Analytical Similarity Assessment

Shein-Chung Chow, PhD

Associate Director for Biosimilar Review
Office of Biostatistics, OTS/CDER
US Food and Drug Administration

May 4th, 2018

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

Outline

- Regulatory approval pathway in US
- FDA recommended stepwise approach
- Current issues
  - Same standard for generics/biosimilars and new drugs (90% CI versus 95% CI)
  - Scientific justification for selection of $1.5\sigma_R$
  - Issue on Tier 2 quality range approach
  - Tolerance interval approach
  - Sample size

Development of regulatory approval pathway in US

- *Biologics Price Competition and Innovation* (*BPCI*) Act
  - Passed by US Congress in 2009 and enacted on March 23, 2010
- Regulatory requirements/guidance
  - *FDA Public Hearing* (November 2-3, 2010)
  - Various *User Fees Stakeholders' meetings* within the FDA between November 2-3, 2010 and December 16, 2011
  - *FDA Public Meeting* (December 16, 2011)
  - *Three FDA draft guidance's* (February 9, 2012; 2015)
  - *FDA Public Hearing* (May 11, 2012)
  - *FDA draft guidance on interchangeability* (January, 2017)
  - *FDA draft guidance on analytical similarity assessment* (Sept, 2017)

Scientific Considerations in Demonstrating Biosimilarity to a Reference Product

Guidance for Industry

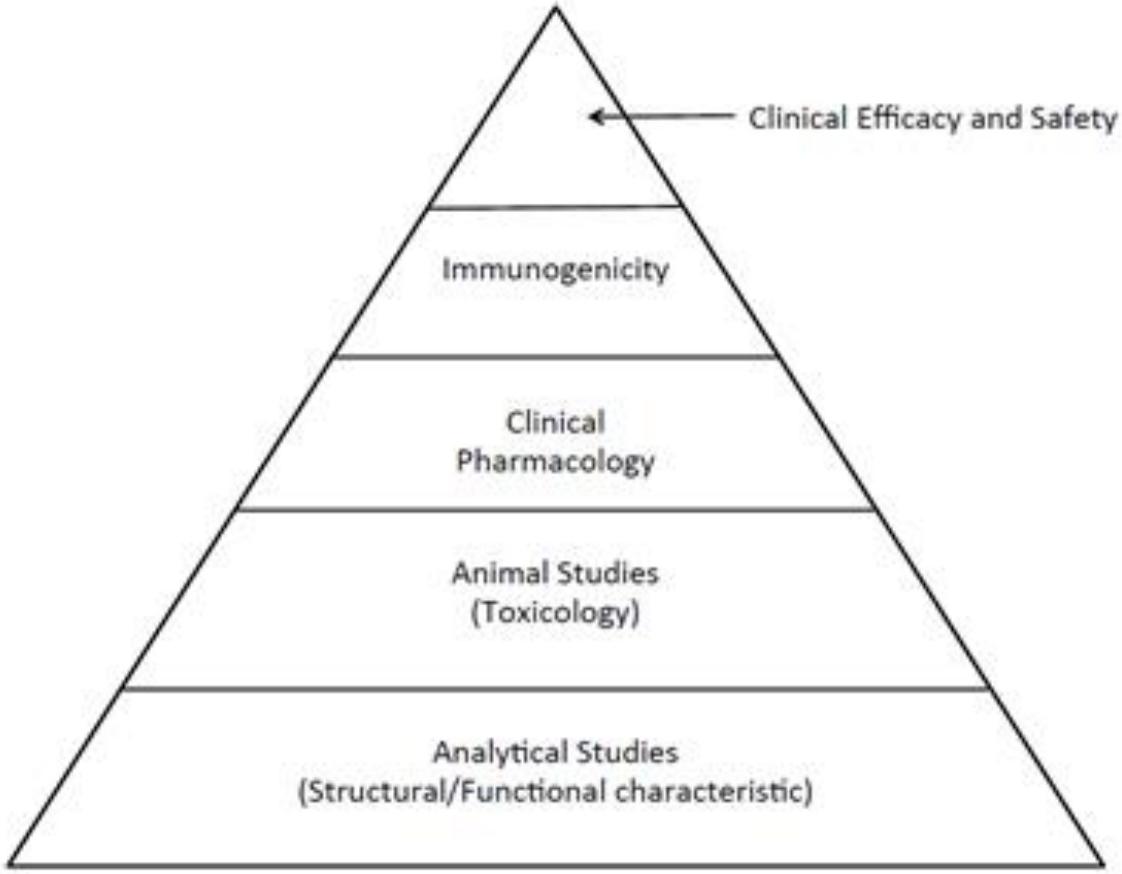
Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

