Communications for Statistical Applications and Methods
2015, Vol. 22, No. 6, 605–613

DOI: http://dx.doi.org/10.5351/CSAM.2015.22.6.605
Print ISSN 2287-7843 / Online ISSN 2383-4757

# Statistical Assessment of Biosimilarity based on the Relative Distance between Follow-on Biologics in the $(k + 1)$-Arm Parallel Design

Seung-Ho Kang[1,a], Wooyoung Shin[a]

[a]Department of Applied Statistics, Yonsei University, Korea

**Abstract**

A three-arm parallel design has been proposed to assess the biosimilarity between a biological product and a reference product using relative distance (Kang and Chow, 2013). The three-arm parallel design consists of two arms for the reference product and one arm for the biosimilar product. This paper extended the three-arm parallel design to a $(k + 1)$-arm parallel design composed of $k\,(\geq 3)$ arms for the reference product and one arm for the biosimilar product. A new relative distance was defined based on Euclidean distance; consequently, a corresponding test procedure was developed based on asymptotic distribution. Type I error rates and powers were investigated both theoretically and empirically.

Keywords: similarity, biosimilar, equivalence trial, delta method

## 1. Introduction

Biological drug products are manufactured by a living system (or organism) and can be used for the treatment, prevention, and cure of human diseases. These are important drug products for patients with unmet medical needs; however, some biological drug products can be very expensive. This high cost makes it difficult for many patients to access these products. A number of biological products are due to expire in the next few years and the subsequent production of follow-on products has attracted significant attention from the pharmaceutical industry and regulatory agencies. The European Medicines Agency (EMA) has taken the initiative to develop regulatory guidelines for the approval pathway of biosimilar products (for example; EMA, 2003, 2005), while the Food and Drug Administration (FDA) has published a draft guidance for the demonstration of biosimilarity to a reference product (US FDA, 2012). These guidelines define a biosimilar product as one that is highly similar to the reference product, notwithstanding minor differences in clinically inactive components, with no clinically meaningful differences in terms of safety, purity, and potency. The guidelines do not contain sufficient discussions of the criteria for biosimilarity; consequently, several articles have been devoted to the search for biosimilarity criteria and corresponding measures for the assessment of the biosimilarity of biosimilar products (Chow and Liu, 2009; Chow *et al.*, 2009; Chow, 2013; Chow *et al.*, 2013; Chow *et al.*, 2014; Dong and Tsong, 2014; Dong *et al.*, 2014; Hsieh *et al.*, 2013; Kang and

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2013R1A 1A200).

[1] Corresponding author: Department of Applied Statistics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. E-mail: seungho@yonsei.ac.kr

Published 30 November 2015 / journal homepage: http://csam.or.kr

Chow, 2013; Kang and Kim, 2014; Liu *et al.*, 2013; Lu *et al.*, 2014; Shin and Kang, 2014; Yang and Lai, 2014).

The manufacturing process for reference biological products may undergo important periodic life cycle changes and small manufacturing changes can affect clinical efficacy and the safety of biologics. Hence, it is important to conduct studies to compare the reference product to itself. In these studies, the reference products can come from different batches or different manufacturing processes (or locations) in order to find the variability associated with the reference product. In order to incorporate this variability, Kang and Chow (2013) proposed a new three-arm parallel design and statistical methods to assess the biosimilarity between a biological product and a reference product. Lu *et al.* (2014) studied the power comparisons between the frequency estimator method, the ratio method, and the linearization method under the three-arm parallel design. Shin and Kang (2015) extended the three-arm parallel design to a binary endpoint.

One of the most commonly used designs for the assessment of biosimilarity is a two-arm equivalence trial with an alternative hypothesis that states that the difference between the two population means of a biosimilar product and a reference product is smaller than a pre-specified margin. A disadvantage of this method is that the variability of the reference product from different batches or different manufacturing processes cannot be incorporated.

