Enliven Archive
Intensifying Thoughts

Research Article | www.enlivenarchive.org | Enliven: Biosimilars and Bioavailability

Current Issues in Bioequivalence/Biosimilarity Studies

Shein-Chung Chow[1], Susan S. Chow[2*], and Annpey Pong[3]

[1]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina, USA

[2]Kansas City University College of Osteopathic Medicine, Kansas City, Missouri, USA

[3]Merck & Co., Inc., Kenilworth, NJ, USA

*Corresponding author: Susan Chow, Kansas City University College of Osteopathic Medicine, Kansas City, MO,
E-mail: susan.chow@kansascity.edu

Received Date: 01st November 2021
Accepted Date: 15th December 2021
Published Date: 26th December 2021

Citation: Chow S-C, Chow SS, Pong A. Current Issues in Bioequivalence/Biosimilarity Studies. Enliven: Biosimilars Bioavailab. 2020; 6(1): 003.

Abstract

In the United States, when an innovative (brand-name) drug product is going off patent, pharmaceutical and/or generic/biotech companies may file an abbreviated new drug application (ANDA) for approval of generic copies or a Biologic License Application (BLA) for approval of a biosimilar product of the innovative drug product. This article provides a comprehensive review of the criteria, design, and analysis for assessment of bioequivalence (generic drugs) and biosimilarity (biosimilar products). In addition, fundamental differences between the assessment of bioequivalence for generic drugs and biosimilarity for biosimilar products are compared. Some commonly encountered challenging issues and recent development in bioequivalence and biosimilarity assessment are also discussed.

Keywords: Highly variable drugs; Two one-sided tests procedure (TOST); Confidence interval approach; Switching and alternation; Complete n-of-1 trial design

3) Introduction

For traditional chemical (small molecule) drug products in the United States (US), when an innovative (brand-name) drug product is going off patent, pharmaceutical and/or generic companies may file an abbreviated new drug application (ANDA) for approval of generic copies of the brand-name drug product. In 1984, the US Food and Drug Administration (FDA) was authorized to approve generic drug products under the *Drug Price Competition and Patent Term Restoration* Act [1], which is known as the Hatch and Waxman Act. For approval of small molecule generic drug products, the FDA requires that evidence in *average* of bioavailability in terms of the rate and extent of drug absorption be provided. The assessment of bioequivalence as a surrogate endpoint for quantitative evaluation of drug safety and efficacy is based on the *Fundamental Bioequivalence Assumption* that if two drug products are shown to be bioequivalent in average bioavailability, it is assumed that they will reach the same therapeutic effect or they are therapeutically equivalent. Under the Fundamental Bioequivalence Assumption, regulatory requirements, study design, criteria, and statistical methods for assessment of bioequivalence have been well established [see, e.g., 2-7].

Unlike small molecule drug products, the *generic versions* of biologic products are viewed similar to biological drug products (SBDP). The SBDP are *not* generic drug products, which are drug products with contain active ingredient(s) *identical* to the innovative drug product. Thus, the concept for development of SBDP, which are made of living cells, is very different from that of the generic drug products for small molecule drug products. The SBDP are usually referred to as biosimilars or biosimilar products. In 2009, the FDA was authorized to approve biosimilar products under the *Biologics Price Competition and Innovation* (BPCI) Act [1] (as part of the *Affordable Care Act*). For approval of biosimilar products, FDA recommends a stepwise approach when providing the totality-of-the-evidence for the safety and efficacy of the proposed biosimilar products. The totality-of-the-evidence can be obtained through biosimilar studies for analytical similarity assessment (in terms of critical quality attributes relevant to clinical outcomes), pharmacokinetics and pharmacodynamics (PK/PD) similarity assessment (in terms of extent and rate of drug absorption), and clinical and immunogenicity similarity assessment (in terms of safety and efficacy study endpoints).

In drug research and development, it is well recognized that generic drugs and biosimilar products are fundamentally different in terms of their molecules, syntheses and structures. For example, generic drugs which contain identical active ingredient(s) are small-molecule chemical compounds, while biosimilar products are large-molecule biologic products which are made of living cells or living organisms. Additionally, generic drugs have well defined structures which are easy to characterize and relatively stable, while biosimilar products have heterogeneous structures (mixture of related molecules), which are variable and often difficult to characterize. Unlike generic drugs, biosimilar products may cause an unwanted immune response. To provide a better understanding, Table 1 summarizes fundamental differences between generic drugs and biosimilar products.