and/or

Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-7800
Email: ocod@fda.hhs.gov
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

Stepwise approach for obtaining totality-of-the-evidence



Stepwise approach

- Analytical studies
  - Critical quality attributes at various stages of manufacturing process
- Animal studies
  - Include the assessment of toxicity
- Clinical pharmacology
  - Include pharmacokinetics (PK) or pharmacodynamics (PD)
- Additional clinical studies
  - The assessment of immunogenicity
  - Safety/tolerability
  - Efficacy

FDA 2015 Guidance

Statistical Approaches to Evaluate Analytical Similarity
Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Scott N. Goldie at 301-796-2055 or (CBER) Office of Communication, Outreach and Development at 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

September 2017
Biosimilars

Analytical similarity assessment

- FDA recommended 3-tier approach
- Classification of critical quality attributes (CQAs) into three tiers according to their criticality or risk ranking relevant to clinical outcomes
- An appropriate statistical model or scoring system based on
  - mechanism of action (MOA) or
  - pharmacokinetic/pharmacodynamics (PK/PD)
  - Information available in the literature

  should be used whenever possible

Analytical similarity assessment

- Tier 1 CQAs
  - Most relevant to clinical outcomes
  - Equivalence test
- Tier 2 CQAs
  - Mild-to-moderate relevant to clinical outcomes
  - Quality range approach
- Tier 3 CQAs
  - Least relevant to clinical outcomes
  - Raw data and graphical comparison

Same standard

90% CI for generics/biosimilars versus 95% CI for new drugs

Generic/biosimilar drugs

- Interval hypotheses testing
- Two one-sided tests (TOST) at α level of significance

  TOST is a size-α test
- Operationally equivalent to $(1\text{-}2\alpha)$ CI approach
- α is well-controlled
- FDA's recommendation is TOST **not** $(1\text{-}2\alpha)$ CI approach

New drugs

- Point hypotheses testing
- Two-sided test (TST) at α level of significance

  Each side is $\alpha/2$
- TST is equivalent to $(1\text{-}\alpha)$ CI approach
- α is well-controlled

ELSEVIER

Controlled Clinical Trials 23 (2002) 515–520

Controlled Clinical Trials

A note on statistical methods for assessing therapeutic equivalence

Shein-Chung Chow, Ph.D.[a,*], Jun Shao, Ph.D.[b]

[a]*Statplus, Inc., Heston Hall, Yardley, Pennsylvania, USA*

[b]*Department of Statistics, University of Wisconsin, Madison, Wisconsin, USA*

Manuscript received October 29, 2001; manuscript accepted April 29, 2002

Abstract

The two one-sided tests procedure and the confidence interval approach are two commonly used statistical approaches for testing therapeutic equivalence or assessing bioequivalence. However, some confusion arises. For example, what is the difference between the two approaches, given the fact that in some cases the two approaches produce the same test? Should we use level $1-\alpha$ or $1-2\alpha$ when applying the confidence interval approach? When different confidence intervals are available, which confidence interval should be used? The purpose of this paper is to clarify this confusion. It is shown that the approach of using $1-\alpha$ confidence intervals produces level α tests, but the sizes of these tests may be smaller than α, and that the use of $1-2\alpha$ confidence intervals generally does not ensure that the corresponding test be of level α, although there are exceptional cases. The sizes of several tests obtained using different confidence intervals are also evaluated.

Keywords: Two one-sided tests; Confidence intervals; Significance level; Size

Scientific justification of $1.5\sigma_R$

- Under similar assumptions that $\mu_T - \mu_R$ is proportional to σ_R and $\mu_T - \mu_R = \frac{1}{8}\sigma_R$, Chow et al. (2016) provided a scientific justification by arriving the same margin of $1.5\sigma_R$.

