Enliven Archive
Intensifying Thoughts

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

On Totality-of-the-Evidence in Biosimilar Product Development

Chow SC[1], Lee SJ[2]*

[1] *Duke University School of Medicine, Durham, North Carolina, USA.*
[2] *Celltrion Inc., Incheon, Republic of Korea.*

[*]Corresponding Author: Shein-Chung Chow PhD, Professor, Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina, USA. E-mail: sheinchung.chow@duke.edu

Received Date: 08th August 2019
Accepted Date: 30th August 2019
Published Date: 31st August 2019

Citation: Chow SC, Lee SJ. On Totality-of-the-Evidence in Biosimilar Product Development. Enliven: Biosimilars Bioavailab.2019; 3(1): 001-004.

Abstract

For approval of a proposed biosimilar product, the United States Food and Drug Administration (FDA) requires totality-of-the-evidence be provided to support a demonstration of biosimilarity between the proposed biosimilar product and the US-licensed drug product. However, as indicated in Section 314.126 of 21 CFR (Code of Federal Regulation) that substantial evidence needs to be provided to support the claims of new drugs. In practice, although there is no clear distinction between the concept of substantial evidence (with legal basis) in new drug development and the concept of totality-of-the-evidence (without legal basis) in biosimilar drug product development, it is a concern whether the totality-of-the-evidence in terms of analytical similarity, pharmacokinetic (PK) and pharmacodynamic (PD) similarity and clinical similarity can provide substantial evidence to support the demonstration of biosimilarity between the proposed biosimilar product and the US-licensed drug product. A couple of recent regulatory submissions were presented to demonstrate the concern.

Keywords: Substantial Evidence; Stepwise Approach; Analytical Similarity; PK/PD Similarity; Clinical Similarity.

Introduction

For approval of a proposed biosimilar product, the United States (US) Food and Drug Administration (FDA) requires that totality-of-the-evidence be provided to support a demonstration that the proposed biosimilar product is highly similar to the US-licensed product, notwithstanding minor differences in clinically inactive components, and that there are no clinically meaningful differences between the proposed biosimilar product and the US-licensed product in terms of the safety, purity and potency of the product.

To assist the sponsor in biosimilar product development, FDA recommends a stepwise approach for obtaining the totality-of-the-evidence for demonstrating biosimilarity between the proposed biosimilar product and its innovative drug product in terms of safety, purity, and efficacy [1-5]. The stepwise approach starts with similarity assessment in critical quality attributes (CQAs) in analytical studies, followed by the similarity assessment in pharmacological activities in pharmacokinetic and pharmacodynamic (PK/PD) studies and similarity assessment in safety and efficacy in clinical studies. For analytical similarity assessment in CQAs, FDA further recommends tiered approach which classifies CQAs into three tiers depending upon their criticality or risk ranking relevant to clinical outcomes. For determination of criticality or risk ranking, FDA suggests establishing a predictive (statistical) model based on either mechanism of action (MOA) or PK relevant to clinical outcome. Thus, the following assumptions are made for the stepwise approach for obtaining the totality-of-the-evidence.

(1) Analytical similarity is predictive of PK/PD similarity;
(2) Analytical similarity is predictive of clinical outcomes;
(3) PK/PD similarity is predictive of clinical outcomes.

These assumptions, however, are difficult (if not impossible) to verify in practice. For assumptions (1) and (2), although many *in vitro* and *in vivo* correlations (IVIVC) have been studied in the literature, the correlations between specific CQAs and PK/PD parameters or clinical endpoints are not fully studied and understood. In other words, most predictive models are not well established or are established but not validated. Thus, it is not clear how a (notable) change in a specific CQA can be translated to a change in drug absorption or clinical outcome. For (3), unlike bioequivalence assessment for generic drug products, there does not exist *Fundamental Biosimilarity Assumption* indicating that PK/PD similarity implies clinical similarity in terms of safety and efficacy. In other words, PK/PD similarity or dis-similarity may or may not lead to clinical similarity. Note that the assumptions (1) and (3) combined does not lead to the validity of assumption (2) automatically. The validity of assumptions (1)-(3) is critical for the success of obtaining total-

ity-of-the-evidence for assessing biosimilarity between the proposed biosimilar and the innovative biological product. This is because the validity of these assumptions ensures the relationships among analytical, PK/PD, and clinical similarity assessment and consequently the validity of the overall biosimilarity assessment. Table 1 illustrates relationships among analytical, PK/PD, and clinical assessments in the stepwise approach for obtaining the totality-of-the-evidence in biosimilar product development.