Table 1. Fundamental Differences between Generic Drugs and Biosimilar Products

Generic Drugs	Biosimilar Products
Chemical drugs	Biologic drugs
Small molecules	Large molecules
Made by chemical synthesis	Made by living organisms
Defined structure	Heterogeneous structure
	Mixture of related molecules
Easy to characterize	Difficult to characterize
Relative stable	Variable
No issue of immunogenicity	Issue of immunogenicity
Usually taken orally	Usually injected
Often prescribed by general practitioners	Usually prescribed by specialists

As a result, study endpoint, equivalence/similarity criterion, study design, and statistical analysis for assessment of biosimilarity for biosimilar products are similar but different from those for assessment of bioequivalence for generic drugs, which are well established since 1984 [4,5]. Thus, standard methods for assessment of bioequivalence for generic drugs cannot be directly applied to the assessment of biosimilarity for biosimilar products due to the fundamental differences as outlined in Table 1. In this article, our focus will not only be placed on the fundamental differences between small molecule drug products and biologic products, but also current and challenging issues surrounding quantitative evaluation of bioequivalence (for small molecule drug products) and biosimilarity (for biosimilar products).

In the next section, bioavailability assessment for generic drug products is briefly described, while a description of biosimilarity assessment for biosimilar products is given in Section 3. Section 4 provides a head-to-head comparison between bioequivalence assessment for generic drugs and biosimilarity assessment for biosimilar products. Some challenging issues that are commonly encountered in the bioequivalence/biosimilarity assessment are discussed in Section 5. Section 6 gives some concluding remarks.

Bioequivalence Assessment for Generic Drug Products

Before approving small molecule generic drug products, the FDA requires evidence of average bioequivalence in drug absorption in terms of some pharmacokinetic (PK) parameters, including the area under the blood and/or plasma concentration-time curve (AUC) and peak concentration (C_{max}), which can be provided through the conduct of bioequivalence studies. In practice, we may claim that a test drug product is bioequivalent to an innovative (reference) drug product if the 90% confidence interval for the ratio of geometric means of the primary PK parameter is completely within the bioequivalence limits of (80%, 125%). The confidence interval for the ratio of geometric means of the primary PK parameter is obtained based on log-transformed data. In the following, study designs and statistical methods that are commonly considered in bioequivalence studies are briefly described.

Study Design

As indicated in the *Federal Register* [Vol. 42, No. 5, Sec. 320.26(b) and Sec. 320.27(b), 1977], a bioavailability study (single-dose or multidose) should be crossover in design, unless a parallel or other design is more appropriate for valid scientific reasons. Thus, in practice, a standard two-sequence, two-period (or 2×2) crossover design is often used for a bioavailability or bioequivalence study. The test product and reference product are denoted by T and R, respectively. Thus, a 2×2 crossover design can be expressed as (TR, RT), where TR is the first sequence of treatments and RT denotes the second sequence of treatments. Under the (TR, RT) design, qualified subjects who are randomly assigned to sequence 1 (TR) will receive the test product T first and then receive the reference product R after a sufficient length of wash-out period. Similarly, subjects who are randomly assigned to sequence 2 (RT) will receive the reference product (R) first and then receive the test product (T) after a sufficient length of wash-out period. One of the limitations of the standard 2×2 crossover design is that it does not provide independent estimates of intra-subject variabilities since each subject will receive the same treatment only once. In the interest of assessing intra-subject variabilities, the following alternative higher-order crossover designs for comparing two drug products are often considered: (i) Balaam's design, i.e., (TT, RR, RT, TR), (ii) two-sequence, three-period dual design, e.g., (TRR,RTT), and (iii) four-sequence, four-period design, e.g., (TTRR, RRTT, TRTR, RTTR).

In addition to the assessment of average bioequivalence (ABE), there are other types of bioequivalence assessment including the population bioequivalence (PBE) which is intended to address drug prescibability and individual bioequivalence (IBE) which is intended to address drug switchability. For assessing IBE/PBE, the FDA recommends that a

replicated design be considered for obtaining independent estimates of intra-subject and inter-subject variabilities and variability due to subject-by-drug product interaction. A commonly considered replicate crossover design is the replicate of a 2×2 crossover design given by (TRTR, RTRT). In some cases, an incomplete block design or an extra-reference design such as (TRR, RTR) may be considered depending upon the study objectives of the bioavailability/bioequivalence studies [8].

Statistical Methods

As indicated earlier, ABE is claimed if the ratio of average bioavailabilities between test and reference products is within the bioequivalence limit of (80%, 125%) with 90% assurance based on log-transformed data. Along this line, commonly employed statistical methods are the confidence interval approach and interval hypotheses testing. In the confidence interval approach, a 90% confidence interval for the ratio of means of the primary pharmacokinetic response such as AUC or C_{max} is obtained under an analysis of variance model. We claim bioequivalence if the obtained 90% confidence interval is totally within the bioequivalence limit of (80%, 125%). For the method of interval hypotheses testing, the interval hypotheses that sets of one-sided hypotheses. The first set of hypotheses verifies that the average bioavailability of the test product is not too low, whereas the second set of hypotheses verifies that the average bioavailability of the test product is not too high. Under the two one-sided hypotheses, Schuirmann's two one-sided tests procedure is commonly employed for testing ABE [2].

In practice, other statistical methods such as Westlake's symmetric confidence interval approach, confidence interval based on Fieller's theorem, Chow and Shao's joint confidence region approach, Bayesian methods, and non-parametric methods such as Wilcoxon-Mann- Whitney two one-sided tests procedure, distribution-free confidence interval based on the Hodges-Lehmann estimator, and bootstrap confidence interval are sometimes considered [7].

