not being appropriately regulated and monitored or bypassed, can result in altered properties or quality that may limit drug efficacy and harmfully affect patient safety . For the remainder of the chapter we will be occasionally using the term “biocopy”. For the sake of better understanding of readers, its necessary to define biocopy drugs and highlight the key differences from biosimilars. The varying policies of regulatory agencies in different countries to regulate a biocopy/biosimilar can often create confusion with regards to definition and terminology. As a matter of fact, a biosimilar in one country may not be biosimilar but merely biocopy or even bioquestionable in the other country. Biocopy products can be designated biosimilar only if the analytical, clinical and non-clinical data demonstrate similarity of active proteins and other components of the products and have undergone the step-wise science based similarity assessment as described in earlier chapters of this book. Further it is also expected that biosimilars are thoroughly reviewed for their quality, safety and efficacy and appropriately regulated by health agencies.

The complex structure of biotherapeutic agents raise concerns for immunogenicity. Injection of a recombinant protein, especially repeated injection can have a vaccine like effect. In vaccinology, modification of the native protein or adding an adjuvant results in an effect, which increase the immunogenicity of vaccine. On the contrary, the impurities in biotherapeutic agents can have an unwanted adjuvant effect and cause adverse immunologic reactions. The impurity or impaired quality can lead to the generation of anti-drug antibodies and cause loss of efficacy. The anti-drug antibody causes patients’ morbidity and mortality by its neutralization effect. This chapter focuses mainly on the immunogenicity aspects related to biosimilar recombinant human erythropoietin (r-HuEpo) and rationalizes the need for a robust product quality and regulation program for such biosimilar HuEpo products particularly in developing nations. The chapter also discusses how the strategic approach focused on analytical characterization, similarity and product quality in the licensing process and post-marketing pharmacovigilance will help to wisely develop affordable biosimilars. This strategic approach for biosimilars will lead to increase access to drugs with a favorable health economics and patient safety in low- and middle-incomes countries.

Immunogenicity of Recombinant Human Erythropoietin

Since late 1980s, r-HuEpo has been used for treatment of anemia of chronic kidney disease. Correction of anemia improve cardiovascular as well as non-cardiovascular morbidity and mortality. Despite these beneficial effects of r-HuEpo, some chronic kidney disease patients who have previously or are currently using r-HuEpo have been reported adverse drug reaction of pure red cell aplasia (PRCA) (Peces et al. 1996; Prabhakar and Muhlfelder 1997; Lacombe 1996; Bennett et al. 2004) These

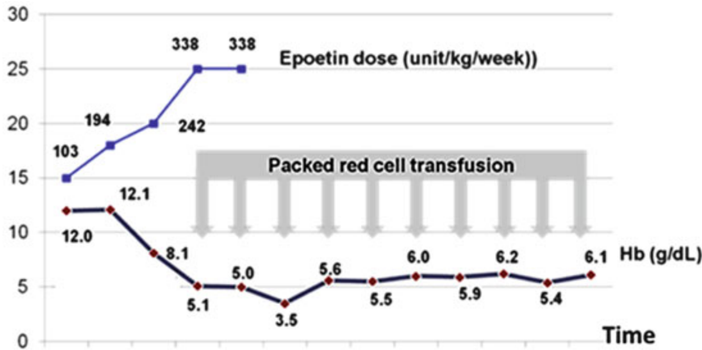


Fig. 25.1 A typical anti-r-HuEpo associated PRCA patient, developed sudden loss of r-HuEpo efficacy despite increase of r-HuEpo dose, resulted frequent blood transfusion dependent to relieve anemic symptoms

patients developed an unexplained sudden decrease in their hemoglobin (Hb) level. Anti-r-HuEpo antibody, which has been demonstrated in several studies, to be the proximate cause of the PRCA (Casadevall 2002; Casadevall et al. 2002). Typically, the anti-r-HuEpo associated PRCA patients had initially satisfactory responded to r-HuEpo and subsequently developed sudden loss of efficacy, refractory unexplained anemia despite an increase of r-HuEpo dose (Fig. 25.1). Reticulocytopenia was incongruent with the degree of anemia (reticulocyte count <10,000 cells/mm^3) and can be used to indicate suspicious cases. The bone marrow examinations demonstrated PRCA based upon the absence of erythroid precursor in an otherwise normocellular marrow.

Several platforms have been used to detect anti-r-HuEpo antibody namely radio-immunoprecipitation test (RIP) using radiolabelled Epo, bridging ELISA, BIAcore® surface plasmon resonance assay and bioassay (Urra et al. 1997; McLeod et al. 1974). The anti-r-HuEpo antibodies are directed against the peptide of r-HuEpo molecule rather than the carbohydrate moiety and cause loss of r-HuEpo efficacy by the neutralizing effect. Patients have severe anemia as the anti-r-HuEpo antibody cross reactivity to native erythropoietin by the homology of the molecule, blocks the interaction of both r-HuEpo and endogenous erythropoietin with the Epo receptor.

Anti-r-HuEpo associated PRCA patients suffer from anemic symptom and require frequent blood transfusion to alleviate severity. Patients may require one unit of packed red cell transfusion every one or 2 weeks, having risk of iron overload, hemochromatosis, alloantibody to red blood cell and transfusion related infection.

The effective treatment for anti-r-HuEpo associated PRCA has been kidney transplantation. Case report demonstrated that none of the patients required blood transfusion post-transplantation and the recovery of refractory anemia can be prompt after the recovery of kidney allograft function (Praditpornsilpa et al. 2005). Besides renal transplantation, the optimal therapy for anti-r-HuEpo associated PRCA is uncertain; previous data have shown that failures are common.

Steroids, cytotoxic agents, intravenous immunoglobulin (IVIG) or cyclosporine yields response rates ranging from 30 to 55% (Erslev 1995). A case report showed that rituximab (4 cycles of 375 mg/m^2/week) successfully recovered anti-r-HuEpo associated PRCA (Mandreoli et al. 2004).

There are no reports of spontaneous recovery of anti-r-HuEpo associated PRCA. After the reversal of anti-r-HuEpo antibody, the patients continue to have renal anemia and require long-term blood transfusion, albeit less frequently than when the antibody is positive. There was a case report intravenous re-challenge of r-HuEpo after the reversal of r-HuEpo antibody. After re-challenging, all patients responded to r-HuEpo, target level of Hb was maintained, blood transfusion was not required, and anti-r-HuEpo was consistently negative which suggest that re-challenge with intravenous r-HuEpo can successfully treat renal anemia after the reversal of anti-r-HuEpo antibody (Praditpornsilpa et al. 2014).

Risk Factors of Immunogenicity

Established risk factors for anti-r-HuEpo-associated PRCA include subcutaneous administration (Macdougall et al. 2012), immunogenetic and product qualities (Porter 2001). Intravenous administration association with r-HuEpo immunogenicity has been rarely observed (Shimizu et al. 2011). Compared to the circulating blood, antigen-presenting dendritic cells are abundant at skin and subcutaneous tissue, these professional antigen presenting cells encounter with naive T cells and migrate to draining lymph nodes to elicit allo-recognition (Randolph et al. 2005). Major histocompatibility complex encoded proteins, predominate at antigen presenting cells, play roles to immune recognition, break of tolerance and production of anti-drug antibodies. The specific HLA has been shown association with drug hypersensitivity, demonstrating the significance of HLA in the immune allo-recognition process, HLA-B*5701, HLA-DR-7, and HLA-DQ3 alleles as genetic markers in HIV-1 reverse-transcriptase inhibitor abacavir and HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol (Mallal et al. 2002; Hung et al. 2005). Study in Thai population showed the association of HLA-DR-B1*09 and incidence of anti-r-HuEpo associated PRCA with repeated r-HuEpo subcutaneous exposure (Praditpornsilpa et al. 2009). There was a significant difference of HLA-DRB1*09 gene frequency linkage with HLA-DQ-B1*0309 between anti-r-HuEpo associated PRCA cases, and potential cadaveric kidney transplantation in the waiting list or potential national stem cell registry donor which represent background prevalence of HLA-DRB1*09 gene frequency of general population. The odd ratio of HLA-DRB1*09 allele for anti-r-HuEpo associated PRCA was 2.89 [95% CI: 1.88–4.46]. HLA DR-B1*09 allele is more common in Asian population but rare in Caucasian (<1%). However, a case-control study of risk factors associated with anti-r-HuEpo-associated PRCA, performed in Caucasian predominated population from Europe and Canada showed that the allele frequency of HLA-DRB1*09 occurs at a significantly higher frequency in

anti-r-HuEpo-associated PRCA cases than in controls. The odd ratio of HLA-DRB1*09 allele for anti-r-HuEpo associated PRCA in Caucasian population was 10.89 [95% CI: 2.2–53.7] (Fijal et al. 2008). Although HLA-DRB1*09 occurs at a significantly higher frequency in anti-r-HuEpo-associated PRCA cases, carrying the HLA-DRB1*09 allele was neither necessary nor sufficient to cause anti-r-HuEpo-associated PRCA.

Complexity of r-HuEpo Manufacturing

The manufacturing of biotherapeutics agents is complex and variable, require industrialize cell culture system, chemical and physical protocols for isolation and purification (Fig. 25.2) (Demain and Vaishnav 2009). The modification steps such as glycosylation of therapeutic proteins are considered a critical quality attribute and play an important role in favorable pharmacokinetic profile (Solá and Griebenow 2009). In addition, product stabilization against aggregation and chemical breakdown, packaging and cold-chained logistics can be issues for product efficacy and patient safety.

Alteration in manufacturing process may cause impurities or aggregation of the protein that may augment the immune response and can have major biological impact (Combe et al. 2005). The capital-intensive technology required for bio-therapeutic manufacturing results in expensive products. With the expiration of patent protection for the innovative r-HuEpo, biocopy r-HuEpo became available (Schellekens 2004). These biocopy r-HuEpos are more affordable, allowing chronic kidney disease patients with renal anemia easy access to treatment worldwide. European Medicines Agency (EMA) released regulatory guideline to evaluate

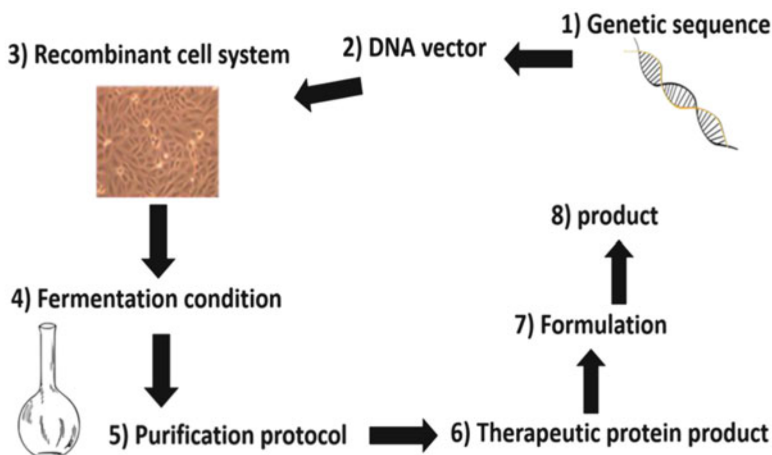


Fig. 25.2 The steps of biotherapeutic manufacturing

biocopy r-HuEpo focused on the quality attribute, pre-clinical and clinical data and classify as “Biosimilar r-HuEpo”. Verification of the similarity of biosimilar to innovator products has been a key challenge. Critical safety issue and the immunogenicity of biotherapeutic drugs has been highlighted, confirming a need for comprehensive immunogenicity testing prior to approval and extended post-marketing surveillance.