Relationships among Analytical, PK/PD, and Clinical Similarity

Relationships among CQAs, PK/PD responses, and clinical outcomes can be described in Figure 1. In practice, for simplicity, CQAs, PK/PD responses, and clinical outcomes are usually assumed linearly correlated. For example, let x, y, and z be the test result of a CQA, PK/PD response, and clinical outcome, respectively. Under assumptions (1)-(3), we have

(1) $y = a_1+b_1x+e_1$;
(2) $z = a_2+b_2y+e_2$;
(3) $z = a_3+b_3x+e_3$;

where e_1,e_2, and e_3 follow a normal distribution with mean 0 and variances σ_1^2,σ_2^2, and σ_3^2, respectively. In practice, each of the above models is often difficult, if it is not impossible, to be validated due to lack of insufficient data collected during the biosimilar product development. Under each of the above models, we may consider the criterion for examination of the closeness between an observed response and its predictive value to determine whether the respective model is a good predictive model. As an example, under model (1), we may consider the following two measures of closeness, which are based on either the *absolute* difference or the *relative* difference between an observed value y and its predictive value $\hat{y}$.

Criterion I. $p_1 = P\{|y-\hat{y}|<\delta\}$,
Criterion II. $p_2 = P\{|(y-\hat{y})/y|<\delta\}$.

It is desirable to have a high probability that the difference or the relative difference between y and $\hat{y}$, given by p_1 and p_2, respectively, is less than a clinically meaningful difference δ.

Suppose there is a well-established relationship between x (e.g., test results of a given CQA) and y (e.g., PK/PD response). Model (1) indicates that a change in CQA, say Δ_x corresponds to a change of $a_1+b_1\Delta_x$ in PK/PD response. Similarly, model (2) indicates that a change in PK/PD response, say Δ_y corresponds to a change of $a_2+b_2\Delta_y$ in clinical outcomes. Models (2) and (3) allows us to evaluate the impact of the change in CQA (i.e.,x) on PK/PD (i.e., y) and consequently clinical outcome (i.e., z).

Under models (2) and (3), we have

$a_2+b_2y+e_2 = a_3+b_3x+e_3$.

This leads to

$a_1 = (a_3-a_2)/b_2$, $b_1 = b_3/b_2$, *and* $e_1 = (e_3-e_2)/b_2$.

with

$b_2^2\,\sigma_1^2 = \sigma_3^2+\sigma_2^2$

or

$\sigma_1 = 1/b_2\,\sqrt{(\sigma_2^2 + \sigma_3^2)}$.

In practice, the above relationships can be used to verify primary assumptions as described in the previous section provided that models (1)-(3) have been validated. Suppose models (1)-(3) are well-established, validated, and fully understood. A commonly asked question is whether PK/PD studies and/or clinical studies can be waived if analytical similarity and/or PK/PD similarity have been demonstrated. Note that the above relationships hold only under linearity assumption. When there is a departure from linearity in each one of models (1)-(3), the above relationships are necessarily modified. Considering multiple CQAs and several endpoints in PK/PD and clinical outcomes, the model (1)-(3) can be easily extended to general linear models such as

(4). $Y = B_1X+E_1$;
(5). $Z = B_2Y+E_2$;
(6). $Z = B_3X+E_3$;

where, E_1, E_2 and E_3 follow a multivariate normal distribution, $N(0, \sigma_1^2 I)$, $N(0, \sigma_2^2 I)$, *and* $N(0, \sigma_3^2 I)$, respectively.

Then, we have

$B_1 = B_2^{-1}\,B_3$,

With

$\sigma_1^2\,I = (\sigma_2^2+\sigma_3^2)\,B_2^{-1}$,

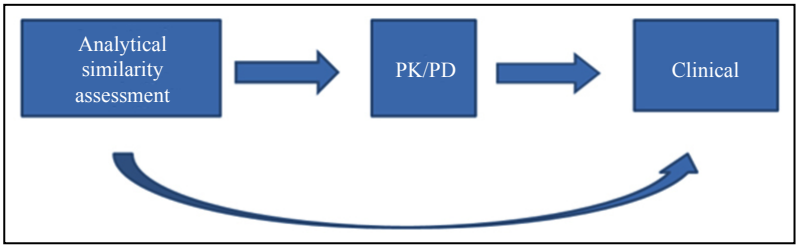
provided by

$B_1 = (X'X)^{-1}\,X'Y$,
$B_2 = (Y'Y)^{-1}\,Y'Z$,

and

$B_3 = (X'X)^{-1}\,X'Z$.

Figure 1. Relationships among analytical, PK/PD, and clinical assessment.