In Vitro Bioequivalence Testing

For generic approval, as indicated in 21 CFR 320.24, bioavailability and bioequivalence may be established by *in vivo* [4] and *in vitro* studies or with suitable justification by *in vitro* studies alone (FDA, 2003b). Thus, in practice, there are two types of studies that are commonly conducted for bioequivalence assessment for generic approval. These two types of bioequivalence assessment are *in vivo* studies and *in vitro* studies (see also Table 2). In vivo studies are referred to (i) PK/PD studies, (ii) bioavailability/ bioequivalence (BA/BE) studies, and (iii) clinical studies. On the other hand, *in vitro* studies are referred to dissolution test and dissolution profile comparison. *In vitro* tests include, but are not limited to, testing for content uniformity, prime/re-prime, spray pattern, plume geometry, and droplet distribution.

Table 2. Comparison between *In Vivo* and *In Vitro* Bioequivalence Testing

In Vivo BE Testing	*In Vitro* BE Testing
Drug absorption	Drug release/delivery
Healthy volunteers	Bottles
Small sample size	Large sample size
Large variability	Less variability
BE limits of (80%, 125%)	BE limits of (90%, 111%)
Not controllable - Inter- & intra-subject variabilities	Controllable - Between-batch & within-batch Between-bottle variabilities
Fundamental BE Assumption	IVIVC *In vitro* drug release/deliver is predictive of in vivo drug absorption

It should be noted that bioequivalence assessment using either *in vivo* studies or *in vitro* studies is based on the Fundamental Bioequivalence Assumption that (i) BA/BE is predictive of clinical outcomes and (ii) *in vivo* and *in vitro* correlation (IVIVC). In other words, there is a well-established correlation between *in vivo* test results and *in vitro* test results, e.g., drug release/delivery is predictive of drug absorption.

Remarks

As indicated by the regulatory agencies, a generic drug can be used as a substitution of the brand-name drug if it has shown bioequivalence to the brand-name drug. Current regulations do not indicate that two generic copies of the same brand-name drug can be used interchangeably, even if they are bioequivalent to the same brand-name drug. Bioequivalence between generic copies of a brand-name drug is not required. Thus, one of the controversial issues is whether these approved generic drug products can be used safely and interchangeably.

Biosimilarity Assessment for Biosimilar Products

As indicated earlier, the assessment of bioequivalence is possible under the Fundamental Bioequivalence Assumption. Due to the fundamental differences between the small molecule drug products and biological products (Table 1), the Fundamental Bioequivalence Assumption and the well-established standard methods may not appropriately be directly applied for assessment of biosimilarity. Based on the *Biologics Price Competition and Innovation* (BPCI) Act [1] (as part of the *Affordable Care Act*) passed by the US Congress on March 23, 2010, quantitative evaluation of biosimilarity includes the concepts of biosimilarity and drug interchangeability, which will be briefly described below.

Definition of Biosimilarity

In the BPCI Act [1], a biosimilar product is defined as a product that is *highly similar* to the reference product notwithstanding minor differences in clinically inactive components and there are no clinically meaningful differences in terms of safety, purity, and potency. Based on this definition, a biological medicine is considered biosimilar to a reference biological medicine if it is highly similar to the reference in safety, purity (quality) and efficacy. However, little or no discussion regarding that *'How similar is considered highly similar*?' in the BPCI Act is given.

The BPCI Act seems to suggest that a biosimilar product should be highly similar to the reference drug product in all spectrums of good drug characteristics such as identity, strength, quality, purity, safety, and stability. In practice, however, it is almost impossible to demonstrate that a biosimilar product is high similarity to the reference product in all aspects of the good drug characteristics in a single study. Thus, to ensure a biosimilar product is highly similar to the reference product in terms of these good drug characteristics, different biosimilar studies may be required. For example, if safety and efficacy is a concern, then a clinical trial must be conducted to demonstrate no clinically meaningful differences in terms of safety and efficacy. On the other hand, to ensure highly similar in quality, assay development/validation, process control/validation, and product specification of the reference product are necessarily established. In addition, testing for comparability in manufacturing process between biosimilars and the reference must be performed. In some cases, if a surrogate endpoint such as pharmacokinetic (PK), pharmacodynamics (PD), or genomic marker is predictive of the primary efficacy/safety clinical endpoint, then a PK/PD or genomic study may be used to assess biosimilarity between biosimilars and the reference product. It should be noted that current regulatory requirements are guided based on a case-by-case basis by the following basic principles of (i) the extent of the physicochemical and biological characterization of the product, (ii) nature or possible changes in the quality and structure of the biological product due to changes in the manufacturing process (and their unexpected outcomes), (iii) clinical/regulatory experiences with the particular class of the product in question, and (iv) several factors that need to be considered for biocomparability.