As for Thailand, the first biocopy r-HuEpo became available in 1997 for treatment of renal anemia in chronic kidney disease. Since then, there has been an expansion of biocopy r-HuEpo products in Thai drug market. More than 20 biocopy r-HuEpo products have been approved and licensed following the generic small molecule chemical drug pathway before the execution of biosimilar regulatory licensing pathway. Thus biocopy r-HuEpo products approved in Thailand cannot be classified as “Biosimilar r-HuEpo”. This unfortunate scenario in Thailand also demonstrates that biocopy r-HuEpo, if not regulated by step-wise science based biosimilar pathway which looks at totality of evidence (chemistry, analytical, clinical and safety), predisposes patients to immunogenicity related issues that arise from such uncharacterized products (Praditpornsilpa et al. 2011). Biocopy r-HuEpo should not be approved following the conventional small molecule generic pathway, which mainly focuses on bioequivalence study, and dossier based on the quality attribute was required. Study in Thailand showed the immunogenicity of biocopy r-HuEpos in Thai market (Joung et al. 2008), and it has been impossible to delineate which biocopy products were the cause due to a small sample size of the study. The findings of these studies mandates the need for re-evaluation of all biocopy r-HuEpo products in Thailand. The lack of stepwise quality, preclinical and clinical similarity exercises between biocopy r-HuEpo and its reference product can create potential gaps in overall quality and patient safety. As for most biotherapeutic products, data from pre-licensing are usually too limited to identify all potential adverse effects; pharmacovigilance plans and risk management plan should be submitted and integrated as parts of evaluation.

Pharmacovigilance Study in Thailand

Pharmacovigilance study, prospective, immunogenicity surveillance registry of r-HuEpo with subcutaneous exposure has been conducted in Thailand to monitor and ensure patient safety for biocopy r-HuEpo while the re-evaluation of r-HuEpo products has been ongoing, and address the question whether dissimilarity of product qualities affect immunogenicity profile. This question required analytical study of product qualities together with pharmacovigilance.

The pharmacovigilance aimed to estimate the incidence rate of anti-r-HuEpo associated PRCA among subjects who had subcutaneous exposure to any currently marketed r-HuEpo products in Thai market. The prospective cohort design allowed robust recording of exposure for individual subjects to all r-HuEpo products an accurate time sequence that might otherwise be difficult to determine retrospectively

and provided the framework for the direct calculation of incidence and for sensitivity analyses concerning various assumptions of latency and susceptibility. The registry was initially planned to address the risk of anti-r-HuEpo-associated PRCA. Subsequently, it also intended to investigate whether immunogenicity related to product qualities.

The registry enrolled subjects who were receiving duration of less than 1 month or about to receive (within 1 month) a marketed r-HuEpo product in Thailand by the subcutaneous route. Potential subjects were excluded if had a history of PRCA or aplastic anemia, history of unexplained loss or lack of effect (LOE) to a r-HuEpo product ongoing at the time of enrollment, had history of anti-r-HuEpo antibody prior to enrollment, and were currently receiving immunosuppressive medication (e.g., cyclosporine, tacrolimus, sirolimus, mycophenolic mofetil, azathioprine, or monoclonal antibodies) and prednisolone. The enrolled subjects were followed for unexplained LOE, hemoglobin level, iron study and any serious adverse event suspected to be causally related to use of any r-HuEpo product. Cases of unexplained LOE were investigated by standard of care and if remained unexplained LOE bone marrow aspiration/biopsy and anti-r-HuEpo antibody assay were investigated.

The preliminary result of the registry confirmed immunogenicity risk of r-HuEpo biocopy in Thailand. More than four thousand enrolled cases were chronic kidney disease patients who had renal anemia. The etiology of chronic kidney disease (CKD) mirrored the etiology of CKD in Thailand as diabetic nephropathy, hypertensive nephrosclerosis, chronic glomerulonephritis, obstructive uropathy and uric acid nephropathy were the leading cause. Majority of cases were CKD stage IV and V.

The registry protocol indicated non-interventional study; however, there was a predilection that enrolled subjects should receive only a single r-HuEpo product and subcutaneous administration throughout the study period. The preliminary results also showed that over 36 months, about 80–85% of registered cases used single r-HuEpo product over time. Although however, about 10–15% and 3–5% of registered cases used two and three r-HuEpo products by visit respectively (Fig. 25.3). The changing of each hospital drug formulary policy inevitably mandated the switching of products. The majority of enrolled cases achieved target hemoglobin of more than 10.0 g/dL.

Immunogenicity of Biocopy r-HuEpo in Thailand

Previous study (Praditpornsilpa et al. 2011) estimated risk for anti-r-HuEpo-associated PRCA in Thailand by using actual cases using biosimilar r-HuEpo denominator with PRCA cases. The estimation indicated that 1 out of 2608 patients using biosimilar r-HuEpo would develop PRCA which was likely to underestimate the risk. The final results of the pharmacovigilance will provide more accurate risk estimation and hope that the result will facilitate the national license re-evaluation of r-HuEpo products in Thailand.

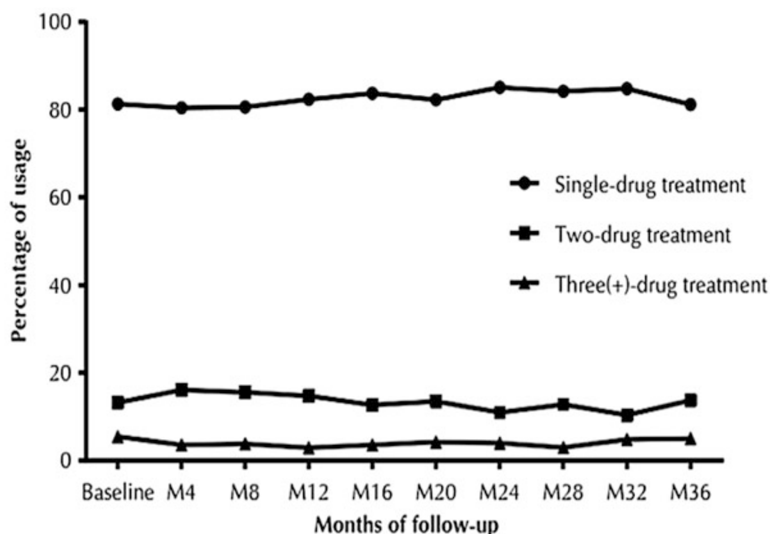


Fig. 25.3 Percentage of patients with different ESA treatment over time

Prospective Immunogenicity Surveillance Registry (PRIMS) study (Macdougall et al. 2015) enrolled 15,333 patients to estimate the incidence of anti-r-HuEpo associated PRCA in patients with renal anemia in Europe, mostly Caucasian population, all received innovative r-HuEpo. The median age of the patients was 73 years. Based on exposed time, the rate of PRCA was 35.8/100,000 patient-year (95% CI 7.4–104.7) for Eprex® versus 14.0/100,000 patient-year (95% CI 1.7–50.6) for Aranesp®/NeoRecormon® combined. The difference of incidence of anti-r-HuEpo associated PRCA between study in European and Thai populations raised issues of immunogenetic background, logistics, handling, cold chain maintenances of r-HuEpo (Locatelli et al. 2007), and product quality. The higher prevalence of HLA-DR*B09 may predispose to higher incidence of anti-r-HuEpo associated PRCA. In pharmacovigilance study prospective, immunogenicity surveillance registry of r-HuEpo with subcutaneous exposure in Thailand, logistics, handling, and cold chain maintenances of r-HuEpo were focused, closely monitored.

Product Quality Issues

r-HuEpo product quality issues (both reference and biosimilars) are known to be associated with immunogenicity. There was an upsurge of anti-r-HuEpo associated PRCA cases by Eprex® to incidence of 46.1/100,000 patient-years in 2001–2003. These patients used subcutaneous Eprex® prefilled syringes preserved with polysorbate-80 (PS-80) and uncoated stoppers (Boven et al. 2005). This incidence

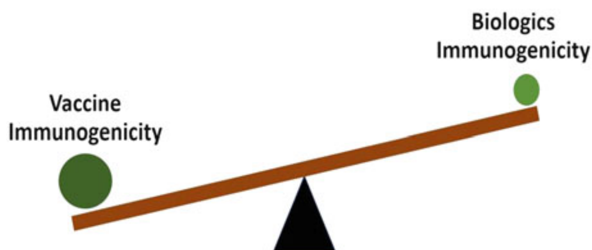


Fig. 25.4 The balance between biologics vs vaccine immunogenicity

occurred after PS-80 a stabilizer surfactant for formulation in prefilled syringes with uncoated rubber stopper was introduced to replace human serum albumin (HSA) to avoid the hypothetical risk of virus and prion transmission. Investigation for the root-cause of this incidence suggested that PS-80 directly exposed rubber materials in uncoated stoppers leached organic adjuvants and induced anti-r-HuEpo antibody (Jenke 2002; Ryan et al. 2006).

However, it has been unclear whether adjuvants alone are sufficient for breaking tolerance or induce allo-recognition and causing antibody response as in vaccination. Other mechanisms such as protein denaturation and alternation of protein conformation such as aggregation from contaminants may play role to enhance the susceptibility to immunogenicity. A root-cause investigation for a biosimilar in Europe (HX575) induced two neutralizing anti-r-HuEpo by subcutaneous exposure and cause PRCA cases (Haag-Weber et al. 2012) suggested that tungsten detected in suspected batch induced dimerization and aggregation with r-HuEpo product (Seidl et al. 2012). The inorganic tungsten leachate from the needle of pre-filled syringe can bind to r-HuEpo, denature the r-HuEpo structure, cause unfolding and aggregation. This aggregate molecule might directly elicit the immune response similar to adding aluminum salt as vaccine adjuvant into the traditional vaccine products (Fig. 25.4). These examples of organic and inorganic leachate suggested that careful attention for should not just be limited to active drugs (HuEpo) but also to the all component of drug products including formulation components (HSA or PS-80), syringe and tungsten needle type and extractables and (or) leachables from rubber components (Fujimori et al. 2013).

Besides the step-wise science driven similarity assessment that ensures the development and approval of a biosimilar product, appropriate manufacturing controls and conformance to Good Manufacturing Processes (GMP) are also required post approval and during life cycle management of the approved biosimilar. As manufacturing and processing controls determine the ultimate product quality of a biosimilar, only products that have been developed and manufactured following the GMP standards should be considered for licensing and commercialization approval. Noncompliance to GMP is definitely considered risk to product quality and hence a threat to patient safety. High standards of GMP ensure regulatory compliance in manufacturing, quality control, quality assurance and also change management and adequate reporting of post approval changes (Sahoo et al. 2009). While the regulation of biosimilars (during submission, pre-approval and post-approval) in

US, EU and Japan is now well defined, the regulation in low- and middle-income countries and some developing nations is not well defined which raises quality concerns for such products. It also raises concerns regarding patient safety if appropriate change management and post change assessment procedures are not in place once a product is approved.

Quality Issues of Biocopy r-HuEpo

r-HuEpo is a complex biologic and has several structural elements which need to be carefully controlled for development of a biosimilar. A study comparing the two alpha r-HuEpo products from Epogen® and Eprex®, suggested that biophysical characteristics namely hydrodynamic structure and protein stability of the same protein from two products manufactured by different manufacturers can be significantly different (Deechongkit et al. 2006).

A comparative study of physical and chemical properties of biocopy alpha r-HuEpo from Korea, China, and India with the reference product (Epogen®) showed that r-HuEpo from Korea, China, and India contained more glycoforms and other product related impurities. Moreover, in vitro testing revealed variation of relative potency for each product when based on the labeled concentration (Park et al. 2009). The literature evaluation of both analytic and clinical studies conducted with biocopy r-HuEpo products currently marketed outside the United States and Europe showed these products differed widely in composition, did not always meet self-declared specifications, and exhibited batch-to-batch variation. Clinical studies of these products were small sample size, of short duration which was impossible to detect the immunogenicity (Combe et al. 2005). Such findings signify the importance of product quality and clinical cohort for a rationally developed biosimilar r-HuEpo product.