The existence of unique solution depends on the rank of matrices, X and Y. One way to obtain those solutions is to use numerical computations. In this case, no clinical meaningful difference might be obtained if the minimum of the $P\{ norm(Z-Z(X'X)^{-1} X'Z)<\delta\}$ and $P\{norm(Z-Z(X'X)^{-1} Y'Z)<\delta\}$ is sufficiently large.

Practical Issues

For biosimilar product development and regulatory review and approval, FDA recommends a stepwise approach by performing analytical similarity assessment, PK/PD similarity test, and clinical similarity assessment in terms of safety, tolerability, and efficacy for obtaining the totality-of-the-evidence. FDA's recommended stepwise approach focuses on three major domains, namely, analytical, PK/PD, and clinical similarity, which highly correlated under models (1)-(3). Some pharmaceutical scientists interpret the stepwise approach as a scoring system (perhaps, with appropriate weights) that includes the domains of analytical, PK/PD, and clinical similarity assessment. In this case, the totality-of-the-evidence can be assessed based on information regarding biosimilarity obtained from each domain. In practice, for each domain, we may consider either FDA's recommended binary response (i.e., similar or dis-similar) or the use of the concept of biosimilarity index Chow et al., 2011 [6] to assess similarity information and consequently the totality-of-the-evidence across domains. For the FDA's recommended approach, Table 1 provides possible scenarios when performing analytical similarity assessment, PK/PD similarity test, and clinical similarity assessment. As it can be seen from Table 1, if the proposed biosimilar product passes similarity test in all domains, FDA considers the sponsor has provided totality-of-the-evidence for demonstration of highly similarity between the proposed biosimilar and the innovative biological product. On the other hand, if the proposed biosimilar product fails to pass any of the suggested similarity assessments (i.e., analytical similarity, PK/PD similarity, and clinical similarity), then regulatory agency will reject the proposed biosimilar product. In practice, it is uncommon to see that the proposed biosimilar may fail in one of the three suggested similarity assessments, namely analytical similarity, PK/PD similarity, and clinical similarity assessments. In this case, the regulatory agency may have hard time to grant approval of the proposed biosimilar product. A typical example is that notable differences in some CQAs between the proposed biosimilar product and the innovative biological product may be observed in analytical similarity assessment. In this case, the sponsors often provide scientific rationales/justifications to indicate that the notable differences have little or no impact on clinical outcomes. This is probably the most debatable issue between FDA and the Advisory Committee during the review/approval process of the proposed biosimilar product because it is not clearly stated in the FDA guidance whether a proposed biosimilar product is required to passes all similarity tests regardless they are Tier 1 CQAs or Tier 2/Tier 3 CQAs before the regulatory agency can grant approval of the proposed biosimilar product. In this case, if FDA and the Oncologic Drug Advisory Committee (ODAC) panel accept sponsors' scientific rationales and justifications that the notable differences have little or no impact on the clinical outcomes, the proposed biosimilar is likely to be granted for approval. This, however, has raised an interesting question whether the proposed biosimilar product is required to pass all similarity tests (i.e., analytical similarity, PK/PD similarity, and clinical similarity) for regulatory approval.

Examples

For illustration purpose, consider two FDA recent biosimilar regulatory submissions, i.e., Avastin biosimilar (ABP215 sponsored by Amgen) and Herceptin biosimilar (MYL-1401O sponsored by Mylan). These two regulatory submissions were reviewed and discussed at an ODAC meeting held on July 13th, 2017 in Silver Spring, Maryland. Table 2 briefly summarizes the results of the review based on the concept of totality-of-the-evidence (Table 2).

For ABP215, a proposed biosimilar to Genentech's Avastin, although ABP215 passed both PK/PD similarity and clinical similarity tests, several quality attribute differences were noted. These notable differences include glycosylation content, FcγRIIIa binding and product related species (aggregates, fragments, and charge variants). The glycosylation and FcγRIIIa binding differences were addressed by means of *in vitro* cell based on antibody dependent cell-mediated cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC) activity, which were not detected for all products (ABP215, US-licensed Avastin, and EU-approved Avastin).The ODAC panel considered the submission has provided totality-of-the-evidence for demonstration of highly similarity between ABP215 and the US-licensed Avastin, notwithstanding minor differences in clinically inactive components, and support that there are no

Table 1. Assessment of Totality-of-the-Evidence.