Definition of Interchangeability

As indicated in the Subsection (b)(3) amended to the Public Health Act Subsection 351(k)(3), the term *interchangeable or interchangeability* in reference to a biological product that meets the standards described in subsection (k)(4), means that the biological product may be substituted for the reference product without intervention from the health care provider who prescribed the reference product. Along this line, in what follows, definition and basic concepts of interchangeability (in terms of switching and alternating) are given.

As indicated in the Subsection (a)(2) amends the Public Health Act Subsection 351(k)(3), a biological product is considered interchangeable with the reference product if (i) the biological product is biosimilar to the reference product and (ii) it can be expected to produce the same clinical result in *any given patient*. For a biological product that is administered more than once to an individual, the risk in terms of safety or diminished efficacy of alternating or switching between use of the biological product and the reference product is not greater than the risk of using the reference product without such alternation or switch. Thus, there is a clear distinction between biosimilarity and interchangeability. In other words, biosimilarity does not imply interchangeability which is much more stringent. Intuitively, if a test product is judged to be interchangeable with the reference product then it may be substituted, even alternated, without a possible intervention, or even notification, of the health care provider. However, interchangeability is expected to produce the same clinical result in *any given patient*, which can be interpreted as that the same clinical result can be expected in *every single patient*. In reality, conceivably, lawsuits may be filed if adverse effects are recorded in a patient after switching from one product to another. It should be noted that when FDA declares the biosimilarity of two drug products, it may not be assumed that they are interchangeable. Therefore, labels ought to state whether a follow-on biologic is biosimilar to a reference product and if interchangeability has or has not been established. However, payers and physicians may, in some cases, switch products even if interchangeability has not been established.

Concepts of Switching and Alternating

Let T and R stand for a proposed biosimilar (test) product and an innovative biological (reference) product, respectively. Most researchers interpret switching as a switch from (T to R), (R to T), (R to R) or (T to T) and alternation as a switch from (R to T to R), (R to R to R), (T to R to T) or (T to T to T). In other words, switching generally refers to a switch from one product (R or T) to another (R or T), where T could be a different interchangeable biosimilar product that has been demonstrated as highly similar to the same reference product. Alternation may begin with one product (R or T) and—after a few switches—return to the same product, where T could be a different interchangeable biosimilar product that has been demonstrated as highly similar to the same reference product. The recent FDA guidance on interchangeability FDA [9] distinguished the concepts of switching and alternation for interchangeable biosimilar products. Switching generally refers to a single switch from one product to another, such as (R to T) and (R to R). On the other hand, alternation refers to multiple switches, such as (R to T to R) and (R to R to R) for two switches and (R to T to R to T) and (R to R to R to R) for three switches.

As indicated, an interchangeable biosimilar product is expected to produce the same clinical results as the reference product in any given patient with the disease under study. To determine whether the proposed interchangeable biosimilar product can produce the same clinical results in any given patient, a switching design with the nature of crossover within individual subjects is necessary [10,11]. Furthermore, an adequate switching design should be able to evaluate the potential risk of safety and efficacy (e.g., increase of severe adverse events and/or diminished efficacy) with and without switching and/or alternation [9]. In the following sections, we describe two useful switching designs: a two-sequence crossover design recommended by the FDA and a n-of-1 trial design [10-12].

The recent FDA guidance on interchangeability recommended a 2 × (m + 1) crossover design, where m is the number of switches [9]. With a single switch, i.e., m = 1, the FDA recommends that a crossover design consists of two sequences of RT and RR, denoted by (RT, RR). As can be seen, the (RT, RR) design allows the evaluation of the effect of the switch from R to T and the effect of the switch from R to R (i.e., no switch). The relative risk between the switch from R to T and no switch (i.e., the switch from R to R) can also be assessed. Note that the 2 × 2 crossover design with a single switch, i.e., (RT, RR), is a partial design of the 4 × 2 Balaam design, i.e., (RR, TT, RT, TR). When m = 2 (i.e., two switches), the FDA suggests a 2 × 3 crossover design that consists of the two sequences of RTR and RRR, denoted by (RTR, RRR). The 2 × 3 crossover design with two switches allows the evaluation of the effect of the switch from R to T, the effect of the switch from T to R, and the effect of the switch from R to R (i.e., no switch). In addition, the relative risk between the switch from R to T and no switch (i.e., the switch from R to R) and between the switch from T to R and no switch can also be assessed. When m = 3 (i.e., there are three switches), the FDA suggests a 2 × 4 crossover design that consists of the two sequences of RTRT and RRRR, denoted by (RTRT, RRRR). Similar to the 2 × 3 crossover design with two switches, the 2 × 4 crossover design with three switches allows the evaluation of the effect of the switch from R to T, the effect of the switch from T to R, and the effect of the switch from R to R (i.e., no switch). The relative risk between the switch from R to T and no switch (i.e., the switch from R to R) and the relative risk between the switch from T to R and no switch can also be assessed.