The three-arm parallel design consists of two arms for the reference product and one arm for the biological product. This paper extended the three-arm parallel design to a $(k + 1)$-arm parallel design composed of $k\,(\geq 3)$ arms for the reference product and one arm for the biological product, to be applied to cases in which $k\,(\geq 3)$ different batches or different manufacturing processes (or locations) are available. A new relative distance was defined based on Euclidean distance. An asymptotic distribution of the point estimator for the new relative distance was derived using the central limit theorem and delta method. A new test procedure to test biosimilarity was developed based on asymptotic distribution. Type I error rates and powers were investigated theoretically and empirically.

## 2. Statistical Assessment of Biosimilarity

### 2.1. The $(k + 1)$-arm parallel design

Let $T$ denote a biosimilar product and $R_i$ represent the reference biological product from $k$ different batches, respectively, $i = 1, 2, \ldots, k$. It is assumed that the $N$ patients are randomized into the following $k + 1$ groups. The patients allocated to the first group receive the biosimilar product $T$, and the number of patients is denoted with $n_1$. The patients assigned to the $(i+1)^{th}$ group receive the reference product, $R_i$, respectively, $i = 1, 2, \ldots, k$, and the number of patients in each group is denoted with $n_2$ for simplicity. The randomization ratio $k : 1 : 1 : \cdots : 1$ is used, so that we have $n_1 = kn_2$, and the total sample size is $N = n_1 + kn_2$. Let $Y$ denote a continuous primary endpoint.

When developing a relative distance to assess the biosimilarity between one biosimilar product and $k$ reference products, the distances should satisfy the following conditions.

1. The distances should make sense intuitively.

2. The mathematical form of the relative distance should be simple enough to be used to develop a good statistical hypothesis testing procedure.

Considering these two factors, we define the distance $d(T, R)$ between a biosimilar product and $k$ reference products as the Euclidean distance between the two averages of each population mean(s) as

follows.

$$d(T,R) = |\mu_T - \mu_R|, \quad \mu_R = \frac{1}{k}\sum_{i=1}^{k}\mu_{R_i},$$

where $\mu_T$ is the population mean of $Y$ in patients with the biosimilar product $T$ and $\mu_{R_i}$ is the population mean of $Y$ in patients who receive the reference product $R_i$, respectively, $i = 1, 2, \ldots, k$. We define the distance $d(R_1, R_2, \ldots, R_k)$ among $k$ same reference products from different batches as the corresponding Euclidean distance as follows.

$$d(R_1, R_2, \ldots, R_k) = \sqrt{\sum_{i=1}^{k}(\mu_{R_i} - \mu_R)^2}.$$

We then consider the following relative distance in order to assess the biosimilarity between the biosimilar product and the reference product

$$rd = \frac{d(T,R)}{d\,(R_1, R_2, \ldots, R_k)}. \tag{2.1}$$

When $k = 2$, the relative distance above is the same as that in Kang and Chow (2013) except at $1/\sqrt{2}$ times. Since the equivalence margin can be redefined to include $1/\sqrt{2}$, the relative distance in (2.1) can be considered as a natural extension of the relative distance in Kang and Chow (2013).

If the relative distance $rd$ is less than a prespecified margin $\delta(\delta > 0)$ in the $(k+1)$-arm parallel design, we claim that the two products are biosimilar. Therefore, the hypotheses of interest are given by

$$H_0 : rd \geq \delta \quad \text{versus} \quad H_A : rd < \delta. \tag{2.2}$$

Then the hypotheses in (2.2) can be rewritten as

$$H_0 : \theta \leq -\delta \ \text{ or } \ \theta \geq \delta \quad \text{versus} \quad H_A : -\delta < \theta < \delta, \tag{2.3}$$

where

$$\theta = \frac{\mu_T - \frac{1}{k}\sum_{i=1}^{k}\mu_{R_i}}{\sqrt{\sum_{i=1}^{k}(\mu_{R_i} - \mu_R)^2}}.$$

It is well known that the hypotheses in (2.3) can be decomposed into two one-sided hypotheses as follows.

$$H_{01} : \theta \leq -\delta \quad \text{versus} \quad H_{A1} : -\delta < \theta \tag{2.4}$$

and

$$H_{02} : \theta \geq \delta \quad \text{versus} \quad H_{A2} : \theta < \delta. \tag{2.5}$$