Twelve brands of biocopy r-HuEpo that were highly prescribed in Thailand were investigated for product quality attributed by analytic methodology namely; high performance size-exclusion chromatography (HP-SEC), asymmetrical flow field-flow fractionation (AF4), sodium dodecyl sulfate polyacrylamide (SDS-PAGE) gel electrophoresis in combination with Western blotting (WB) and test for host cell protein (HCP) impurities as well as endotoxin contamination (Halim et al. 2014). These twelve products contained human serum albumin (HSA) as product stabilizer. By SDS-PAGE/WB, under non-reducing and reducing condition, all products showed the presence of dimeric r-HuEpo-r-HuEpo and dimeric r-HuEpo-HSA (Fig. 25.5). In addition, the profiles of r-HuEpo showed peaks corresponding to higher molecular weight (HMW) impurities in both HP-SEC and AF4, suggesting aggregation of either r-HuEpo or r-HuEpo/HSA. HP-SEC of twelve r-HuEpo products showed the presence of HSA dimer, trimer and HSA-r-HuEpo aggregation.

The aggregation of proteins (therapeutic or HSA like excipient) is generally linked to immunogenicity (Moussa et al. 2016). The low molecular weight species were also detected in all r-HuEpo products with difference proportion. Among twelve r-HuEpo products, there were over limit of HCP and endotoxin contamina-

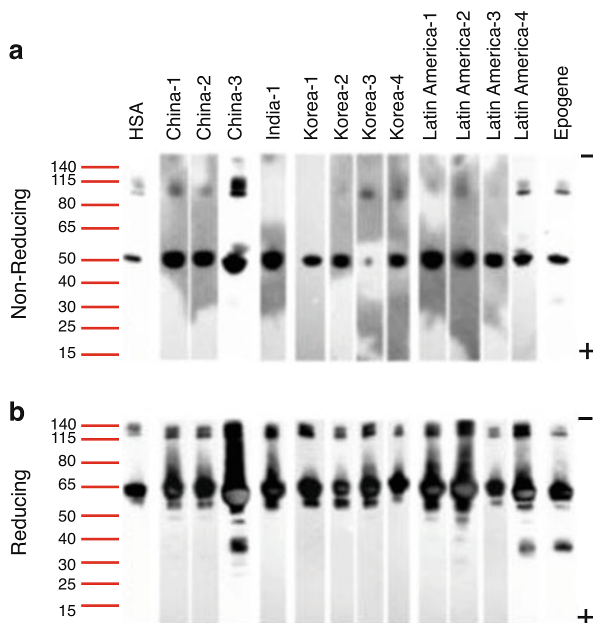


Fig. 25.5 Performance of various r-HuEpo copy products on SDS-PAGE (2 IU/lane) detected with western blot using anti-HSA antibody under (**a**) non-reducing and (**b**) reducing condition. (Image reproduced with permission from (Halim et al. 2014))

tions. HCPs impurities if co-purified with the active protein may trigger an immune response, which can be directed against the active protein leading to unwanted clinical consequences (Chirino and Mire-Sluis 2004). The limitation to detect HCP is specific HCP assays using anti-HCP antibodies, which were specifically generated against the cell line. In case the product coming from a specific process was not available, the HCP values can be underestimated (Bracewell et al. 2015). Endotoxin contamination of products cause adverse effect for example chronic inflammation in hemodialysis patients (Matsuhashi and Yoshioka 2002).

The presence of complex multimeric species (dimeric r-HuEpo-r-HuEpo and dimeric r-HuEpo-HSA) in r-HuEpo products in Thai market not only raise concerns regarding immunogenicity, but also raise a technical on how to perform analytic comparison for protein similarity between innovative products and biocopy r-HuEpo products. The US-FDA guidance on use of excipients in biosimilar formulations, does mention that it may be possible for a proposed product formulated without HSA to demonstrate biosimilarity to a reference product formulated with HSA. Furthermore the guidance also mentions that differences in certain posttranslational modifications or differences in certain excipients (e.g., HSA) might not preclude a finding of biosimilarity provided there is sufficient data submitted to prove that the proposed product is highly similar to reference product.

The assessment of analytical data in the presence of such excipients like HSA and comparison of products (biosimilar and reference product) after removal of such excipients is a complex process. A follow-up study to the study described by Deechongkit et al. (Deechongkit et al. 2006) discusses that a simple isolation

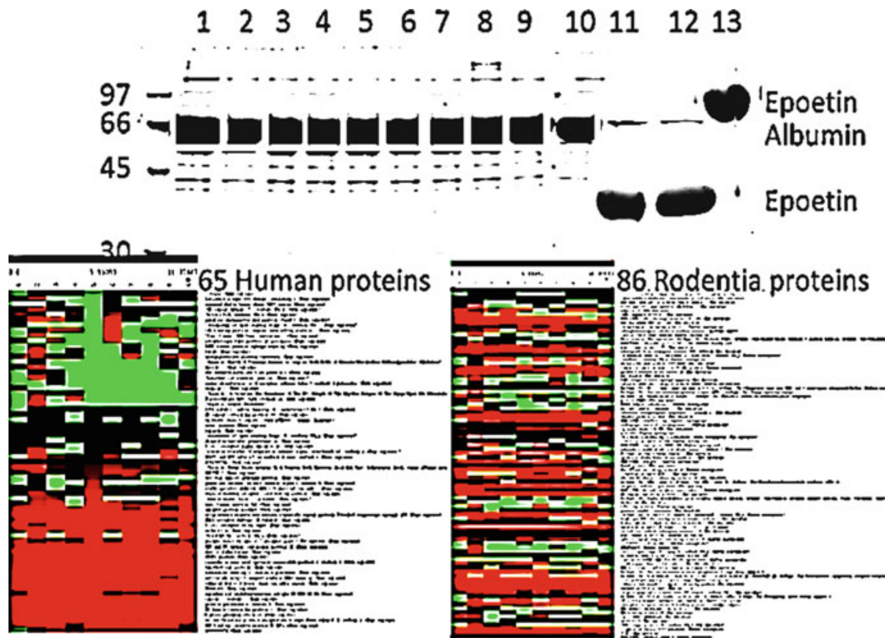


Fig. 25.6 The analysis of SDS-PAGE with GeLC-MS/MS to detect the protein impurities of r-HuEpo product (unpublished data)

procedure adopted to isolate the drug substance from Eprex® and Epogen®, can significantly compromise the active drug substance protein, which raises serious questions about such an approach and the interpretation of such analytical data (Heavner et al. 2007).

By proteomics analysis using one-dimensional SDS-PAGE followed by liquid chromatography-tandem mass spectrometry (GeLC-MS/MS) for qualitative and quantitative protein impurities of the products (Unpublished data), human proteins other than HSA and r-HuEpo and rodent proteins were detected (Fig. 25.6). Human protein impurities may come in the samples via HSA and rodent protein impurities may contaminate via rodent host cell culture. This finding mandated the quality and source of HSA if added as stabilizer should be regulated in addition of r-HuEpo quality.

Biosimilar r-HuEpo in US, Lessons Learned

The US Food and Drug Administration's (FDA) Oncologic Drugs Advisory Committee (ODAC) on 25th May 2017, voted 14-to-1 in support of approving a proposed biosimilar to Amgen's anemia drug Epogen (FDA Briefing Document 2017). This

process has been strictly in compliance with “The Biologics Price Competition and Innovation Act of 2009 which was enacted as part the Affordable Care Act”. “Epoetin Hospira”, proposed biosimilar r-HuEpo, was evaluated and compared to US-licensed Epogen product by using multiple orthogonal physicochemical and functional analytical methods. The comparative analytical similarity assessment for critical quality attribute (CQA) included immunogenicity by specific attribute measurement and method for primary structure, post translational modification (glycosylation), high order structure and biological activity. The manufacturing process was also evaluated for potential viral and other adventitious agent contamination. The assessment satisfactorily concluded that the proposed biosimilar r-HuEpo is highly similar to US-licensed Epogen, notwithstanding minor differences in clinically inactive components and that there are no clinically meaningful differences between the reference and proposed biosimilar r-HuEpo products. The stepwise approach was used to address the residual uncertainty at each step about the similarity of the proposed product to the reference product and identify next steps to address that uncertainty. Additional data, such as non-clinical pharmacokinetic/pharmacodynamic, toxicology) and clinical data (efficacy and safety including immunogenicity), were then designed to overcome those residual uncertainty. The combination of quality, nonclinical and clinical data then were finalized as “totality of evidence” presented for licensing approval. This approach adopted in US can be an important stepwise template of how biocopy r-HuEpo should be evaluated by regulatory agencies including regulators in low- and middle-income countries.

Conclusion

How r-HuEpo biocopy products were approved and used in Thailand confirms the need to carefully control the process of approval and the quality of biocopy or biosimilar products. The availability of biosimilars should be seen as chances to enhance the accessibility of drugs to patients especially in the low- and middle-income country where sometime “better than nothing idea” do apply. The implementation of universal health coverage scheme in Thailand allows quality, equitable and accessible health care for all. However, this ambitious policy causes economic burden of the country. Universal health coverage in Thailand covered biocopy reimbursement such as r-HuEpo, interferon and monoclonal antibody (Treerutkuarkul 2010). Biocopy compared to innovative biotherapeutic products alleviated the burden of national health fiscal. However, quality issues with biocopy/biosimilars, its adverse consequences especially when serious, should be avoided and should be preventable. One way to solve the dilemma to use or not to use biocopy is to find an acceptable quality biocopy by strengthening the regulaton in low- and middle-income countries do intensive evaluation, do similarity exercise for quality attribute, pre-clinical and clinical data and provide a wise selection based on totality of evidence plus a comprehensive pharmacovigilance and risk management plan. Only biocopy, which is closest to or is biosimilar

via biosimilar evaluation pathway, should be licensed. The comparative exercise must start with the quality attribute, and if quality attribute comparison signal questionable issues, full dossier must be required for the licensing procedure (Casadevall et al. 2011). For biotherapeutics already licensed by generic pathway, these products must be intensively re-evaluated and withdrawal of license should also be considered if they fail to ensure patient safety. The experience of biosimilar r-HuEpo in Thailand also reiterates the importance of pharmacovigilance. The adverse effects of biotherapeutics due to long term administration may sometimes not be captured in a phase III clinical trial. Pharmacovigilance, risk management plan must be presented and be evaluated as important part of licensing process, in one package with the quality, pre-clinical and clinical comparative exercise. In this regards, regulators in low- and middle-income countries need to be trained to cope with the expectations to ensure quality, safety and efficacy of biocopy products (Thanaphollert and Tungsanga 2011). The most urgent action required though is the re-evaluation of existing biocopies on the market. Such an action will certainly enhance the credibility of therapeutic equivalence of biotherapeutic products among health care professionals and patients.

References

Bennett CL, Luminari S, Nissenson AR, et al. Pure red-cell aplasia and epoetin therapy. New Engl J Med. 2004;351:1403–8.

Blackstone E, Fuhr J. The economics of biosimilars. Am Health Drug Benefits. 2013;6(8):469–78.

Boven K, Stryker S, Knight J, Thomas A, van Regenmortel M, Kemeny DM, Power D, Rossert J, Casadevall N. The increased incidence of pure red cell aplasia with an Eprex formulation in uncoated rubber stopper syringes. Kidney Int. 2005;67(6):2346–53.