General Design for Biosimilar/Interchangeable Studies

In recent years, the n-of-1 trial design has become a very popular design for evaluating the difference in treatment effect within the same individual when n treatments are administered at different dosing periods [10,11]. An n-of-1 trial is a clinical trial in which a single subject is the entire trial. A trial in which random allocation can be used to determine the order in which a test treatment and a control (e.g., a standard of care or an active control agent) are given to a subject is an n-of-1 randomized controlled trial. Thus, the n-of-1 trial design is in fact a crossover design. Following similar ideas for switching designs with single switch and/ or multiple switches, Chow et al. [12] proposed the use of a so-called complete n-of-1 trial design to assess the relative risk between switching/alternation and without switching/ alternation.

The construction of a complete n-of-1 trial design depends on m, the number of switches. For example, if m = 1 (single switch), the complete n-of-1 trial design consists of m + 1 = 2 periods. Each dosing period involves two choices (i.e., either R or T) and, thus, a total of $2^{m+1}=2^2=4$ sequences (i.e., combinations of R and T). This results in a 4 × 2 Balaam design, i.e., (RR, TT, RT, TR). When m = 2 (two switches), the complete n-of-1 trial design consists of m + 1 = 3 periods. Each dosing period involves two choices (i.e., either R or T) and, thus, a total of $2^{m+1}=2^3=8$ sequences. This results in an 8 × 3 crossover design. Similarly, where there are three switches (i.e., m = 3), the complete n-of-1 trial design consists of m + 1 = 4 periods. Each dosing period involves two choices (i.e., either R or T) and, thus, a total of $2^{m+1}=2^4=16$ sequences (i.e., combinations of R and T). This results in a 16 × 4 crossover design. To assist with understanding, Table 3 lists a complete n-of-1 trial design with m = 1 (single switch), m = 2 (two switches) and m = 3 (three switches) that maybe useful for biosimilar switching studies.

As can be seen from Table 3, the switching designs with a single switch, i.e., (RT, RR), two switches, i.e., (RTR, RRR), and three switches, i.e., (RTRT, RRRR), are partial designs of the complete n-of-1 trial designs with a single switch (two periods), two switches (three periods) and three switches (four periods), respectively.

Table 3. Complete n-of-1 Trial Design with m=1,2,and 3

Group	Period I	Period II	Period III	Period IV
1	R	R	R	R
2	R	T	R	R
3	T	T	R	R
4	T	R	R	R
5	R	R	T	R
6	R	T	T	T
7	T	R	T	R
8	T	T	T	T
9	R	R	R	T
10	R	R	T	T
11	R	T	R	T
12	R	T	T	R
13	T	R	R	T
14	T	R	T	T
15	T	T	R	T
16	T	T	T	R

m=1 (single switch with two periods), m=2 (two switches with three periods), m=3 (three switches with four periods)

6.2.2 Is 80% to 125% a reasonable margin?

With small molecule drug products, bioequivalence generally reflects therapeutic equivalence. Drug prescribability, switching, and alternating are generally considered reasonable. With biologic products, however, variations are often higher (other than pharmacokinetic factors which may be sensitive to small changes in conditions). Thus, only parallel-group design rather than crossover kinetic studies can often be performed. It should be noted that very often, with follow-on biologics, biosimilarity does *not reflect* therapeutic comparability. Therefore, switching and alternating should be pursued only with substantial caution.

A Comparison between Bioequivalence and Biosimilarity Studies

As indicated in Table 1, there are fundamental differences between generic drugs and biosimilar products. As a result, study endpoint, criteria for BE and/or BS assessment, study design, and statistical method for data analysis are similar but different.

Fundamental Biosimilarity Assumption

Similar to Fundamental Bioequivalence Assumption for bioequivalence assessment, Chow et al. [13] proposed the following *Fundamental Biosimilarity Assumption* for follow-on biologics:

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

This Fundamental Biosimilarity Assumption is based on the assumptions that (i) analytical similarity is predictive of PK/PD similarity and (ii) PK/PD similarity is predictive of clinical similarity, i.e., Fundamental Bioequivalence Assumption (see also Figure 1). It should, however, be noted that the Fundamental Biosimilarity Assumption is difficult, if not impossible, to verify in practice.

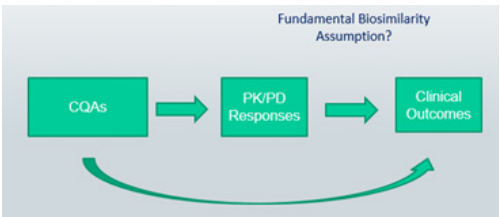
Study Endpoint

For approval of generic drugs, a bioequivalence trial is required to demonstrate that the rate and extent of drug absorption in the bloodstream of the test product is similar to that of the reference product. Thus, the study endpoints considered in bioequivalence trials are AUC (area under the blood concentration time curve) which is the measurement of the extent of drug absorption and C_{max} (peak concentration) which is the measurement of the rate of absorption. On the other hand, for approval of biosimilar products, FDA requires that analytical similarity, PK/PD similarity, and clinical similarity be established for providing totality-of-the-evidence in support of similarity between the test product and the reference product. For assessment of analytical similarity, PK/PD similarity, and clinical similarity, different study endpoints are used. For example, critical quality attributes (CQAs) that relevant clinical outcomes are necessarily assessed for analytical similarity, while safety, efficacy, and/or immunogenicity response are evaluated for clinical similarity.