## 2.2. Statistical tests

Let $Y_{T,j}$ $(j = 1, 2, \ldots, n_1)$ and $Y_{R_i,j}$ $(i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, n_2)$ denote the continuous primary endpoints from the biosimilar product in the first group and the reference product in the $(i+1)^{th}$ group, respectively. We assume that $Y_{T,j}$'s is a random sample from a distribution whose mean and variance are $\mu_T$ and $\sigma^2$ $(< \infty)$, respectively. Similarly, $Y_{R_i,j}$'s is assumed to be a random sample from a distribution whose mean and variance are $\mu_{R_i}$ and $\sigma^2$ $(< \infty)$, respectively. It is further assumed that $Y_{T,j}$'s and $Y_{R_i,j}$'s are mutually independent and that the sample size is adequate for the central limit theorem to be used.

A natural way to estimate $\theta$ in (2.3) is to replace the unknown population means with the corresponding sample means. So an estimator of $\theta$ is given by

$$\hat{\theta} = \frac{\bar{Y}_T - \bar{Y}_R}{\sqrt{\sum_{i=1}^{k} \left(\bar{Y}_{R_i} - \bar{Y}_R\right)^2}},$$

where

$$\bar{Y}_T = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_{T_i}, \qquad \bar{Y}_{R_i} = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_{R_i,j}, \qquad \bar{Y}_R = \frac{1}{k}\sum_{i=1}^{k} \bar{Y}_{R_i}.$$

In the Appendix, we derive the asymptotic distribution of $\sqrt{n_1}(\hat{\theta} - \theta)$ using the central limit theorem and delta method as follows.

$$\sqrt{n_1}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(0, \sigma_\theta^2\right),$$

where the specific form of $\sigma_\theta^2$ is provided in the Appendix.

By using this asymptotic normality of $\sqrt{n_1}(\hat{\theta} - \theta)$, we can conduct hypothesis testing and establish an asymptotic confidence interval for $\theta$. The null hypothesis $H_{01}$ in (2.4) is rejected if $Z_1 > z_\alpha$ where

$$Z_1 = \frac{\hat{\theta} + \delta}{\hat{\sigma}_\theta / \sqrt{n_1}},$$

where $\hat{\sigma}_\theta$ is an estimator of $\sigma_\theta$ whose specific form is provided in the Appendix and $z_\alpha$ is the upper $\alpha$ quantile of the standard normal distribution (for example, $z_{0.05} = 1.645$). In this paper, $\hat{\sigma}_\theta^2$ is obtained by replacing $\mu_T$, $\mu_{R_i}$, $\sigma^2$ in $\sigma_\theta^2$ with $\bar{Y}_T$, $\bar{Y}_{R_i}$, and $s^2$ (pooled sample variance). Similarly, the null hypothesis $H_{02}$ in (2.5) is rejected if $Z_2 < -z_\alpha$ where

$$Z_2 = \frac{\hat{\theta} - \delta}{\hat{\sigma}_\theta / \sqrt{n_1}}.$$

We claim that the two products are biosimilar if each null hypothesis in both (2.4) and (2.5) is rejected at the significance level $\alpha$.

Alternatively, a two-sided asymptotic confidence interval for $\theta$ can also be used to assess the biosimilarity between the two products. Because a $100(1-\alpha)\%$ asymptotic confidence interval for $\theta$ is given by

$$\left(\hat{\theta} \pm z_{\frac{\alpha}{2}} \frac{\hat{\sigma}_\theta}{\sqrt{n_1}}\right),$$

we claim that the two products are biosimilar if

$$-\delta < \hat{\theta} - z_{\frac{\alpha}{2}}\hat{\sigma}_\theta \quad \text{and} \quad \hat{\theta} + z_{\frac{\alpha}{2}}\hat{\sigma}_\theta < \delta.$$

### 2.3. Type I error rate and power

In Section 2.2, we proposed a statistical test to test the hypothesis in (2.3). In this section we investigate asymptotic type I error rates and powers based on the asymptotic distribution and empirical type I error rates and powers based on simulation studies.