Bracewell DG, Francis R, Smales CM. The future of host cell protein (HCP) identification during process development and manufacturing linked to a risk-based management for their control. Biotechnol Bioeng. 2015;112:1727–37.

Casadevall N. Antibodies against r-HuEPO: native or recombinant. Nephrol Dial Transplant. 2002;17(Suppl 5):42–7.

Casadevall N, Nataf J, Viron B, et al. Pure red cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. New Engl J Med. 2002;346:469–75.

Casadevall N, Thorpe R, Schellekens H. Biosimilars need comparative clinical data. Kidney Int. 2011;80:553.

Chirino AJ, Mire-Sluis A. Characterizing biological products and assessing comparability following manufacturing changes. Nat Biotechnol. 2004;22(11):1383–91.

Combe C, Tredree RL, Schellekens H. Biosimilar epoetins: an analysis based on recently implemented European medicines evaluation agency guidelines on comparability of biopharmaceutical proteins. Pharmacotherapy. 2005;25:954–62.

Deechongkit S, Aoki KH, Park SS, Kerwin BA. Biophysical comparability of the same protein from different manufacturers: a case study using Epoetin alfa from Epogen® and Eprex®. J Pharm Sci. 2006;95:1931–43.

Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv. 2009;27:297–306.

Dorey E. How the biologics landscape is evolving. Clin Pharm. 2014;6(9):1–6.

Erslev AJ. Pure red cell aplasia. In: Beutler E, Lichtman MA, Coller BS, Kipps TJ, editors. Williams hematology. New York: McGraw-Hill; 1995. p. 448.

FDA Briefing Document. Onchologic Drugs Advisory Committee Meeting, May 25, 2017, BLA 125545 “Epoetin Hospira”, a proposed biosimilar to Epogen/Procrit 9epoetin alfa). Applicant: Hospira Inc., Pfizer Company.

Fijal B, Ricci D, Vercammen E, Palmer PA, Fotiou F, Fife D, Lindholm A, Broderick E, Francke S, Wu X, Colaianne J, Cohen N. Case-control study of the association between select HLA genes and anti-erythropoietin antibody-positive pure red-cell aplasia. Pharmacogenomics. 2008;9(2):157–67.

Fujimori K, Lee H, Phillips J, Nashed-Samuel Y. Development of an inductively coupled plasma mass spectrometry method for quantification of extracted tungsten from glass prefilled syringes used as a primary packaging for pharmaceutical and therapeutic protein products. PDA J Pharm Sci Technol. 2013;67(6):670–9.

Haag-Weber M, Eckardt KU, Hörl WH, Roger SD, Vetter A, Roth K. Safety, immunogenicity and efficacy of subcutaneous biosimilar epoetin-α (HX575) in non-dialysis patients with renal anemia: a multi-center, randomized, double-blind study. Clin Nephrol. 2012;77(1):8–17.

Halim LA, Brinks V, Jiskoot W, Romeijn S, Praditpornsilpa K, Assawamakin A, Schellekens H. How bio-questionable are the different recombinant human erythropoietin copy products in Thailand? Pharm Res. 2014;31(5):1210–8.

Heavner GA, Arakawa T, Philo JS, Calmann MA, LaBrenz S. Protein isolated from biopharmaceutical formulations cannot be used for comparative studies: follow-up to “a case study using Epoetin Alfa from Epogen and EPREX”. J Pharm Sci. 2007;96:3214–25.

Hung S, Chung W, Liou L, et al. HLA-B * 5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. Proc Natl Acad Sci U S A. 2005;102:4134–9.

Jenke D. Extractable/leachable substances from plastic materials used as pharmaceutical product containers/devices. PDA J Pharm Sci Technol. 2002;56:332–71.

Joung J, Robertson JS, Griffiths E, WHO Informal Consultation Group, et al. WHO informal consultation on regulatory evaluation of therapeutic biological medicinal products held at WHO Headquarters, Geneva, 19–20 April 2007. Biologicals. 2008;36:269–76.

Lacombe C. Resistance to erythropoietin. New Engl J Med. 1996;334:660–2.

Locatelli F, Del Vecchio L, Pozzoni P. Pure red-cell aplasia ‘epidemic’– mystery completely revealed? Perit Dial Int. 2007;27(Suppl 2):S303–7.

Macdougall IC, Roger SD, de Francisco A, et al. Antibody-mediated pure red cell aplasia in chronic kidney disease patients receiving erythropoiesis-stimulating agents: new insights. Kidney Int. 2012;81:727–32.

Macdougall IC, Casadevall N, Locatelli F, Combe C, London GM, Di Paolo S, Kribben A, Fliser D, Messner H, McNeil J, Stevens P, Santoro A, De Francisco ALM, Percheson P, Potamianou A, Foucher A, Fife D, Mérit V, Vercammen E, PRIMS Study Group. Incidence of erythropoietin antibody-mediated pure red cell aplasia: the prospective immunogenicity surveillance registry (PRIMS). Nephrol Dial Transplant. 2015;30(3):451–60.

Mallal S, Nolan C, Witt C, et al. Association between presence of HLA-B * 5701, HLA-DR-7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. Lancet. 2002;359:727–32.

Mandreoli M, Finelli C, Lopez A, Ascani S, Vianelli N, Baccarani M, Santoro A. Successful resumption of epoetin alfa after rituximab treatment in a patient with pure red cell aplasia. Am J Kidney Dis. 2004;44(4):757–61.

Matsuhashi N, Yoshioka T. Endotoxin-free dialysate improves response to erythropoietin in hemodialysis patients. Nephron. 2002;92:601–4.

McLeod D, Shreeve M, Axeirad A. Improved plasma culture system for production of erythrocytic colonies in vitro: quantitative assay method for CFU-E. Blood. 1974;44:517–34.

Moussa EM, et al. Immunogenicity of therapeutic protein aggregates. J Pharm Sci. 2016;105: 417–30.

Park SS, Park J, Ko J, Chen L, Meriage D, Crouse-Zeineddini J, Wong W, Kerwin BA. Biochemical assessment of erythropoietin products from Asia versus US Epoetin Alfa manufactured by Amgen. J Pharm Sci. 2009;98:1688–99.

Peces R, Torre M, Alcazar R, Urra J. Antibodies against recombinant human erythropoietin in a patient with erythropoietin resistance anemia. New Engl J Med. 1996;3356:523–4.

Porter S. Human immune response to recombinant human proteins. J Pharm Sci. 2001;90:1–11.

Prabhakar SS, Muhlfelder T. Antibodies to recombinant human erythropoietin causing pure red cell aplasia. Clin Nephrol. 1997;47:331–5.

Praditpornsilpa K, Buranasot S, Bhokaisuwan N, Avihingsanon Y, Pisitkul T, Kansanabuch T, Eiam-Ong S, Chusil S, Intarakumtornchai T, Tungsanga K. Recovery from anti-recombinant-human-erythropoietin associated pure red cell aplasia in end-stage renal disease patients after renal transplantation. Nephrol Dial Transplant. 2005;20(3):626–30.

Praditpornsilpa K, Kupatawintu P, Mongkonsritagoon W, Supasyndh O, Jootar S, Intarakumthornchai T, Pongskul C, Prasithsirikul W, Achavanuntakul B, Ruangkarnchanasetr P, Laohavinij S, Eiam-Ong S. The association of anti-r-HuEpo-associated pure red cell aplasia with HLA-DRB1$_*$09-DQB1$_*$0309. Nephrol Dial Transplant. 2009;24:1545–9.

Praditpornsilpa K, Tiranathanagul K, Kupatawintu P, Jootar S, Intragumtornchai T, Tungsanga K, Teerapornlertratt T, Lumlertkul D, Townamchai N, Susantitaphong P, Katavetin P, Kanjanabuch T, Avihingsanon Y, Eiam-Ong S. Biosimilar recombinant human erythropoietin induces the production of neutralizing antibodies. Kidney Int. 2011;80(1):88–92.

Praditpornsilpa K, Tiranathanakul K, Jootar S, Tungsanga K, Eiam-Ong S. Rechallenge with intravenous recombinant human erythropoietin can be successful following the treatment of anti-recombinant erythropoietin associated pure red cell aplasia. Clin Nephrol. 2014;81(5):355–8.

Randolph GJ, Angeli V, Swartz MA. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. Nat Rev Immunol. 2005;5(8):617–28.

Ryan MH, Heavner GA, Brigham-Burke M, et al. An in vivo model to assess factors that may stimulate the generation of an immune reaction to erythropoietin. Int Immunopharmacol. 2006;6:647–55.

Sahoo N, Choudhury K, Manchikanti P. Manufacturing of biodrugs: need for harmonization in regulatory standards. BioDrugs. 2009;23(4):217–29.

Schellekens H. When biotech proteins go off-patent. Trends Biotechnol. 2004;22:406–41.

Seidl A, Hainzl O, Richter M, Fischer R, Böhm S, Deutel B, Hartinger M, Windisch J, Casadevall N, London GM, Macdougall I. Tungsten-induced denaturation and aggregation of epoetin Alfa during primary packaging as a cause of immunogenicity. Pharm Res. 2012;29:1454–67.

Shimizu H, Saitoh T, Ota F, Jimbo T, Tsukada Y, Murakami H, Nojima Y. Pure red cell aplasia induced only by intravenous administration of recombinant human erythropoietin. Acta Haematol. 2011;126(2):114–8.

Solá RJ, Griebenow K. Effects of glycosylation on the stability of protein pharmaceuticals. J Pharm Sci. 2009;98:1223–45.

Thanaphollert P, Tungsanga K. Towards regulation of similar biotherapeutic products: Thailand's perspective. Biologicals. 2011;39:346–7.

Treerutkuarkul A. Thailand: health care for all, at a price. Bull World Health Organ. 2010;88(2):84–5.

Urra J, Torre M, Alcaza P, Peces R. Rapid method for detection of anti-recombinant human erythropoietin antibodies as a new form of erythropoietin resistance. Clin Chem. 1997;43: 848–9.

Walsh G. Biopharmaceutical benchmarks 2014. Nat Biotechnol. 2014;32(10):992–1000.