A Comparison

To provide a better understanding, Table 4 summarizes comparison between *in vivo* bioequivalence testing and biosimilarity assessment in terms of study endpoint, associated variability, criterion (i.e., bioequivalence limit or biosimilarity margin), study design, and regulatory requirement.

Figure 1. Relationship between *in vitro* testing and *in vivo* testing



Note: The question mark indicates that regulatory agency has not yet confirmed this fundamental assumption.

Table 4. Comparison between *In Vivo* BE Testing and Biosimilarity Testing

Characteristics	*In Vivo* BE Testing	Biosimilarity Testing
Study endpoint	Drug absorption	Drug safety/efficacy
Variability	20% - 30%	40% - 50%
Criterion	(80%, 125%)	SABE[1] (proposal)
Study design	Crossover	Parallel/crossover
Statistical method	TOST[2] or CI[3]	TOST or CI
Regulatory requirement	BE trial	Quality, purity, and efficacy

[1]SABE=scaled average bioequivalence
[2]TOST=two one-sided tests procedure
[3]CI=90% confidence interval approach

Current and Challenging Issues

In this section, there are several challenging issues that are commonly encountered when performing bioequivalence testing or biosimilarity assessment. These challenging issues include, but are not limited to (i) the mixed use of TOST and 90% CI approach, (ii) margin selection for bioequivalence/biosimilarity assessment, (iii) the issue of drug interchangeability, and (iv) bridging study with multiple references, which will be briefly described below.

Mixed use of TOST and 90% CI approach

One of the major challenges in bioequivalence assessment for generic drugs and biosimilarity assessment for biosimilar products is starting with the official method for interval hypotheses testing, i.e., Schuirmann's two one-sides tests procedure (TOST) and ends up with a 90% confidence interval approach for evaluation of bioequivalence and biosimilarity. Chow and Zheng [14] pointed out that TOST is a size α test and is operationally equivalent to a 90% confidence interval (CI) approach in some cases. However, TOST is not equivalent to the 90% CI approach in general, especially when the study endpoint is binary response variable. Thus, it is suggested that TOST and the 90% CI approach should not be mixed up when performing either bioequivalence testing or biosimilarity assessment regardless of the study design used.

Margin Selection for Bioequivalence/Biosimilarity

For approval of generic drugs which contain identical active ingredient, one-size-fits-all bioequivalence limit (margin) of (80%, 125%) after log-transformation is often used. Unlike generic drugs, biological products are made by living cells or organisms with much larger variation. In this case, the one-size-fits-all criterion may not applicable. For approval of biosimilar products, FDA recommends a stepwise approach including assessment of analytical similarity, PK/PD similarity, and clinical similarity used for providing the totality-of-the-evidence in support of regulatory evaluation and approval. As analytical similarity, PK/PD similarity, and clinical similarity assessment will be performed based on different study endpoints with different variabilities, the selection of biosimilarity margin has become very critical. In practice, it is suggested that the assessment of biosimilarity of biosimilar products should take into consideration variability in addition to average as standard in bioequivalence testing for small molecule drugs. Zhang et al. [15] explored the impact of variability on similarity margin for biosimilarity assessment. On the basis of the derived relationship between variability and similarity margin that result in the same power given all other parameters fixed, Zhang et al. [15] proposed several scaled biosimilarity margins to incorporate highly variable biological products.

To provide a better understanding, Table 5 summarizes some suggested bioequivalence and biosimilarity margins based on the study of the relationship between the variability and the scaled margins for bioequivalence/biosimilarity testing [15].

Table 5. Variability Versus Margin

Type	Variability	Margin
In vitro BE testing	<10%	(90%, 111%)
In vivo BE testing	20%-30%	(80%, 125%)
Highly variable drugs	>30%	(70%, 143%) or SABE[1]
Biosimilarity testing	40%-50%	SABE (proposal)

[1]SABE=scaled average bioequivalence

The issue of drug interchangeability

Interchangeability of drug products has very different features with small molecules and with biologicals. With small-molecule drugs, a statement of bioequivalence generally indicates therapeutic equivalence and interchangeability. In contrast, with the much more sensitive and complicated biological drugs, a declaration of biosimilarity emphatically does not imply that a patient could be switched from one product to another. Both formulations may be prescribed and administered to subjects who have not yet received the drug in any of its forms. However, regulatory agencies have been very cautious about enabling and permitting interchangeability.