- Chow SC, Song FY, and Bai H. (2016). *AAPS Journal*, 18(3), 670-677, doi:10.1208/s12248-016-9882-5.

The AAPS Journal, Vol. 18, No. 3, May 2016 (© 2016)
DOI: 10.1208/s12248-016-9882-5

Research Article

Analytical Similarity Assessment in Biosimilar Studies

Shein-Chung Chow,[1,4] Fuyu Song,[2] and He Bai[3]

Received 22 December 2015; accepted 27 January 2016; published online 12 February 2016

Abstract. For assessment of biosimilarity, the US Food and Drug Administration (FDA) recommends a stepwise approach for obtaining the totality-of-the-evidence for demonstrating biosimilarity between a proposed biosimilar product and an innovative (reference) biological product. The stepwise approach starts with analytical studies for functional and structural characterization at various stages of manufacturing process of the proposed biosimilar product. Analytical similarity assessment involves identification of critical quality attributes (CQAs) that are relevant to clinical outcomes. FDA proposes first classifying the identified CQAs into three tiers according to their criticality or risking ranking relevant to clinical outcomes and then performing equivalence test (for CQAs in Tier 1), quality range approach (for CQAs in Tier 2), and raw data or graphical presentation (for CQAs in Tier 3) for obtaining totality-of-the-evidence for demonstrating biosimilarity between the proposed biosimilar product with the reference product. In practice, some debatable issues are evitably raised due to this complicated process of analytical similarity assessment. In this article, these debatable are described and discussed.

KEY WORDS: equivalence test; fixed SD approach; quality range approach; stepwise approach; tiered approach; totality-of-the-evidence.

INTRODUCTION

In recent years, the assessment of biosimilarity for biosimilar products has received much attention by scientists, researchers, and reviewers from the pharmaceutical

ensure final product outputs remain within acceptable quality limits (see, e.g., (3)). For the analytical studies, FDA suggests that CQAs should be identified and classified into three tiers according to their criticality or risk ranking based on mechanism of action (MOA) or PK using

Scientifically meaningful difference

- The mean difference of $\mu_T - \mu_R = \frac{1}{8}\sigma_R$ is the maximum difference allowed for claiming highly similar.
- Sample size calculation was performed for detecting such a difference for achieving an 80% power.

Tier 2 quality range approach

- Quality range set by the reference product
- Derived based on population rather than population mean
  - We expect there are 95% (99%) of test values will fall within 2 (3) SD below and above mean
- Issue regarding Tier 2 quality range approach
  - Especially when there is a significant mean difference between test and reference product

Tier 2 quality range approach

- The quality range approach is a useful method for assessing similarity of non-Tier 1 quality attributes especially when both reference and test products have similar means and variances.
- In cases where there are notable differences in means and/or heterogeneity in variances, the quality range method provides non-statistician such as biologist a visual examination of the seriousness of the mean difference and/or the degree of heterogeneity in variability associated with the proposed biosimilar product for assessment of similarity.

Tolerance interval approach

- Questions regarding tolerance interval and min-max approaches
- Both methods attempt to take variability into consideration
- EMA considers tolerance interval approach is acceptable

Sample size

- Sample size is a function of $\alpha, \beta, \delta, \Delta, \sigma,$ and k, i.e.,

  $n = f(\alpha, \beta, \delta, \Delta, \sigma, \text{and } k)$

  where α = significance level

  β = type II error

  δ = $\mu_T - \mu_R$

  Δ = margin

  σ = standard deviation of reference product

  k = treatment allocation factor

Sample size

- A typical approach is to fix $\alpha,\ \delta, \Delta, \sigma,$ and k, and then select n for achieving a desired power $1-\beta$
- In practice, it is not possible to select n and at the same time controlling all of the parameter
- Sample size may be obtained by
  - Maintaining clinically meaningful difference or margin (δ)
  - Controlling variability (σ)
  - Achieving a desired power (β)
  - Having a high probability of reproducibility (δ and σ)