Index

A
Abasaglar, 226
Abatacept, 500
Abbreviated Biologic License Application (aBLA), 76
Abbreviated new drug application (ANDA), 7
Abbreviated new drug submission (ANDS) pathway, 139
AbbVie, 28
ABP501, 294
  on adalimumab PK profile, 593–594
  analytical similarity, 592
  clinical evidence, 593
  of efficacy, 594–595
  functional similarity, 592
  immunogenicity, 595–596
  inflammatory bowel disease, 592
  PK similarity study, 593
  safety, 596
ABP-501, 386
Abseamed, 226
Absolute neutrophil count (ANC), 132
Accofil, 226
Accountable Care Organizations (ACOs), 40
Active pharmaceutical ingredient (API), 225
Adalimumab, 498–499, 514–517, 654, *see also* ABP501
Adaptive analytical characterization strategy
  adaptive characterization strategy, 313
  biosimilar product characterization, 311–313
  FDA regulatory guidelines, 306
  lifecycle management
    continued procedure performance verification, 315
    FDA process validation guidance, 314
    glycosylation, 314
    ICH Q2(R1), 314
    ICH Q8 (2009c), 314
    monoclonal antibody biosimilar products, 316–318
    procedure design, 315
    procedure performance qualification, 315
    QbD, 314
    reference and biosimilar products, 314
  protein product characterization, 307–309
  reference product characterization, 309–311
  351(k) licensure pathway, 306
Addressing assay variability
  assay training, 414
  assay troubleshooting, 415
  cell type/cell line selection, 412–413
  data analysis, 414
  procedural accuracy, 413–414
Adverse drug reactions (ADR), 133
Affordable Care Act, 688
Agência Nacional de Vigilância Sanitária (ANVISA), 5
Agila Biotech Pvt. Ltd., 658
AlphaLISA, 396
Alpheon, 227
ambr®, 177, 178
Amgen, 27, 34–35, 93, 96
Amgen Inc. v. Hospira, Inc., 94–95
Amgen v. Apotex, 34–35
Amgen v. Hospira, supra, 98
Amgevita, 227
Amjevita®, 240

© American Association of Pharmaceutical Scientists 2018

H. J. Gutka et al. (eds.), *Biosimilars*, AAPS Advances in the Pharmaceutical Sciences Series 34, https://doi.org/10.1007/978-3-319-99680-6

Amjevita syringes, 244
Analysis of the variance (ANOVA), 562, 563
Analytical development
QbD and, 212
analytical method capability and continual improvement, 214
CMA, 212–213
control strategy, 213–214
design space, 213
QTAMP, 212
risk assessments, 213
Analytical methods, 213, 214
Analytical similarity
adjustment for age differences, 442–445
assessment, 434–435
biotechnology-derived drugs, 432
CQA
selection of K for Tier 2, 447–451
Tier 1 equivalence margin, 445–447
Tier 1 equivalence test, 438–441
description, 473–474
EMA statistical approach, 435–437
FDA approach, criticisms of (*see* FDA approach, criticisms of)
FDA statistical approach, 437–438
reference product, 433
stepwise approach, 433–434
test product, 433
Tier 2 testing, 441–442
Tier 3 testing, 445
Analytical similarity assessment
analytical testing plan
gold standards, 282
head-to-head testing, 284
intrinsic variability, 281
method qualification, 282
method suitability parameter selection, 283
method validation, 282
resolving power, 281
sensitivity of detection, 281
state-of-the-art analytical technologies, 281
biosimilar product, 274–275
data assessment
acceptable difference, 285
core fucosylation of N-glycans, 285
EMA, 288–290
Fcγ3A receptor binding affinity, 285
U.S. FDA, 286–288
EMA guidelines, 262
forced degradation
ABP 501's aggregation rate, 294
analytical method screening and selection, 292
analytical response measures, 296
CCD, 296, 297
comparative stress studies, 291, 299–301
contour plot, 297, 298
control samples, 297
experimental noise, 298, 299
generic experimental conditions, 291
head-to-head characterization tests, 291
LC-MS analysis, 295
product's development lifecycle, 290
response measure, 295
selection of screening factors, 294
side-by-side analysis, 291
simultaneous analysis, 295
stressed materials, 291
test materials, 292–293
two-staged approach, 292
quality attribute, selection and ranking of
assessment hierarchy, 280
criticality assessment, 277–278
definition, 275
degree of variability, 278
host cell proteins, 280
information sources, 278
molecular variants, 280
product-related substances/impurities, 278
qualitative *vs.* quantitative, 278–279
selection of, 276–277
Tool 1, 280
worst-case situation, 280
reference product
batch size, 268
comparator product, 268
legislative provisions, 268
Monte Carlo simulation, 269
non-domestic reference product, 268
obvious factor of risk, 268
spacing reference product procurement, 268
statistical analysis, 269
Tier 1 and Tier 2 assessments, 270–272
Tier 2 assessments, 272
test materials
analytical feasibility, 266–267
different presentations, 263–264
EU guideline, 263
incremental variability, 263
process changes, 262
product age, 264–266
statistical considerations, 267

Analytical strategy, CMC, 182–185
Anaphylaxis, 367
Anion exchange (AEX) column purification, 346
Antibody dependent cell cytotoxicity (ADCC), 605
Antibody-dependent cell-mediated cytotoxicity (ADCC) assay, 401, 402, 414, 419
Antibody-dependent cellular cytotoxicity (ADCC), 224
Antibody dependent cellular phagocytosis (ADCP), 401
Antibody drug conjugates (ADCs), 401
Antibody therapeutics
  binding assays
    CLB, 396–397
    LBA, 396
  bioassays
    in vitro cell line based bioassays, 398
    in vitro tissue based bioassays, 398
    in vivo bioassay, 397–398
    reporter gene based bioassays, 398–399
Anti-CD20 monoclonal antibody, 123
Anti-drug antibodies (ADAs), 135, 136, 366
  European perspective, on biosimilars, 159–161
  QbD, 203
Anti-receptorMAbs, 400
Anti-r-HuEpo, 678
Apoptosis assay, 422
Apotex, 34–35
Aranesp®, 264
Assay development, 136
Assay repeatability, 405
Assay troubleshooting, 415
Association for Accessible Medicines (AAM), 4
Avastin®, 241
Average sales price (ASP), 55–57
Avonex®, 367

B
Basaglar, 113
Bemfola, 227
Benepali, 227
Betaseron®, 367
Bevacizumab, 480–483
BIA evaluation software, 427
Binding assays
  CLB, 396–397
  LBA, 396
Binocrit, 228
Bioassays, 395
  in vitro cell line based bioassays, 398
  in vitro tissue based bioassays, 398
  in vivo bioassay, 397–398
  reporter gene based bioassays, 398–399
Biocon, 36, 658
Biocopy, 677
Biograstim, 228
Biological activity assays
  ADCC and CDC, 418
  ADCC assay, 419
  addressing assay variability
    assay training, 414
    assay troubleshooting, 415
    cell type/cell line selection, 412–413
    data analysis, 414
    procedural accuracy, 413–414
  analytical biosimilarity assessment, 417–418
  antibody therapeutics (*see* Antibody therapeutics)
  apoptosis assay, 422
  bioassays, 395
  biological dose-response modeling, 416–417
  biosimilar product life-cycle management
    assay replacement/comparability, 406–407
    binding or cell-based functional assays, 402
    cell based assay optimization, 403–404
    critical reagents, 406
    legacy potency assay, 406–407
    locked down, 402
    method remediation, 403–404
    method tech transfer, 408
    method validation, 404–405
    QTPP, 402
  biosimilar products, 394
  CDC assay, 419–422
  cell surface antigen binding assay, 422–424
  C1q binding assay, 425
  critical quality attributes, 415–416
  Fc$^{\gamma}$R I binding assay, 425
  Fc$^{\gamma}$R IIb binding assay, 425–426
  Fc$^{\gamma}$RIIIa-Phe binding assay, 427
  Fc$^{\gamma}$R IIIa-Val binding assay, 427
  FcRn binding assay, 427–428
  functional assays, 409
  monoclonal antibodies, 394
  physicochemical analyses, 394
  physicochemical and functional studies, 411–412

Biological activity assays (*cont.*)
potency assays, 395, 409–410
Reference Standard, 410–411
regulatory pathway, 408
robust analytical testing, 408
therapeutic MoA
ADCs, 401
agonistic MoA, 399
antagonistic MoA, 399–400
effector function of MAbs, 401
multiple/novel MoA, 400–401
Biological medicines, 107
drug development, 108
Biological Price Competition and Innovation Act of 2009 (BPCIA), 30, 114
Biologicals Working Party (BWP), 146, 147
Biologic medicines, 123
Biologics, 7, 38, 43, 123, 222
R&D costs, 24–26
Biologics license application (BLA), 8, 115, 572
Biologics market, 16
Biologics Price Competition and Innovation Act (BPCI Act), 8, 10, 76–77, 471
biosimilar's inaction, 83
biosimilar's response, 80
commercialization, 180-day notice of, 83
confidential information, biosimilar's disclosure of, 79
confidentiality provisions, 78–79
exchanged lists, 81–82
Hatch-Waxman Act, 84
infringement, act of, 78
litigation, preparation for
biosimilar's preparation, 89–90
sponsor's preparation, 86–88
negotiated list, 81
patent exchange/patent dance time line, 82
sponsor's initial patent list, 80
sponsor's reply, 81
Biologics Pricing Competition and Innovation (BPCI) Act, 544
Biologics Working Party (BWP), 9
Biopharma industry, 25
Biosimilar biologic drug approvals
authorizing indications, 137–138
Canadian guidance, 137
patent hold, 138–139
safety and efficacy, 137–138
biologics, 123–124
biotechnological/biologic products, comparability of, 124–125
clinical development, 130
efficacy, 133–135
immunogenicity, 135–137
pharmacodynamic studies, 132–133
pharmacokinetic studies, 130–132
safety, 133
development paradigm, 128–129
generic pharmaceuticals, 124
interchangeability, 139–141
non-clinical development, 130
post-market monitoring, 142
reference biologic drug, 126–128
regulatory framework, 125–127
switching, 141–142
Biosimilar development
bridging studies, 119–120
chemistry, manufacturing, and controls, 116–117
clinical studies, 117–119
global regulatory strategy framework, 108, 109
interest/competitive intelligence, countries of, 108–110
nonclinical studies, 117
product target profile, 110
regulatory pathways, 110
EMA guidelines, 111–112
FDA guidelines, 112–115
WHO, 115–116
regulatory strategy, 120
strategic plan, 120
Biosimilar, evaluation process, 472–473
Biosimilar mAb HOS comparability analysis, 343–344
Biosimilar market, 23–24
approvals, FDA lag in, 29
barriers to entry, 26–29
biologics R&D costs, 24–26
biologics and, 43
competition, 24
data and market exclusivity, 30–31
interchangeability, 37
Mylan, 36
patent dance, litigation
Amgen v. Apotex, 34–35
Janssen v. Celltrion, 35–36
Sandoz v. Amgen, 32–34
patents
as barriers to entry, 31–32
issues, 37
litigation before application, 32
patents and exclusivity, economic justification for, 29–30
potential gains, 41–42

pricing and market opportunities, 38–39
risks and return on investment, 42–43
in U.S., 39–41
Biosimilar Medicinal Products Working Party (BMWP), 9
Biosimilar MedicinesWorking Party (BMWP), 146
Biosimilar products, 394
Biosimilar r-HuEpo, 681
Biosimilars
AAM, 4
biologics market, 16
development and regulatory pathway
small *vs.* large molecules, 6–7
totality of the evidence, 8–9
U.S., small *vs.* large molecules, 7–8
drug development, evolution in, 4
Hatch-Waxman Act, 4
regulation
controversial regulatory concepts, 10
extrapolation, 10–12
harmonization, issue of, 15
history of, 9–10
switching *vs.* interchangeability, 12–15
reimbursement strategies and incentives, 18–19
TCSDD, 4
tenets of, 5–6
uptake successes, 17–18
U.S. FDA, 5
Biosimilars drug products, 274–275
amjevita PFS, 250
API, 225
assessing similarity, 222
assessment considerations, 224–225
autoinjector, 251
BDP, 247
biobetter, 246
characterization, 311–313
container/closure system, 223
cyltezo PFS, 250
difference selected formulation, 245
different pH, 243
dosage forms for, 246–247
drug product manufacturing process, 251–252
enbrel PFS, 250
EU approved/refused and withdrawn products, 225–237
formulation, 223
herceptin, 244
"highly similar" molecule, 222
home-use products, 223
immunogenicity, 222
intellectual property, 223
in-use studies, consideration of, 245
life-cycle management
assay replacement/comparability, 406–407
binding or cell-based functional assays, 402
cell based assay optimization, 403–404
critical reagents, 406
legacy potency assay, 406–407
locked down, 402
method remediation, 403–404
method tech transfer, 408
method validation, 404–405
QTPP, 402
manufacturing process, 223
molecule and drug substance, 223–224
natural isotope abundance 2D NMR, 243
non-clinical and clinical safety, 222
PFS, 249
pharmacodynamics and pharmacokinetics, 222
primary packaging and device, 247–248
Pure Red Cell Aplasia, 250
quality parameters, 245
rubber components, 248–249
shelf-life, 244
silicone oil and tungsten, 249
structural stabilizers, 244
surfactants, 244
target concentration, 247
tier 1 equivalence test, 249
tier 3 qualitative comparison, 249
tier 2 quality range criteria, 249
tungsten pins, 250
type 1 glass, 248
US approved products, 225, 238–242
Biosimilars, in India development and commercialization
analytical similarity
filgrastim, 665–667
G-CSF, 654
ICH Q5E and Q6B guidelines, 664
independent in-depth similarity, 665
rituximab biosimilars, 667–669
biosimilar confirmatory, 654–655
biosimilars regulations
CDSCO, 659–660
EMA, 659
key stakeholders, 660
KFDA, 659
MFDS, 659