For generic drugs, the concept of drug interchangeability includes prescribability and switchability. Drug prescribability is usually referred to as the physician's choice for prescribing an appropriate drug for his or her patients between the brand-name drug and its generic copies, while drug switchability is referred to as the switch of a drug (e.g., a brand-name drug or its generic copies) to an alternative drug (e.g., a generic copy) within the same subject for whom the concentration of the drug has been titrated to a steady, efficacious, and safe level. To address drug prescribability and drug switchability, it is suggested that, in addition to average bioequivalence, population bioequivalence and individual bioequivalence, which account for both of average and variability of bioavailability, be established, respectively [7].

For biosimilar products, on the other hand, the concept of drug interchangeability includes switching and alternation. FDA defines an interchangeable biosimilar product as follows. A biological product is considered to be interchangeable with the reference product if (A) the biological product (i) is biosimilar to the reference product; and (ii) can be *expected* to produce the same clinical result in any given patient; and (B) for a biological product that is administered more than once to an individual, the risk in terms of safety or diminished efficacy of alternating or switching between use of the biological product and the reference product is not greater than the risk of using the reference product without such alternation or switch.

Thus, there is a clear distinction between biosimilarity and interchangeability in biosimilarity assessment. In other words, biosimilarity does not apply interchangeability. According to the FDA's definition, the major challenging

issue is that in practice, it is *not* possible to show same clinical result in any given patient. In other words, it is not possible for use to show the same clinical result for every patient. Chow and Liu [16] and Chow [17] interpreted BPCI's definition as it is possible to show same clinical result in any given patient *with certain assurance.*

Bridging Study with Multiple References

When there are multiple reference products, e.g., EU (European Union)-approved product and US (United States)-licensed product, a PK/PD bridging study is often conducted in order to bridge the clinical data from the original region (e.g., EU) to the new region (e.g., US) in support of the biosimilar regulatory submission in the new region. The purpose is to avoid duplicated clinical trials for clinical similarity between a proposed biosimilar product and the reference product in the new region provided that there is no ethnic concern in the two regions. For assessment of the bridging study, FDA recommends that head-to-head pairwise comparisons (i.e., Test product vs US-licensed product, Test vs EU-approved product, and EU-approved vs US-licensed product) be performed for biosimilarity assessment. As pointed out by an Oncologic Drugs Advisory Committee (ODAC) member at an ODAC meeting held on July 12-13 at the FDA in Silvery Spring, the pairwise comparisons suffer from the disadvantages of (i) each comparison of the pairwise comparisons dose not fully utilize data collected from the study and (ii) not all of the comparisons use the same reference. To address these issues, alternatively, Zheng et al. [18] proposed an approach based on simultaneous confidence interval for biosimilarity assessment in bridging studies with multiple references.

Concluding Remarks

As indicated earlier, we claim that a test drug product is bioequivalent to a reference (innovative) drug product if the 90% confidence interval for the ratio of means of the primary PK parameter is totally within the bioequivalence limits of (80%, 125%). This one size-fits-all criterion only focuses on average bioavailability and ignores heterogeneity of variability. Thus, it is not scientifically/statistically justifiable for assessment of biosimilarity of follow-on biologics. In practice, it is then suggested that appropriate criteria, which can take the heterogeneity of variability into consideration, be developed since biosimilars are known to be variable and sensitive to small variations in environmental conditions [13,16,19].

At the FDA public hearing, questions that are commonly asked are "*How similar is considered similar?"* and "*How should the degree of similarity be measured and translated to clinical outcomes (e.g., safety and efficacy)?"* These questions are closely related to drug interchangeability of biosimilars or follow-on biologics which have been shown to be biosimilar to the innovative product [20,21].

For assessment of bioequivalence for chemical drug products, a crossover design is often considered, except for drug products with relatively long half-lives. Since most biosimilar products have relatively long half-lives, it is suggested that a parallel group design be considered. However, parallel group design does not provide independent estimates of variance components such as inter- and intra-subject variabilities and variability due to subject-by-product interaction. Thus, it is a major challenge for assessing biosimilars under parallel group designs. Although EMA of EU has published several product-specific guidance based on the concept papers [22-30], it has been criticized that there are no objective *standards* for assessment of biosimilars because it depends upon the nature of the products. Product-specific standards seem to suggest that a flexible biosimilarity criterion should be considered and the *flexible* criterion should be adjusted for variability and/or the therapeutic index of the innovative (or reference) product.

As described above, there are many uncertainties in assessing the biosimilarity and interchangeability of biosimilars. As a result, it is a major challenge to both clinical scientists and biostatisticians to develop valid and robust clinical/statistical methodologies for assessment of biosimilarity and interchangeability under the uncertainties. In addition, addressing the issues of quality and comparability in manufacturing process is another challenge for both the pharmaceutical scientists and biostatisticians. The proposed general approach derived from the biosimilarity index is based on the concept of reproducibility.

References

1. BPCI. Biologic Price, Competition, and Innovation Act. United States Congress. Washington DC. 2009.

2. 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.

3. EMEA. Note for Guidance on the Investigation of Bioavailability and Bioequivalence. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/EWP/QWP/1401/98, London, United Kingdom. 2001.