First, we derive the formula of the asymptotic type I error rate and power based on the asymptotic distribution. Under the alternative hypothesis $H_A : -\delta < \theta < \delta$, the asymptotic power function is given by

$$\begin{aligned} p(\mu_T, \mu_{R_1}, \ldots, \mu_{R_k}) &= P(Z_2 < -z_\alpha \text{ and } Z_1 > z_\alpha | -\delta < \theta < \delta) \\ &\sim P\left(Z < -z_\alpha + \frac{\delta - \theta}{\sigma_\theta/\sqrt{n_1}} \text{ and } Z > z_\alpha - \frac{\delta + \theta}{\sigma_\theta/\sqrt{n_1}}\right) \\ &= \Phi\left(-z_\alpha + \frac{\delta - \theta}{\sigma_\theta/\sqrt{n_1}}\right) - \Phi\left(z_\alpha - \frac{\delta + \theta}{\sigma_\theta/\sqrt{n_1}}\right), \end{aligned} \tag{2.6}$$

where the random variable $Z$ follows the standard normal distribution, and $\Phi$ is the cumulative distribution function of the standard normal distribution. In order to have positive powers, the following condition should be satisfied.

$$-z_\alpha + \frac{\delta - \theta}{\sigma_\theta/\sqrt{n_1}} > z_\alpha - \frac{\delta + \theta}{\sigma_\theta/\sqrt{n_1}}.$$

Therefore, the margin $\delta$ should satisfy the following constraint.

$$\delta > z_\alpha \frac{\sigma_\theta}{\sqrt{n_1}}.$$

The asymptotic type I error rate is the same as (2.6), except that it is evaluated under the null hypothesis $H_0$ in (2.3). The asymptotic type I error rates and powers are calculated with (2.6) and presented in Tables 1 and 2.

We need to investigate the empirical type I error rates and powers in finite samples since the statistical test developed in Section 3 employs large sample theory. Random samples of $Y_{T,j}$ and $Y_{R_i,j}$ $(i = 1, 2, \ldots, k)$ are generated from normal distributions under the null hypothesis in (2.3). We compute two test statistics $Z_1$ and $Z_2$. If $Z_1 > z_{0.05}$ and $Z_2 < -z_{0.05}$, then the null hypothesis in (2.3) is rejected. We generate 5,000 simulation samples and calculate the empirical type I error rates as the proportion of samples in which the null hypothesis in (2.3) is rejected. Similarly, the empirical powers are also computed.

Table 1 displays asymptotic and empirical type I error rates at $n_1 = 100$ and $n_1 = 200$, respectively. It shows that the asymptotic type I error rates are close to the nominal level (5%) except in the cases of $\sigma^2 = 2$ and $n_1 = 100$, and that the empirical type I error rates are much smaller than the nominal level (5%). Table 2 presents the asymptotic and empirical powers at $n_1 = 100$ and $n_1 = 200$, respectively.

## 3. Concluding Remarks

In this article, we focused on biosimilarity testing for a biosimilar product and a reference product. We proposed a biosimilarity testing procedure based on the relative distance under a $(k + 1)$-arm parallel design. We investigated the performance of the proposed method theoretically and empirically. Shin and Kang (2014) extended the three-arm parallel design to binary endpoints for three popular metrics

Table 1: The asymptotic and empirical type I error rates (%)