Biosimilars, in India development and commercialization (*cont.*)
2016 guidelines, 661–664
US-FDA, 659
WHO, 659
CAGR, 654
clinical similarity, 669–670
Filgrastim, 654
global biosimilar market, 656–657
indian biosimilar market, 657–658
Indian biotech sector's revenue, 654
Indian players with global partnerships, 658
insulin analogs, 654
monoclonal antibodies, 654
prices of biologics, 656
scientific bridge, 654
small molecule generics *vs.* biosimilars, 655
Biosimilar sponsor's testing, 417
Blitzima, 228
B-lymphocytes, 491
Broadest reasonable interpretation, 101
Bulk drug product (BDP), 247

C

Canada, biosimilar biologic drug, *see* Biosimilar biologic drug approvals
Canadian reference product (CRP), 139
Capable process, 210
Cation-exchange chromatography, 179
Cation exchange (CEX) column purification, 346
CD spectroscopy, 331
Cell-based binding assays (CBA), 396
Cell surface antigen binding assay, 422–424
Celltrion, 35–36, 93, 94
Celltrion's follow-up data package, 591–592
Center for Medicare and Medicaid Innovation (CMMI), 71
Centers for Medicare and Medicaid Services (CMS), 51, 54
Central composite design (CCD), 296, 297
Central Drug Standard Control Organization (CDSCO), 657
Certificate of Analysis, 174
Chaperone proteins, 352
Chemistry, manufacturing and controls (CMC), 173
BDS, specifications for, 181
expression systems and clonal selection, 175–177
implementation and analytical strategy, 182–185
product quality
and control strategy, manufacturing conditions on, 180–182
raw materials effect, 179–180
reference drug, 174–175
upstream and downstream processes and scale-up to manufacturing, 177–179
Cipla Ltd., 657, 658
Circular dichroism (CD), 331, 340
Clinical Commissioning Groups (CCGs), 18, 67
Clinical development
biosimilar biologic drug approvals, 130
efficacy, 133–135
immunogenicity, 135–137
pharmacodynamic studies, 132–133
pharmacokinetic studies, 130–132
safety, 133
Clinical studies
European perspective, biosimilars, 155
clinical comparability, 158–159
extrapolation of efficacy and safety, 162–163
immunogenicity, 159–161
interchangeability, 163–165
pharmacodynamic comparability, 157–158
pharmacokinetic comparability, 155–157
Clinical trial materials (CTM), 189
Clonal selection, CMC, 175–177
Commercialization, 180-Day Notice of, 89
Committee for Human Medicinal Products (CHMP), 146, 147
Committee for Medicinal Products for Human Use (CHMP), 9
Comparative clinical studies, design considerations
bevacizumab, 480–483
clinically meaningful differences, 475–477
etanercept, 478–479
filgrastim, 479–480
safety signals, 477–478
stepwise approach, 474
Competitive intelligence, countries of, 108–110
Competitive ligand binding (CLB) assay, 396–397, 400
Complement-dependent cytotoxicity (CDC) assay, 402, 414, 419–422
"Composition of matter" patent, 31

Compound annual growth rate (CAGR), 654
Confirmatory phase III-type safety, 619–620
Consumer Price Index-Urban (CPI-U) adjustment, 62
Continual improvement, 196, 214
Control strategy, QbD, 195–196
Conventional cleanroom, 253
Corynebacteria, 175
Coverage gap, 58
Coverage Gap Discount Program (CGDP), 58
C1q binding assay, 425
Critical formulation attributes (CFAs), 215
Critical material attributes (CMAs), 190
Critical process parameters (CPP), 194
Critical quality attributes (CQAs), 129, 174, 175, 180, 182–185, 366, 688
  QbD, 191–192, 200–201
  selection of K for Tier 2, 447–451
  Tier 1 equivalence margin, 445–447
  Tier 1 equivalence test, 438–441
CT-P13
  in afucosylated species, 585–586
  analytical characterization, 582–583
  Celltrion's follow-up data package, 591–592
  clinical development program, 587
  efficacy and safety comparison, 587–588
  functional characterization, 582–583
  AS patients, PK equivalence study, 584, 587
  RA and AS indications, 589–591
  residual uncertainty of afucosylated glycans, 583–584
Cyltezo™, 228, 241, 246
Cyltezo syringes, 244

D
Darbepoetin, 499–500
DasBox, 177
Decision rules, 287
Design of Experiments (DOE) approach, 193–194, 207–208
Design space
  QbD, 193–195
  QTPP, 207–209
Differential scanning calorimetry (DSC), 332
District Level Committee (DLC), 660
DM Corporation, 658
Donut hole, 58
Draft Reflection Paper, 289, 290
Dr. Reddy's Laboratories, 657
Drug Controller General of India (DCGI), 660
Drug product (DP), 189
Drug substance (DS), 189
Dynamic light scattering (DLS), 378–380

E
ELISA-based assay, 352
ELISA-based binding assay, 344
ELISA-based method, 343
Enbrel®, 27, 239
Enbrel PFS, 250
Endogenous proteins, 491
Enzyme linked immuno-sorbent assay (ELISA), 396
Epoetin alfa-ADAs, 159
Epoetin Alfa Hexal, 229
Epoetin Hospira, 688
Epogen®, 93, 687
Epogen/Procrit, 242
Eprex®, 683, 687
Erelzi®, 229, 239, 246
Erelzi® methods, 614
Erythropoiesis-stimulating agent (ESA), 387
Erythropoietin (EPO), 16
Etanercept®, 17, 96, 353, 478–479
EudraVigilance, 643
European Crohn's and Colitis Organization (ECCO), 11
European Economic Area (EEA), 151
European Medicines Agency (EMA), 5, 6, 111–112, 128, 262, 288–290, 659, 680
  guidelines and scientific assessment reports, 146
European Medicines Evaluation Agency (EMEA), 146
European perspective on biosimilars
  assessment of biosimilars, 149–150
  clinical studies, 155
    clinical comparability, 158–159
    extrapolation of efficacy and safety, 162–163
    immunogenicity, 159–161
    interchangeability, 163–165
    pharmacodynamic comparability, 157–158
    pharmacokinetic comparability, 155–157
  EU, 146
    evolution of, 147
  guidelines, for biosimilars, 148–149
  non–clinical development, 153–155

European perspective on biosimilars (*cont.*)
post-marketing surveillance, 167
product information, 165–166
quality, 151–153
reference product, 151
European Pharmacopeia (EP), 200
European Union (EU), 146
evolution of, 147
Expression systems
CMC, 175–177
Extrapolation
ABP501
on adalimumab PK profile, 593–594
analytical similarity, 592–593
clinical evidence, 593
of efficacy, 594–595
functional similarity, 592–593
immunogenicity, 595–596
inflammatory bowel disease, 592
PK similarity study, 593
safety, 596
biosimilar biologic drug approvals, 137–138
Canadian guidance, 137
patent hold, 138–139
safety and efficacy, 137–138
biosimilars, regulation, 10–12
BLA pathway, 572
CT-P13
in afucosylated species, 585–586
analytical characterization, 582–583
Celltrion's follow-up data package, 591–592
clinical development program, 587
efficacy and safety comparison, 587–588
functional characterization, 582–583
AS patients, PK equivalence study, 584, 587
RA and AS indications, 589–591
residual uncertainty of afucosylated glycans, 583–584
European perspective on biosimilars, 162–163
immunogenicity comparisons, 580–581
MOAs
antibody drugs, 576–577
drug delivery carriers, 577
efficacy and safety, 578–579
glycosylation impacts, 577
hormonal protein drugs, 576
structural residual uncertainty, 577–578
modeling and simulation, potential, 581–582
PK considerations, 579–580
regulatory requirement
EMA, 575
FDA, 574
Health Canada, 575–576
scientific principles, 573
TMDD, 581
Totality of Evidence, 621
Extrinsic particles, 365

F
Fc$^{\gamma}$R I binding assay, 425
Fc$^{\gamma}$R IIb binding assay, 425–426
Fc$^{\gamma}$RIIIa-Phe binding assay, 427
Fc$^{\gamma}$R IIIa-Val binding assay, 427
FcRn binding assay, 427–428
FDA approach, criticisms of
correlation effects, 460–465
reformulation of hypotheses, Tier 1 test
equal variances, 456–457
numerical data, 458–459
simulated type I error rates, 459–460
unequal variances, 457–458
Tier 1 test
impact of correlation, 453–454
inflation of type I error, 452–453
Tier 2 test
impact of correlation, 455
selection of K, 454–455
Federal Food, Drug, and Cosmetic Act (FD&C Act), 112
Federal Medical Assistant Percentage (FMAP), 61–62
Fee-for-service (FFS) model, 70
Filgrastim, 200, 479–480
Filgrastim hexal, 229
Filgrastim ratiopharm, 229
Fingerprint-like similarity, 312, 328–330
Flixabi, 230
Flow imaging microscopy (FIM) analysis, 376–377
Food and Drug Administration (FDA) guidelines, 112–115, 128
immunogenicity assessment, therapeutic indications, 513
protein particulates
additional particle detection, 369
Amgen's submission, 370
antigenic proteins, 368
light obscuration analysis, 370, 375
particle characterization methods, 370
safety and efficacy, 369
similarity assessment, 371–374

totality of evidence, 370
USP 788, 369
USP 1787, 369
Food and Drug Regulations, 125
Food and Drugs Act, 125
Food and Drugs Control Administration (FDCA), 660
Formulation development
QbD and, 214–215
CFAs, 215
control strategy, 216
design space, 216
formulation capability and continual improvement, 216–217
prior knowledge, 217
risk assessments, 216
Fortical, 113
Fourier transform infrared (FTIR) spectroscopy, 331–332, 352
4PL model, 417
FTIR spectroscopy, 324
Fulphila, 242
Functional assays, 153, 181

G
Generic/automatic substitution, 139
Genetic Engineering Approval Committee (GEAC), 660
Global biosimilar market, 656–657
Glycosylated molecules, 180
Glycosylation, 174, 176, 314
GMP compliant software package, 417
Good manufacturing practices (GMP), 26
Good manufacturing processes (GMP), 684
Granulocyte colony stimulating factor (G-CSF), 664
Grastofil, 230

H
Harmonization of biosimilars, 15
Hatch-Waxman Act, 4, 85
Health Canada, 126, 128, 130, 132, 138, 142
Healthcare Common Procedure Coding System (HCPCS), 51
Herceptin®, 241
Herceptin (Trastuzumab), 343
Herceptin vial, 247
Higher order structure (HOS) methods, 340
current industry landscape, 322–324
FDA and EMA guidance, 322
industry perception, 324–326
method strategy
circular dichroism, 331
1-D NMR, 332–333
2-D NMR, 333
DSC, 332
FTIR spectroscopy, 331–332
HDX, 333
intrinsic fluorescence, 334
qualifying HOS methods, 330
UV CD, 330
X-ray crystallography, 334
polymorphism (*see* HOS polymorphism)
selection considerations
fingerprint-like similarity, 328–330
FTIR spectroscopy, 326
hypothetical relationship, 327, 329
method capabilities, 326
risk assessment models, 326
risk mitigation, 326
spectroscopic methods, 326
value and cost, 327, 328
High performance size-exclusion chromatography (HP-SEC), 685
Hospira, 93
HOS polymorphism
chaperone proteins, 352
four different molecules, 355
mAb HOS stability, 352
8M urea-treated sample, 353
PCA technology
chaperone proteins, 352
four different molecules, 355
mAb HOS stability, 352
8M urea-treated sample, 353
refolding process, 354
serum-derived human IgG, 353
three different molecules, 354
refolding process, 354
serum-derived human IgG, 353
three different molecules, 354
Host-cell DNA (HCDNA), 174
Host-cell proteins (HCP), 174, 685
Human gamma globulin (HGG), 366
Human MHC Class II proteins, 491
Human serum albumin (HSA), 367, 684
Humira®, 240, 250, 380, *see also* ABP499
Humira (Adalimumab), 344
Hydrogen-Deuterium exchange (HDX), 333
Hylenex, 113