4. FDA. Guidance for Industry – Bioavailability and Bioequivalence Studies for Orally Administrated Drug Products – General Considerations, Center for Drug Evaluation and Research, the US Food and Drug Administration, Rockville, Maryland, USA. 2003a.

5. FDA. Guidance for Industry – Bioavailability and Bioequivalence Studies for Nasal Aerosols and Nasal Sprays for Local Action, Center for Drug Evaluation and Research, the US Food and Drug Administration, Rockville, Maryland, USA. 2003b.

6. WHO. World Health Organization Draft Revision on Multisource (Generic) Pharmaceutical Products: Guidelines on Registration Requirements to Establish Interchangeability, Geneva, Switzerland. 2005.

7. Chow SC, Liu JP. Design and Analysis of Bioavailability and Bioequivalence Studies. 3rd edition, Chapman Hall/CRC Press, Taylor & Francis, New York. 2008.

8. Chow SC, Shao J, Wang H. Individual bioequivalence testing under 2x3 crossover designs. Stat Med. 2002, 21: 629-648.

9. FDA. Guidance for industry – Considerations in Demonstrating Interchangeability with a Reference Product. The United States Food and Drug Administration, Silver Spring. 2019.

10. Lillie EO, Patay B, Diamant J, Issell B, Topol E, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med. 2011, 8: 161-173.

11. Davidson KW, Cheung YK, McGinn T, Wang YC. Expanding the role of n-of-1 trials in the precision medicine era: action priorities and practical consideration. Commentary. National Academy of Medicine, Washington DC. 2018.

12. Chow SC, Song FY, Cui C. On hybrid parallel–crossover designs for assessing drug interchangeability of biosimilar products. J Biopharm Stat. 2017, 27: 265-271.

13. Chow SC, Hsieh TC, Ch, E, Yang J. A comparison of moment-based and probability-based criteria for assessment of follow-on biologics. J Biopharm Stat. 2010, 20: 31-45.

14. Chow SC, Zheng J. The use of 90% CI or 95% CI in biosimilar product development – a controversial issue? J Biopharm Stat. 2019, 29: 834-844.

15. Zhang N, Yang J, Chow SC, Chi E. Impact of variability on the choice of biosimilarity limits in assessing follow-on biologics. Stat Med. 2013, 32: 424-433.

16. Chow SC, Liu JP. Statistical assessment of biosimilar products. J Biopharm Stat. 2010, 20: 10-30.

17. Chow SC. Biosimilars: Design and Analysis of Follow-on Biologics. Chapman and Hall/CRC Press, Boca Raton. 2013.

18. Zheng J, Yin D, Yuan M, Chow S.C. Simultaneous confidence interval approach for analytical similarity assessment. J Biopharm Stat. 2019, 29, 920-940.

19. Hsieh TC, Chow SC, Liu JP, Hsiao CF, Chi E. Statistical test for evaluation of biosimilarity of follow-on biologics. J Biopharm Stat. 2010, 20: 75-89.

20. Roger SD. Biosimilars: How similar or dissimilar are they? Nephrology (Carlton). 2006, 11: 341-346.

21. Roger SD, Mikhail A. Biosimilars: opportunity or cause for concern? J Pharm Sci. 2007, 10: 405- 410.

22. EMEA. Note for Guidance on Comparability of Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Non Clinical and Clinical Issues. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/3097/02, London, United Kingdom. 2003a.

23. EMEA. Rev. 1 Guideline on Comparability of Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Quality Issues. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/ BWP/3207/00/Rev 1, London, United Kingdom. 2003b.

24. EMEA. Guideline on Similar Biological Medicinal Products. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/437/04, London, United Kingdom. 2005a.

25. EMEA. Draft Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance: Quality Issues. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/49348/05, London, United Kingdom. 2005b.

26. EMEA. Draft Annex Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Non Clinical and Clinical Issues – Guidance on Biosimilar Medicinal Products containing Recombinant Erythropoietins. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/94526/05, London, United Kingdom. 2005c.

27. EMEA. Draft Annex Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Non Clinical and Clinical Issues – Guidance on Biosimilar Medicinal Products containing Recombinant Granulocyte-Colony Stimulating Factor. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP /31329 /05, London, United Kingdom. 2005d.

28. EMEA. Draft Annex Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Non-Clinical and Clinical Issues – Guidance on Biosimilar Medicinal Products containing Somatropin. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/94528/05, London, United Kingdom. 2005e.

29. EMEA. Draft Annex Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Non Clinical and Clinical Issues – Guidance on Biosimilar Medicinal Products containing Recombinant Human Insulin. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/32775/05, London, United Kingdom. 2005f.

30. EMEA. Guideline on the Clinical Investigating of the Pharmacokinetics of Therapeutic proteins. The European Medicines Agency Evaluation of Medicines for Human Use. EMEA/CHMP/89249/04, London, United Kingdom. 2005g.

31. Haidar SH, Davit B, Chen ML, Conner D, Lee L, et al. Bioequivalence approaches for highly variable drugs and drug products. Pharm Res. 2008, 25: 237-241.