| $\mu_T$ | $\mu_{R_1}$ | $\mu_{R_2}$ | $\mu_{R_3}$ | $\sigma^2$ | $\theta$ | $\delta$ | $n = 100$ | | $n = 200$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Asym | Empi | Asym | Empi |
| 13 | 15 | 14 | 9 | 2 | 0.07332 | 0.07332 | 0.3 | 0.4 | 4.9 | 2.3 |
| 13 | 15 | 14 | 9 | 2 | 0.07332 | 0.07330 | 0.2 | 0.4 | 4.9 | 2.3 |
| 13 | 15 | 13 | 10 | 2 | 0.09366 | 0.09366 | 0.2 | 0.6 | 4.9 | 2.2 |
| 13 | 15 | 13 | 10 | 2 | 0.09366 | 0.09360 | 0.2 | 0.4 | 4.9 | 2.0 |
| 13 | 14 | 16 | 8 | 2 | 0.05661 | 0.05661 | 0.4 | 0.3 | 4.9 | 2.3 |
| 13 | 14 | 16 | 8 | 2 | 0.05661 | 0.05660 | 0.3 | 0.2 | 4.9 | 1.8 |
| 13 | 14 | 16 | 10 | 2 | −0.07715 | 0.07715 | 0.3 | 0.3 | 4.9 | 2.5 |
| 13 | 14 | 16 | 10 | 2 | −0.07715 | 0.07700 | 0.2 | 0.3 | 4.9 | 2.3 |
| 13 | 14 | 9 | 17 | 2 | −0.05832 | 0.05832 | 0.3 | 0.4 | 4.9 | 2.1 |
| 13 | 14 | 9 | 17 | 2 | −0.05832 | 0.05800 | 0.2 | 0.2 | 4.9 | 2.0 |
| 12 | 11 | 8 | 18 | 2 | −0.04593 | 0.04593 | 0.4 | 0.3 | 4.9 | 2.5 |
| 12 | 11 | 8 | 18 | 2 | −0.04593 | 0.04550 | 0.1 | 0.1 | 4.9 | 1.8 |
| 12 | 11 | 7 | 17 | 2 | 0.04683 | 0.04683 | 0.4 | 0.4 | 4.9 | 2.2 |
| 12 | 11 | 7 | 17 | 2 | 0.04683 | 0.04650 | 0.1 | 0.0 | 4.9 | 1.8 |
| 12 | 14 | 9 | 14 | 2 | −0.08165 | 0.08165 | 0.3 | 0.5 | 4.9 | 2.2 |
| 12 | 14 | 9 | 14 | 2 | −0.08165 | 0.08130 | 0.1 | 0.4 | 4.9 | 1.8 |
| 12 | 15 | 9 | 13 | 2 | −0.07715 | 0.07715 | 0.3 | 0.4 | 4.9 | 2.2 |
| 12 | 15 | 9 | 13 | 2 | −0.07715 | 0.07700 | 0.2 | 0.3 | 4.9 | 1.8 |
| 10 | 12 | 7 | 10 | 2 | 0.09366 | 0.09366 | 0.2 | 0.6 | 4.9 | 2.1 |
| 10 | 12 | 7 | 10 | 2 | 0.09366 | 0.09340 | 0.1 | 0.4 | 4.9 | 2.0 |
| 13 | 15 | 14 | 9 | 1 | 0.07332 | 0.07332 | 4.9 | 2.3 | 5.0 | 2.3 |
| 13 | 15 | 14 | 9 | 1 | 0.07332 | 0.07000 | 4.8 | 1.9 | 5.0 | 1.5 |
| 13 | 15 | 13 | 10 | 1 | 0.09366 | 0.09366 | 4.9 | 2.6 | 5.0 | 2.2 |
| 13 | 15 | 13 | 10 | 1 | 0.09366 | 0.09000 | 4.8 | 1.8 | 5.0 | 1.3 |
| 13 | 14 | 16 | 8 | 1 | 0.05661 | 0.05661 | 4.9 | 2.2 | 5.0 | 2.3 |
| 13 | 14 | 16 | 8 | 1 | 0.05661 | 0.05000 | 4.7 | 0.9 | 5.0 | 0.6 |
| 13 | 14 | 16 | 10 | 1 | −0.07715 | 0.07715 | 4.9 | 2.6 | 5.0 | 2.2 |
| 13 | 14 | 16 | 10 | 1 | −0.07715 | 0.07000 | 4.5 | 1.0 | 5.0 | 0.7 |
| 13 | 14 | 9 | 17 | 1 | −0.05832 | 0.05832 | 4.9 | 2.1 | 5.0 | 2.5 |
| 13 | 14 | 9 | 17 | 1 | −0.05832 | 0.05000 | 4.1 | 0.7 | 5.0 | 0.4 |
| 12 | 11 | 8 | 18 | 1 | −0.04593 | 0.04593 | 4.9 | 2.4 | 5.0 | 2.3 |
| 12 | 11 | 8 | 18 | 1 | −0.04593 | 0.04000 | 4.3 | 0.8 | 5.0 | 0.6 |
| 12 | 11 | 7 | 17 | 1 | 0.04683 | 0.04683 | 4.9 | 2.0 | 5.0 | 2.3 |
| 12 | 11 | 7 | 17 | 1 | 0.04683 | 0.04000 | 4.1 | 0.8 | 5.0 | 0.5 |
| 12 | 14 | 9 | 14 | 1 | −0.08165 | 0.08165 | 4.9 | 2.2 | 5.0 | 2.6 |
| 12 | 14 | 9 | 14 | 1 | −0.08165 | 0.07000 | 4.1 | 0.6 | 5.0 | 0.5 |
| 12 | 15 | 9 | 13 | 1 | −0.07715 | 0.07715 | 4.9 | 2.5 | 5.0 | 1.9 |
| 12 | 15 | 9 | 13 | 1 | −0.07715 | 0.07000 | 4.5 | 1.1 | 5.0 | 0.9 |
| 10 | 12 | 7 | 10 | 1 | 0.09366 | 0.09366 | 4.9 | 2.2 | 5.0 | 2.4 |
| 10 | 12 | 7 | 10 | 1 | 0.09366 | 0.08000 | 4.1 | 0.6 | 5.0 | 0.3 |