I
ICH Q8(R2), 306
Immune system, 159–161
Immunex, 96
Immunex Corp. v. Sandoz, Inc, 92
Immunogenicity, 162, 222
  clinical development, 135–137
  European perspective, biosimilars, 159–161
Immunogenicity assessment
  acceptable differences
    conditions, 507
    etanercept, 510–512
    FDA guidance, 506
    infliximab, 507–510
    rituximab, 512–513
  cell-mediated, 490
  clinical evidence, 530
  confounding patient-related variables, control of, 518–519
  EU and FDA guidance, 495
  evaluation parameters
    efficacy, 504–505
    immune response parameters, 500–503
    PK and PD parameters, 504
    safety, 505–506
  humoral elements, 490
  immunogenic biosimilar candidates, 532
  individual product quality attributes, 531
  infliximab biosimilar, 517
  in vitro assays, 531
  mitigation of risk, 493–494
  multi-disciplinary exercise, 490
  non-clinical evaluation, 494–495
  post-authorization risk mitigation
    control of risks, originator products, 527–528
    independent post-authorization life-cycles, uncertainty, 529–530
  precision profiling, 531
  product quality variables, 492–493
  reference product, risk profile for
    abatacept, 500
    adalimumab, 498–499
    darbepoetin, 499–500
    immunogenicity-related risks, 496
    pegfilgrastim, 497–498
  risk-based approach, 490
  rituximab biosimilars, 517
  switching and interchangeability
    biosimilarity, 519–520
    caution, 523
    DANBIO registry, 523
    EU and US meaning, 519
    experts expressing opinions, 524
    NOR-SWITCH phase 4 trial, 522
    PLANETAS study, 520–521
    PLANETRA study, 521–522
    retrospective cohort study, 523
    rheumatoid arthritis patients, 523
    in US, 524–527
  therapeutic indications
    adalimumab, 514–517
    EU guidance, 514
    FDA guidance, 513
  therapeutic proteins, 491–492
Immunogenicity package, 161
Immunogenicity testing, 118, 137
Imraldi, 230
Inbiopro Solutions Pvt. Ltd., 658
Indication-specific assays, 153
Industry M4Q, 306
Inflammatory bowel disease (IBD), 11–12, 162
Inflectra®, 230, 238, 244
Infliximab, 11, 69, 136, 141, 153, 162, 507–510, *see* CT-P13
Inherent particles, 365
Inhixa, 231
Inst., Inc. v. Iancu, 101
Institute for Clinical Effectiveness and Research (ICER), 71
Institutional Animal Ethics Committee (IAEC), 660
Institutional Biosafety Committee (IBSC), 660
Intas Biopharmaceuticals, 658
Intellectual property, 223
Interchangeability, 12–15, 118, 519
  biosimilar biologic drug approvals, 139–141
  biosimilar market, 37
  European perspective on biosimilars, 163–165
  stepwise approach, 483–484
Interchangeability study design and analysis
  adverse events, 563
  analysis of the variance, 562, 563
  BPCI Act, 544
  calendar time, 570
  data analysis, 561–563
  drug supply, 565–566
  in European Union, 545–546
  FDA guidance, 545
  modeling approach, 566–567
  number of cycles, estimation, 570
  regulatory approach, 544
  switching and substitution, 545

switching study
CV estimation, 554–557
dropout rate, 557–560
duration, 550
endpoints, 547
immunogenicity sampling and PK, 550
lead-in period, 548–550
patient population, 561
PK sampling period, 547
sample size, 551–554
serum drug concentration, 549
stratification, 561
Weibull model, 559
in United States, 546
in worldwide, 546
Interest/competitive intelligence, countries of, 108–110
International Council on Harmonization (ICH), 125, 128
International Nonproprietary Name (INN), 167
Inter partes review (IPR), 98–99
estoppel, 102
legal standards, 100–101
patent owners, 103–104
petitioners, 103
petition for, 99
statistics, 102–103
timeline for, 100
Intravenous immunoglobulin (IVIG), 679
Intrinsic fluorescence, 334, 352
Intrinsic particles, 365
In vitro cell line based bioassays, 398
In vitro tissue based bioassays, 398
In vivo bioassay, 397–398
IPCA Laboratories, 658
Ixifi™, 239, 244

J

Janssen, 35–36, 93, 94
Janssen Biotech Inc. v. Celltrion Healthcare Co., 96
Janssen v. Celltrion, 97
J-code policy, 51
stakeholder considerations on, 54–56
Johnson & Johnson, 28, 217
Juvenile idiopathic arthritis (JIA), 246

K

Korean Food and Drug Administration (KFDA), 659
Korea's Ministry of Food and Drug Safety (MFDS), 659

L

Ligand binding assay (LBA), 396
Light microscopy, 383
Light obscuration (LO), 370, 375
Lipitor, 41
Litigation-related issues
Amgen Inc. v. Hospira, Inc., 94–95
ANDA, 76
biosimilar product types/naming, 77–78
BPCIA, 76–77
biosimilar's inaction, 83
biosimilar's response, 80
commercialization, 180-day notice of, 83
confidential information, biosimilar's disclosure of, 79
confidentiality provisions, 78–79
exchanged lists, 81–82
Hatch-Waxman Act, 84
infringement, act of, 78
negotiated list, 81
patent exchange/patent dance time line, 82
sponsor's initial patent list, 80
sponsor's reply, 81
BPCIA litigation, preparation for
biosimilar's preparation, 89–90
sponsor's preparation, 86–88
Hatch-Waxman safe harbor, 84–86
Immunex Corp. v. Sandoz, Inc., 92–93
information, confidential exchange provision, 95–98
IPRs, 98–99
estoppel, 102
legal standards, 100–101
patent owners, 103–104
petitioners, 103
petition for, 99
statistics, 102–103
timeline for, 100
patent dance mandatory, 90–92
Luminex-based multiplex assay, 352
Luminex multiplex system, 357–359
Luminex xMAP®Multiplex Technology Platform, 352
Luminex xMAP technology, 355–357
Lupin Ltd., 658
Lusduna, 231

M

mAb biosimilarity analysis, 343–344
mAb glycosylation, 351
Managed Care Organizations (MCOs), 62–63

Mannose, 224
Marketing authorisation (MA) process, 149–150
Marketing Authorization Application (MAA), 112
Mechanism of actions (MOAs), 119, 396
  antibody drugs, 576–577
  drug delivery carriers, 577
  efficacy and safety, 578–579
  glycosylation impacts, 577
  hormonal protein drugs, 576
  structural residual uncertainty, 577–578
Medicaid
  CPI-U adjustment, 62
  FMAP, 61–62
  MCOs, 62–63
  MDRP, 62
  PDLs, 62
  physician-administered drugs, coverage and payment for, 63
Medicaid Best Price, 62
Medicaid Drug Rebate Program (MDRP), 62
Medicare, 52
  blended J-code policy, stakeholder considerations on, 54–56
  coding and payment policy, 56–58
  Medicare Part B, 53–54
  Medicare Part D, 58–61
Medicare Administrative Contractors (MACs), 53
Medicare Modernization Act (MMA), 56
Medicare Part B
  coding, 53–54
  coverage, 53
Medicare Part D
  biosimilars impact, 59–61
  CGDP, 58
  coverage gap, 59
  donut hole, 58
Medicare Physician Fee Schedule (MPFS), 54
Method capabilities, 326
Microfluidic resistive pulse sensing (MRPS), 385–386
Ministry for Health Labour and Welfare (MHLW), 659
Monoclonal antibodies (mAbs), 340, 394
  binding, 399
Monoclonal antibody, QbD, 197–199, 204–205
Monte Carlo simulations, 269, 288
Morphology-directed Raman spectroscopy (MDRS), 383–385
Movymia, 231
Mvasi™, 231, 241
Mylan, 36, 223

N
N-acetyl neuraminic acid (NANA), 179
Nanoparticle tracking analysis (NTA), 380–381
National Drug Codes (NDCs), 51
National Health Service (NHS), 17–18
National Institute of Health and Care Excellence (NICE), 68
National regulatory agencies (NCAs), 146
Neulasta, 242
Neupogen®, 238
New Drug Application (NDA), 7
Nivestim, 232
Non–clinical development
  biosimilar biologic drug approvals, 130
  European perspective, on biosimilars, 153–155
Non-clinical testing, 117
Non-Hodgkin's lymphoma (NHL), 123
Nonmedical switching, 12
NOR-SWITCH study, 13
Norwegian Drug Procurement Cooperation, 17, 69
Notice of Compliance (NOC), 139

O
Office of Patented Medicines and Liaison (OPML), 138
Ogivri™, 241
Omnitrope, 113, 232
1-D NMR, 324, 332–333
One Factor At a Time (OFAT) studies, 193, 208
Ontruzant, 232
Optical microscopy, 383
Orthogonal method, 324
Ovaleap, 232
Overall survival (OS), 134

P
Patent dance, litigation
  biosimilar market
    Amgen v. Apotex, 34–35
    Janssen v. Celltrion, 35–36
    Sandoz v. Amgen, 32–34
  litigation-related issues, 90–92
Patented Medicines (Notice of Compliance) Regulations (PM(NOC) Regulations), 138
Patent Statute, 77
Patent Trial and Appeal Board (PTAB), 99, 102
Patient acceptance, 65

Pediatric Study Plan (PSP), 119
Pegfilgrastim, 497–498
PEGfilgrastim biosimilar, 27
Peginesatide®, 387
PEGylated proteins, 368
Peripheral blood mononucleated cells (PBMCs), 412
Peroxides, 179
Pfizer, 28, 38
Pharmaceutical and Medical Devices Agency (PMDA), 5
Pharmacodynamic studies, clinical development, 132–133
Pharmacokinetic studies, clinical development, 130–132
Pharmacopoeia, 283
Pharmacovigilance
adverse drug reactions, 633–634
AEs/risks associated with drugs
in EU, 642–643
in Japan, 643–644
in South Korea, 644
in Taiwan, 644
in Thailand, 644–645
in US, 641–642
areas of medicine, 645
causality analysis, 639
communication of risk, 639
EudraVigilance, 643
importance of, 634–635
medication errors, 634
multisource biologics market, 635–637
product quality, 633
risk determination, 639
signal detection, 637–638
unique identifiers, 640–641
World Health Organization, 645
Pharmacy benefit managers (PBMs), 64
Phosphate buffered saline, 215
Physicochemical analyses, 394
PLANETAS study, 520–521
PLANETRA study, 521–522
PNS Pharma, 42
Polysorbates, 179
Post-market monitoring
biosimilar biologic drug approvals, 142
Post-translational modifications (PTMs), 128, 174–176
Potency assays, 395, 409–410
Potential risk, 533
Preferred Drug Lists (PDLs), 62
Prefilled syringes (PFS), 249
Prior knowledge, 196, 211–212, 214
Process analytical technology (PAT), 195
Process characterization, 193
Process patent, 31
Process performance qualification (PPQ), 193
Productivity, 176
Product quality attribute assessment (PQAA), 313
Product target profile, 110
Protein A, 174
Protein attribute-process risk classification, 310
Protein-based pharmaceutical products, 306
Protein conformational array (PCA) technology
bioprocess development
Biosimilar-1 drug substance, 347
CEX and AEX column purification, 346
full panel HOS analysis of, 347, 348
heavy chain CH3 amino acid 355–379, 345
HOS analysis of, 346
'mAb heavy chain amino acid 154–179', 345
PCA ELISA, 345
CD, 340
design diagram, 341
development, 341–342
"fingerprint-like" technology, 340
fluorescence spectroscopies, 340
formulation development
HOS analysis of, 348–351
IgG2 sample, 348
IgG2 *vs.* IgG1 molecule, 349
light exposure, 351
mAb glycosylation, 351
mAb HOS stability, 348
MAb stability, 348
N-glycosylation, 352
PCA ELISA, 348
V_L and C_L domains (pAb6), 349
FTIR, 340
HOS polymorphism
chaperone proteins, 352
four different molecules, 355
mAb HOS stability, 352
8M urea-treated sample, 353
refolding process, 354
serum-derived human IgG, 353
three different molecules, 354
immunology epitope scanning principle, 340
Luminex multiplex system, 357–359
Luminex xMAP technology, 355–357

