Enliven Archive
Intensifying Thoughts

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

Equivalent Test with Flexible Margin in Analytical Similarity Assessment

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

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

[*]Corresponding Author: Minkyung Oh, Clinical Development Division, Celltrion, Inc. 19F, IBS Building, 263, Central-ro, Yeonsu-gu, Incheon, Republic of Korea. Tel: +82-32-850-5839, Fax: +82-32-837-1203 E-mail: minkyung.oh@celltrion.com

Received Date: 05th November 2019
Accepted Date: 25th November 2019
Published Date: 02nd December 2019

Citation: Lee SJ, Oh M, Chow SC. Equivalent Test with Flexible Margin in Analytical Similarity Assessment. Enliven: Biosimilars Bioavailab.2019; 3(2): 005-011.

Abstract

For analytical similarity assessment of a given critical quality attribute between a proposed biosimilar (test) product and an innovative (reference) biological product, FDA recommended an equivalence test with a margin of 1.5 σ_R (standard deviation of the reference product) be performed. The FDA recommended similarity margin has been criticized by many authors in the literature due to its inflexibility. In this article, we proposed an equivalence test with flexible margin for controlling possible inflation/deflation of the variability associated with the response. The performance of the proposed equivalence test with flexible margin is evaluated both theoretically and by means of simulation. The results indicate that flexible margin can be selected within the range of (1.575 σ_R, 2.025 σ_R) for achieving reasonable statistical assurance, for example, controlling type I error at the α=5% level of significance and 90% power for analytical similarity assessment.

Keywords: Mean Shift; Inflation of Variability; Sample Size Requirement.

Introduction

For review and approval of a proposed biosimilar product, the United States (US) Food and Drug Administration (FDA) suggested a stepwise approach for obtaining totality-of-the-evidence for demonstrating that the proposed biosimilar product is highly similar to an innovative (reference) biological product [1]. The stepwise approach starts with analytical similarity assessment, followed by the demonstration of pharmacokinetics (PK) and pharmacodynamic (PD) similarity and the assessment of clinical similarity (including efficacy and immunogenicity). For analytical similarity assessment, the FDA recommended equivalence test for critical quality attributes (CQAs) that relevant to clinical outcomes with high criticality or risk ranking, quality range (QR) approach for CQAs that are relevant to clinical outcome with less risk ranking, and graphical presentation for CQAs that are relevant to clinical outcomes with least risk ranking be performed [2]. When performing the equivalence test, FDA recommended an equivalence acceptance criterion (EAC), i.e., a similarity margin of $1.5\sigma_R$ be used for demonstrating that the test product is highly similar to the reference product. This equivalence test is considered the most rigorous test among the three testing procedures.

For a given CQA, FDA suggested the equivalence test be performed based on test results of a single test for each of the randomly selected lots of the test product and reference product. FDA recommended that sample standard deviation of the individual test results of the reference lots be used to estimate σ_R for obtaining EAC for the equivalence test. The FDA recommended procedure for equivalence test, however, has been criticized by many researchers in several ways. First, it is unclear how the coefficient of 1.5 was selected although Chow, Song, and Bai [3] provided a justification based on the concept of the scaled average bioequivalence (SABE) under a multicaptive model. Second, it is a great concern that sample standard deviation is data-dependent which can very depend upon the selected reference lots for the equivalence test. Besides, sample standard deviation is a point estimate and there is variability associated with the point estimate. Thus, the margin (i.e., EAC) selected based on point estimate of σ_R is considered not flexible enough to reflect the true σ_R, which falls within the confidence interval $(\hat{\sigma_L},\hat{\sigma_U})$, where $\hat{\sigma_L}$ and $\hat{\sigma_U}$ are 95% confidence lower and upper limits, respectively. In practice, it is then suggested that a more flexible margin be selected based on an estimate of σ_R selected in the range of $(\hat{\sigma_L},\hat{\sigma_U})$.

The purpose of this article is to (i) explore the range of flexible margin, (ii) propose an optimal but flexible EAC margin for equivalence test, and (iii) perform sample size calculation based on the proposed margin. In the next section, statistical properties of the flexible margin are explored. An optimal but flexible EAC margin is derived in Section 3. Also included in this section is the examination of statistical properties of the

proposed flexible margin. The impact on sample size requirement based on the proposed flexible margin is studied in Section 4. Some concluding remarks are given in the last section of this article.

Range of Flexible EAC Margin

For CQAs that are relevant to clinical outcomes with high criticality or risk ranking, FDA recommended an equivalence test be performed.

Equivalence Test

Let $x_{Ri}, i = 1,\ldots,n_R$ and x_{Ti}, i = 1,…, n_T be the test values of the reference product and test product, respectively. The equivalence test can be summarized in the following steps:

Step 1. Calculate mean and variance for reference product, which are given by

$$\hat{\mu}_R = \frac{1}{n_R}\sum_{i=1}^{n_R} X_{Ri} \text{ and } \hat{\sigma}_R^2 = \frac{1}{n_R - 1}\sum_{i=1}^{n_R}(X_{Ri} - \hat{\mu}_R)^2$$

$$\hat{\mu}_T = \frac{1}{n_T}\sum_{i=1}^{n_T} X_{Ti} \text{ and } \hat{\sigma}_T^2 = \frac{1}{n_R - 1}\sum_{i=1}^{n_T}(X_{Ti} - \hat{\mu}_T)^2$$

Step 2. Establish EAC margin, which is given by $\delta = 1.5 \times \hat{\sigma}_R$