: the risk difference, the log relative risk, and the log odds ratio. It is useful to extend the $(k + 1)$-arm parallel design to binary endpoints. The determination of the biosimilar margin $\delta$ is crucial to the evaluation of biosimilarity. The margin should be pre-specified and justified on statistical and clinical grounds (World Health Organization, 2009; US FDA, 2012); however, the method to determine the margin is not yet known. There have been some studies and relevant guidelines (Li *et al.*, 2013; World Health Organization, 2009; US FDA, 2010, 2012); however, determining the margin for the evaluation of biosimilarity remain an urgent topic for future research.

## Appendix: Central Limit Theorem and Delta Method

By the central limit theorem,

Table 2: The asymptotic and empirical powers (%)

| $\mu_T$ | $\mu_{R_1}$ | $\mu_{R_2}$ | $\mu_{R_3}$ | $\sigma^2$ | $\theta$ | $\delta$ | $n = 100$ | | $n = 200$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Asym | Empi | Asym | Empi |
| 13 | 15 | 14 | 9 | 2 | 0.073 | 0.17 | 70.7 | 74.0 | 92.7 | 96.4 |
| 13 | 15 | 14 | 9 | 2 | 0.073 | 0.18 | 77.9 | 82.2 | 96.2 | 98.5 |
| 13 | 14 | 16 | 8 | 2 | 0.057 | 0.14 | 78.9 | 83.4 | 96.5 | 98.8 |
| 13 | 14 | 16 | 8 | 2 | 0.057 | 0.15 | 86.4 | 90.4 | 98.7 | 99.6 |
| 13 | 14 | 9 | 17 | 2 | −0.058 | 0.14 | 75.3 | 79.9 | 95.0 | 98.0 |
| 13 | 14 | 9 | 17 | 2 | −0.058 | 0.15 | 83.4 | 87.4 | 98.0 | 99.2 |
| 12 | 11 | 7 | 17 | 2 | 0.047 | 0.11 | 72.6 | 76.8 | 93.7 | 96.8 |
| 12 | 11 | 7 | 17 | 2 | 0.047 | 0.12 | 83.0 | 87.9 | 97.9 | 99.4 |
| 12 | 15 | 9 | 13 | 2 | −0.077 | 0.18 | 71.5 | 75.8 | 93.1 | 96.7 |
| 12 | 15 | 9 | 13 | 2 | −0.077 | 0.19 | 78.3 | 81.8 | 96.3 | 98.3 |
| 13 | 15 | 14 | 9 | 1 | 0.073 | 0.15 | 79.1 | 83.3 | 96.6 | 98.6 |
| 13 | 15 | 14 | 9 | 1 | 0.073 | 0.16 | 87.1 | 91.7 | 98.9 | 99.7 |
| 13 | 14 | 16 | 8 | 1 | 0.057 | 0.11 | 71.6 | 74.5 | 93.2 | 96.7 |
| 13 | 14 | 16 | 8 | 1 | 0.057 | 0.12 | 83.8 | 87.9 | 98.1 | 99.3 |
| 13 | 14 | 9 | 17 | 1 | −0.058 | 0.12 | 80.0 | 84.7 | 96.9 | 98.8 |
| 13 | 14 | 9 | 17 | 1 | −0.058 | 0.13 | 89.3 | 93.6 | 99.3 | 99.9 |
| 12 | 11 | 7 | 17 | 1 | 0.047 | 0.09 | 70.0 | 74.2 | 92.3 | 95.9 |
| 12 | 11 | 7 | 17 | 1 | 0.047 | 0.10 | 84.8 | 89.5 | 98.4 | 99.5 |
| 12 | 15 | 9 | 13 | 1 | −0.077 | 0.15 | 71.6 | 76.0 | 93.2 | 96.2 |
| 12 | 15 | 9 | 13 | 1 | −0.077 | 0.16 | 80.9 | 85.8 | 97.2 | 99.0 |