No. of dis-similarities	Analytical similarity assessment	PK/PD similarity assessment	Clinical similarity	Overall assessment
0	Yes	Yes	Yes	Yes
1	Yes	Yes	No	No
1	Yes	No	Yes	*
1	No	Yes	Yes	*
2	Yes	No	No	No
2	No	Yes	No	No
2	No	No	Yes	No
3	No	No	No	No

*Scientific rationale are necessary provided

Table 2. Examples of Assessment of Totality-of-the-evidence.

Regulatory Submission	Innovative Product	Proposed Biosimilar	Totality-of-the-evidence		
			Analytical Similarity	PK/PD Similarity	Clinical Similarity
BLA 761028 (Amgen)	Avastin	ABP215	Notable differences observed in glycosylation content and Fcγ Rllla binding	Pass	Pass
BLA 761074 (Mylan)	Herceptin	MYL-1401O	Subtle shifts in glycosylation (sialic acid, high mannose, and NG-HC	Pass	Pass

clinically meaningful differences between ABP215 and the US-licensed Avastin in terms of the safety, purity and potency of the product. For MYL-1401O, a proposed biosimilar to Genentech's Herceptin, although MYL-1401O passed both PK/PD similarity and clinical similarity tests, there are subtle shifts in glycosylation (sialic acid, high mannose, and NG-HC). However, the residual uncertainties related to increase in total mannose forms and sialic acid and decrease in NG-HC were addressed by ADCC similarity and by the PK similarity. Thus, the ODAC panel determined that the submission has provided totality-of-the-evidence to support a demonstration of highly similarity between MYL-1401O and the US-licensed Herceptin, notwithstanding minor differences in clinically inactive components, and support that there are no clinically meaningful differences between MYL-1401O and the US-licensed Herceptin in terms of the safety, purity and potency of the product.

Concluding Remarks

For regulatory approval of new drugs, Section 314.126 of 21 CFR states that substantial evidence needs to be provided to support the claims of new drugs. For regulatory approval of a proposed biosimilar product, the FDA requires totality-of-the-evidence be provided to support a demonstration of biosimilarity between the proposed biosimilar product and the US-licensed drug product. In practice, it should be noted that there is no clear distinction between the substantial evidence in new drug development and the totality-of-the-evidence in biosimilar drug product development. As discussed in the previous section regarding the two recent regulatory submissions, it is not clear whether totality-of-the-evidence of highly similarity can only be achieved if the proposed biosimilar product has passed all similarity tests across different domains of analytical, PK/PD, and clinical assessment. When notable differences in CQAs (in Tier 1) are observed, the notable differences may be ignored if the sponsors can provide scientific rationales/justification to rule out that the observed difference have an impact on clinical outcomes. This, however, is somewhat controversial because Tier 1 CQAs are considered most relevant to clinical outcomes depending upon their criticalities or risk rankings that impact the clinical outcomes. The criticalities and/or risk rankings may be determined using model (3). If a notable difference is considered having little or no impact on the clinical outcome, then the CQA should not be classified into Tier 1 at the first place. This controversy could be due to classification of CQAs based on subjective judgment rather than objectively statistical modeling. In the two examples concerning biosimilar regulatory submissions of ABP215 (Avastin biosimilar) and MYL-1401O (Herceptin biosimilar), the sponsors also seek for approval across different indications. It has been tremendous discussions regarding whether totality-of-the-evidence observed from one indication or a couple of indications can be used to extrapolate to other indications even different indications have similar mechanism of actions. The ODAC panel expressed their concern of extrapolation without collecting any clinical data from other indications and encouraged further research on scientific validity of extrapolation and/or generalize ability of the proposed biosimilar product be conducted.

References

[1]. Chow SC. Biosimilars: Design and Analysis of Follow-on Biologics. Chapman and Hall/CRC Press, Taylor & Francis, New York; 2013.

[2]. Endrenyi L, Declerck P, Chow SC. Biosimilar Drug Product Development. CRC Press, Taylor & Francis Group, New York.2017.

[3]. FDA. Guidance for Industry – Scientific Considerations in Demonstrating Biosimilarity to a Reference Product. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), the United States Food and Drug Administration, Silver Spring, Maryland; 2015.

[4]. US Food and Drug Administration. Guidance for Industry – Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), Silver Spring, US Food and Drug Administration. Maryland; 2016.

[5]. US Food and Drug Administration. Guidance for Industry – Statistical Approaches to Evaluate Analytical Similarity. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), Silver Spring, US Food and Drug Administration. Maryland; 2017.

[6]. Chow SC, Endrenyi L, Lachenbruch PA, Yang LY, Chi E. Scientific factors for assessing biosimilarity and drug Interchangeability of follow-on Biologics. Biosimilars. 2011; 1:13-26.