Protein conformational array (PCA) technology (*cont.*)
  mAb biosimilarity analysis, 343–344
  mAb HOS, 340
  magnetic beads-based Luminex format, 340
  microplate-based ELISA format, 340
Protein particulates
  ABP-501, 386
  anaphylaxis, 367
  biological consequence of immunogenicity, 367
  degradation processes, 364
  extrinsic, intrinsic and inherent, 365
  FDA guidance
    additional particle detection, 369
    Amgen's submission, 370
    antigenic proteins, 368
    light obscuration analysis, 370, 375
    particle characterization methods, 370
    safety and efficacy, 369
    similarity assessment, 371–374
    totality of evidence, 370
    USP 788, 369
    USP 1787, 369
  immunogenicity, 366–368
  injectable therapeutic protein, 364
  multi-dose erythropoietin, 387
  multiple orthogonal methods, 364
  novel technologies for
    MDRS, 383–385
    MRPS, 385–386
    RMM, 382–383
  pure red cell aplasia, 367
  quantify and size protein aggregates, 364
  submicron particles
    DLS, 378–380
    NTA, 380–381
  subvisible particles
    FIM, 376–377
    LO, 370, 375, 376
Protein product characterization, 307–309
Proteomics analysis, 687
Public Health Service Act (PHS Act), 13, 76, 114
Pure red cell aplasia (PRCA), 677

Q

Q9, 306
Q10, 306
Q-code, 54
Quality attributes and methods
  monoclonal antibodies, 610–613
  Zarxio® methods, 607–609
Quality by design (QbD), 314
  and analytical development, 212
    analytical method capability and continual improvement, 214
    CMA, 212–213
    control strategy, 213–214
    design space, 213
    QTAMP, 212
    risk assessments, 213
  application of, 188, 196–197
    control strategy, 209–211
    CQA development, 200–201
    CQAs, 200–202
    design space, 207–209
    QTPP, 197–199
    risk assessments, 201–207
  control strategy, 195–196
  CQA, 191–192
  CTM, 189
  definition, 189
  design space, 193–195
  elements of, 190
  and formulation development, 214–215
    CFAs, 215
    control strategy, 216
    design space, 216
    formulation capability and continual improvement, 216–217
    prior knowledge, 217
    risk assessments, 216
  QTPP, 191
  quality target product profile, 191
  risk assessments, 192–193
  SISPQ, 189
Quality target product profile (QTPP), 152, 182, 191, 197–199, 223, 306, 402
  control strategy, 209–211
  CQAs, 200–202
  design space, 207–209
  risk assessments, 201–207

R

Radio-immunoprecipitation test (RIP), 678
Raman spectroscopy, 383–385
Ranbaxy Laboratories Ltd., 658
RAND Corporation, 16
Ratiograstim, 233
Rayleigh approximation, 380

rDNA technology, 148
Rebif®, 367
Recombinant DNA Advisory Committee (RDAC), 660
Recombinant human erythropoietin (r-HuEpo)
alternation of protein, 684
biochemical modification, 676
biocopy/follow-on-biologics agents, 676
biosimilar r-HuEpo, US, 687–688
breakdown products of proteins, 676
complexity of manufacturing, 680–681
complex structure, 677
development and manufacture, 676
glycosylation and post-translational modifications, 676
GMP standards, 684
immunogenicity of, 677–679
pharmacovigilance study, Thailand
immunogenicity of biocopy, 682–683
immunogenicity risk of, 682
incidence rate of anti-r-HuEpo, 681
re-evaluation of, 681
registry enrolled subjects, 682
registry protocol, 682
product quality issues, 682–683
protein denaturation, 684
quality issues of, 685–687
risk factors of immunogenicity, 679–680
similarity assessment, 684
stabilization and formulation, 676
Reduction, Refinement and Replacement of animal experiments, 155
Reference biologic drug, 126–128
Reference drug, CMC, 174–175
Reference listed drug (RLD), 188–189, 196, 197, 202
Reference product
batch size, 268
characterization, 309–311
comparator product, 268
European perspective, biosimilars, 151
legislative provisions, 268
Monte Carlo simulation, 269
non-domestic reference product, 268
obvious factor of risk, 268
risk profile for
abatacept, 500
adalimumab, 498–499
darbepoetin, 499–500
immunogenicity-related risks, 496
pegfilgrastim, 497–498
spacing reference product procurement, 268
statistical analysis, 269
Tier 1 and Tier 2 assessments, 270–272
Tier 2 assessments, 272
Reference Standard, 410–411
Reformulation of hypotheses, Tier 1 test
equal variances, 456–457
numerical data, 458–459
simulated type I error rates, 459–460
unequal variances, 457–458
Regulatory pathway *vs.* biosimilar regulatory pathway, 603
Reimbursement landscape, 50
CMMI, 71
coding, 50–51
commercial and biosimilars, 63–65
coverage, 50
European experience
France and Belgium, 68
Germany, 67
interchangeability, 69–70
Italy and Spain, 68
Norway, 69
physician and pharmacists incentive policies, 66
regulatory environment, 66
UK, 67–68
FFS model, 70
Medicaid
CPI-U adjustment, 62
FMAP, 61–62
MCOs, 62–63
MDRP, 62
PDLs, 62
physician-administered drugs, coverage and payment for, 63
medical and pharmacy benefits, 51
Medicare, 52
blended J-code policy, stakeholder considerations on, 54–56
coding and payment policy, 56–58
Medicare Part B, 53–54
Medicare Part D, 58–61
payment, 52
PBM, 64
Remicade®, 54, 141, 238
Remicade® (infliximab), *see also* CT-P13
Remsima®, 69, 233
Renflexis®, 54, 238, 244
Reporter gene based bioassays, 398–399
Residual uncertainty, 288
Resonant mass measurement (RMM), 382–383
Restricted access barrier systems, 253
Retacrit, 233, 242, 243

Review Committee on Genetic Manipulation (RCGM), 660
Rheumatoid arthritis (RA), 134
Risk assessments
  QbD, 192–193
Risk evaluation, 193
Risk evaluation and mitigation strategies (REMS), 27
Risk management plan (RMP), 142
Ritemvia, 234
Rituxan, 28
Rituximab, 112, 123, 352–354, 512–513, 654
Rituzena (prev.Tuxella), 234
Rixathon, 234
Riximyo, 234
Robust analytical testing, 408
Robustness, 195
Roche, 28, 36, 42
Routine pharmacovigilance, 167

S
Safety, identity, strength, potency, and quality (SISPQ), 189
Sandoz, 28, 43
Sandoz v. Amgen, 32–34
Serum Institute of India Pvt. Ltd., 658
Sickness funds, 67
Silapo, 235
Silicone oil and tungsten, 249
Similar Biotherapeutic Products (SBPs), 115
Solid analytical package, 184
Soluble receptors/receptor fusion proteins, 400
Solumarv, 235
Solymbic, 236
Sponsors, 87
Standard and Poor's (S&P), 24
State Biotechnology Coordination Board (SBCB), 660
Statistical analysis software (SAS), 208
Statistical equivalence testing, 287, 288
Stepwise approach, *see* Totality of Evidence
Sterile filtration, 251
Stress agents, 293
Strides Arcolab, 658
Study sensitivity, 134
Summary of product characteristics (SmPC), 165, 166
SureClick autoinjector platform, 251
Surface plasmon resonance (SPR) assays, 396–397
Switching, 12–15, 40, 41
  biosimilar biologic drug approvals, 141–142
Switching and interchangeability
  biosimilarity, 519–520
  caution, 523
  DANBIO registry, 523
  EU and US meaning, 519
  experts expressing opinions, 524
  NOR-SWITCH phase 4 trial, 522
  PLANETAS study, 520–521
  PLANETRA study, 521–522
  retrospective cohort study, 523
  rheumatoid arthritis patients, 523
  in US, 524–527
Switching study, interchangeability
  CV estimation, 554–557
  dropout rate, 557–560
  duration, 550
  endpoints, 547
  immunogenicity sampling and PK, 550
  lead-in period, 548–550
  patient population, 561
  PK sampling period, 547
  sample size, 551–554
  serum drug concentration, 549
  stratification, 561
  Weibull model, 559

T
Target mediated drug disposition (TMDD), 581
Terminal sterilization, 251
Terrosa, 236
Tevagrastim, 236
Therapeutic mechanism of action (MoA)
  ADCs, 401
  agonistic MoA, 399
  antagonistic MoA, 399–400
  effector function of MAbs, 401
  multiple/novel MoA, 400–401
Thorinane, 237
351(k) licensure pathway, 306
Tier-based strategy, 201–202
Tier 1 biological assays, 418
Tier 2 biological assays, 419
T-lymphocytes, 491
TNF-inhibitors, 162
Totality of evidence, 8–9, 115, 473
  analytical and functional evidence, 603–604
  analytical data, 614–616
  clinical studies, 618–619, 622–624
  confirmatory safety and efficacy, 622–624
  confirmatory safety and efficacy studies, 619–620

FDA approval, 606
high similarity, 604
implicit concept, 604
information support for Non US sources, 620–621
molecular characteristics, 606
non-clinical animal testing, 616–618
novel biopharmaceuticals, 602
product and manufacturing quality, 605
quality attributes and methods
monoclonal antibodies, 610–613
Zarxio® methods, 607–609
regulatory pathway vs biosimilar regulatory pathway, 603
valuable immunogenicity data, 619
Transmissible spongiform encephalopathies (TSE), 179
Trastuzumab, 286, 654
Truxima, 237
Tufts Center for the Study of Drug Development (TCSDD), 4
Tumor necrosis factor (TNF), 478
Tumour necrosis factor (TNF) biosimilar drugs, 69
Tungsten pins, 250
Tween®, 179
2-D NMR, 324
Two one sided tests (TOST), 476

U

United States Food and Drug Administration (U.S. FDA), 262, 286–288, 659, 687
United States Pharmacopeia (USP), 180
“Use” patent, 31
US pharmacopeia (USP) monograph, 369
UV CD spectroscopy, 324

V

Valtropin, 113, 237

W

Weighted spectral difference (WSD), 331
World Health Organization (WHO), 115–116, 659

X

X-ray crystallography, 324, 334

Z

Zarxio®, 28, 37, 40, 54, 238, 245, 607–609
Zarzio, 237
Z-ave, 378
Zydus Cadila, 657