$$\begin{aligned}
&\sqrt{n_1}\left(\bar{Y}_T - \mu_T\right) \xrightarrow{d} N\left(0, \sigma^2\right),\\
&\sqrt{n_2}\left(\bar{Y}_{R_i} - \mu_{R_i}\right) \xrightarrow{d} N\left(0, \sigma^2\right), \quad \text{for } i = 1, \ldots, k,\\
&\sqrt{n_1}\left(\bar{Y}_{R_i} - \mu_{R_i}\right) = \frac{\sqrt{n_1}}{\sqrt{n_2}}\sqrt{n_2}\left(\bar{Y}_{R_i} - \mu_{R_i}\right) \xrightarrow{d} N\left(0, k\sigma^2\right).
\end{aligned}$$

Since $(k + 1)$ groups are independent,

$$\sqrt{n_1}\left[\begin{pmatrix} \bar{Y}_T \\ \bar{Y}_{R_1} \\ \vdots \\ \bar{Y}_{R_k} \end{pmatrix} - \begin{pmatrix} \mu_T \\ \mu_{R_1} \\ \vdots \\ \mu_{R_k} \end{pmatrix}\right] \xrightarrow{d} N_{k+1}\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \Sigma\right],$$

where

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{pmatrix}.$$

We consider the following function to use the multivariate delta method.

$$g(a, b_1, \ldots, b_k) = \frac{a - \frac{1}{k}\sum_{i=1}^{k} b_i}{\sqrt{\sum_{l=1}^{k}\left(b_l - \frac{1}{k}\sum_{i=1}^{k} b_i\right)^2}} \equiv \frac{h}{f},$$

where

$$h \equiv h(a, b_1, \ldots, b_k) = a - \frac{1}{k}\sum_{i=1}^{k} b_i,$$
$$f \equiv f(a, b_1, \ldots, b_k) = \sqrt{\sum_{l=1}^{k}\left(b_l - \frac{1}{k}\sum_{i=1}^{k} b_i\right)^2}.$$

Then, the partial derivatives of $g$ with respect to its components are given by

$$\frac{\partial g}{\partial a} = \frac{1}{f},$$
$$\frac{\partial g}{\partial b_i} = \frac{\left[\frac{\partial h}{\partial b_i}\right] f - h\left[\frac{\partial f}{\partial b_i}\right]}{f^2},$$

where

$$\frac{\partial h}{\partial b_i} = -\frac{1}{k},$$
$$\frac{\partial f}{\partial b_i} = \left(\frac{1}{2}f^{-1}\right)\left[\sum_{l=1}^{k} 2\left(b_l - \frac{1}{k}\sum_{i=1}^{k} b_i\right)\left(-\frac{1}{k}\right) + 2\left(b_j - \frac{1}{k}\sum_{i=1}^{k} b_i\right)\right].$$

Then, by the multivariate delta method,

$$\sqrt{n_1}\left(\hat{\theta} - \theta\right) = \sqrt{n_1}\left(g\left(\bar{Y}_T, \bar{Y}_{R_1}, \ldots, \bar{Y}_{R_k}\right) - g\left(\mu_T, \mu_{R_1}, \ldots, \mu_{R_k}\right)\right) \xrightarrow{d} N\left(0, \sigma_\theta^2\right),$$

where

$$\sigma_\theta^2 = B\Sigma B^T$$

and

$$B = \left[\frac{\partial g}{\partial a}, \frac{\partial g}{\partial b_1}, \ldots, \frac{\partial g}{\partial b_k}\right]\Bigg|_{a=\mu_T, b_i=\mu_{R_i}}, \quad i = 1, \ldots, k.$$

## References

Chow, S. C. (2013). *Biosimilars: Design and Analysis of Follow-on Biologics*, CRC Press, Boca Raton, FL..

Chow, S. C., Endrenyi, L., Lachenbruch, P. A. and Mentre, F. (2014). Scientific factors and current issues in biosimilar studies, *Journal of Biopharmaceutical Statistics*, **24**, 1138–1153.

Chow, S. C., Hsieh, T. C., Chi, E. and Yang, J. (2009). A comparison of moment-based and probability-based criteria for assessment of follow-on biologics, *Journal of Biopharmaceutical Statistics*, **20**, 31–45.

Chow, S. C. and Liu, J. P. (2009). Statistical assessment of biosimilar products, *Journal of Biopharmaceutical Statistics*, **20**, 10–30.

Chow, S. C., Wang, J., Endrenyi, L. and Lachenbruch, P. A. (2013). Scientific considerations for assessing biosimilar products, *Statistics in Medicine*, **32**, 370–381.

Dong, X. and Tsong, Y. (2014). Equivalence assessment for interchangeability based on two-sided tests, *Journal of Biopharmaceutical Statistics*, **24**, 1312–1331.

Dong, X., Tsong, Y. and Shen, M. (2014). Equivalence tests for interchangeability based on two one-sided probabilities, *Journal of Biopharmaceutical Statistics*, **24**, 1332–1348.

European Medicines Agency (2003). Guideline on comparability of medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues (EMEA/CPMP/3797/02), Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003963.pdf

European Medicines Agency (2005). Guideline on similar biological medicinal products (CHMP/437/04), Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003517.pdf

Hsieh, T. C., Chow, S. C., Yang, L. Y. and Chi, E. (2013). The evaluation of biosimilarity index based on reproducibility probability for assessing follow-on biologics, *Statistics in Medicine*, **32**, 406–414.

Kang, S. H. and Chow, S. C. (2013). Statistical assessment of biosimilarity based on relative distance between follow-on biologics, *Statistics in Medicine*, **32**, 382–392.

Kang, S. H. and Kim, Y. (2014). Sample size calculations for the development of biosimilar products, *Journal of Biopharmaceutical Statistics*, **24**, 1215–1224.

Li, Y., Liu, Q., Wood, P. and Johri, A. (2013). Statistical considerations in biosimilar clinical efficacy trials with asymmetrical margins, *Statistics in Medicine*, **32**, 393–405.

Lu, Y., Zhang, Z. Z. and Chow, S. C. (2014). Frequency estimator for assessing of follow-on biologics, *Journal of Biopharmaceutical Statistics*, **24**, 1280–1297.

Shin, W. and Kang, S. H. (2014). Statistical assessment of biosimilarity based on the relative distance between follow-on biologics for binary endpoints, *Journal of Biopharmaceutical Statistics*, Advance online publication. http://dx.doi.org/10.1080/10543406.2014.979195

US Food and Drug Administration (2010). Guidance for Industry: Non-inferiority clinical trials, www.fda.gov/downloads/drugs /guidancecomplianceregulatoryinformation/guidances/ucm202140.pdf

US Food and Drug Administration (2012). Scientific consideration in demonstrating biosimilarity to a reference product: guidance for industry, http://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf

World Health Organization (2009). Guidelines on evaluation of similar biotherapeutic products, http://www.who.int/biologicals/areas/biological_therapeutics/BIOTHERAPEUTICS_FOR_WEB_22APRIL2010.pdf

Yang, L. Y. and Lai, C. H. (2014). Estimation and approximation approaches for biosimilar index based on reproducibility probability, *Journal of Biopharmaceutical Statistics*, **24**, 1298–1311.

*Received July 17, 2015; Revised September 4, 2015; Accepted September 5, 2015*