Step 3. Construct 90% confident interval for $\mu_T - \mu_R$, which is given below:

$$\text{CI} = \left(\hat{\mu}_R - \hat{\mu}_T - t_{0.95} * \sqrt{\frac{\hat{\sigma}_R^2}{n_R} + \frac{\hat{\sigma}_T^2}{n_T}}, \hat{\mu}_R - \hat{\mu}_T + t_{0.95} * \sqrt{\frac{\hat{\sigma}_R^2}{n_R} + \frac{\hat{\sigma}_T^2}{n_T}} \right)$$

where $t_{0.95}$ is the 95th percentile of T distribution with df =

$$\frac{\left(\frac{\hat{\sigma}_R^2}{n_R} + \frac{\hat{\sigma}_T^2}{n_T}\right)^2}{\frac{(\hat{\sigma}_R^2/n_R)^2}{n_R - 1} + \frac{(\hat{\sigma}_T^2/n_T)^2}{n_T - 1}}$$

Step 5. Decision-making

If the constructed 90% CI is totally within the EAC margin, then we claim that the mean of test product is equivalent with the mean of reference product for the specific CQA.

Range of EAC Margin

As indicated in the previous section, per FDA's recommendation, EAC margin is determined based on a point estimate of σ_R, which is the sample standard deviation, $\hat{\sigma}_R$. This margin, however, has been criticized due to its inflexibility [4]. In practice, an estimate of σ_R can be selected any values from a 95% confidence interval of σ_R. Let $(\hat{\sigma}_L, \hat{\sigma}_U)$ be the 95% confidence interval of σ_R, where $\hat{\sigma}_U$ is given by

$$\hat{\sigma}_U = \sqrt{\frac{n_R - 1}{\chi^2_{\alpha/2, n_R - 1}}}\,\hat{\sigma}_R$$

Also, let $\hat{\sigma}_R^*$ be a flexibly selected value from the range of $(\hat{\sigma}_R, \hat{\sigma}_U)$ and define a flexible index f as follows

$$\hat{\sigma}_R = \quad \times \hat{\sigma}_R$$

where $1 < f < \hat{\sigma}_U/\hat{\sigma}_R$. Thus, the flexible margin becomes

$$\delta = 1.5 \times \hat{\sigma}_R^* = 1.5 f \times \hat{\sigma}_R \quad \text{----- (1)}$$

In practice, if $\frac{2}{3} < \frac{\sigma_T}{\sigma_R} < \frac{3}{2}, \sigma_T$ considered highly similar to σ_R [5]. For a selection of f, one may consider choosing the value halfway between 1 and 3/2 = 1.5. That is, f is chosen to be 1.25. This is to ensure that $\hat{\sigma}_T/\hat{\sigma}_R$ is less than 1.5 at the worst possible scenario. In this case, EAC margin becomes

$$\delta_c = \delta \times f = 1.5 \times \hat{\sigma}_R \times 1.25 = 1.875\,\hat{\sigma}_R. \quad \text{----- (2)}$$

We will refer to this EAC margin as a conservative flexible margin. On the other hand, since $1 < f < \frac{\hat{\sigma}_U}{\hat{\sigma}_R} < \frac{3}{2}$ for highly similarity, we can select $f = \frac{\hat{\sigma}_U}{\hat{\sigma}_R}$ as the most aggressive flexible margin, denoted by δ_{ma}. Without loss of generality, set $\hat{\sigma}_U/\hat{\sigma}_R = 1.5$. This leads to

$$\delta_{ma} = \delta \times f = 1.5 \times \hat{\sigma}_R \times 1.5 = 2.25\,\hat{\sigma}_R. \quad \text{----- (3)}$$

Thus, based on the above discussion, it is suggested that a flexible EAC margin be selected from the flexible range of $(1.875\hat{\sigma}_R, 2.25\hat{\sigma}_R)$. In practice, as an example, one may select the average of the conservative margin and the most aggressive margin as the flexible margin as follows

$$\delta_{ave} = \frac{1}{2}\left(\delta_c + \delta_{ma}\right) = 2.06\,\hat{\sigma}_R \approx 2.0\,\hat{\sigma}_R \quad \text{----- (4)}$$

This method is consistent with FDA's current thinking that flexible analytical assessment is needed in small samples [6]. To achieve 90% of power for evaluating the similarity between the test and reference lot with allowable mean shift, $\mu_T - \mu_R < \frac{1}{8}\sigma_R$ [2], at least 12 samples are required when using EAC margin 1.5 σ_R, while only 7 samples are required for EAC margin $2.0\sigma_R$.

Optimal and Flexible Margin

As discussed in the previous section, the selection of (4) as flexible EAC margin, it is of interest to investigate whether such a selection is the optimal choice in terms certain desirable statistical properties. For this purpose, for a fixed sample size n_R, we may select f for achieving maximum power for testing the following interval hypotheses for equivalence or similarity.

$H0 : |\varepsilon| \geq \delta$ versus H_a: $|\varepsilon| < \delta$, ----- (5)

where $\varepsilon = \mu_T - \mu_R$ and $\delta = 1.5 f \hat{\sigma}_R$

The test drug is concluded to be equivalent to the reference drug on average if the null hypothesis is rejected at significance level α. Assuming that $n_T = n_R = n$ and $\sigma_T \approx \sigma_R$, when $\sigma_R = \hat{\sigma}_R$ is known, the null hypothesis H_0 is rejected at significance level α if

$$\frac{\left(\varepsilon - 1.5 f \hat{\sigma}_R\right)}{\hat{\sigma}_R \sqrt{\frac{2}{n}}} < -z_\alpha \text{ and } \frac{\left(\varepsilon + 1.5 f \hat{\sigma}_R\right)}{\hat{\sigma}_R \sqrt{\frac{2}{n}}} > z_\alpha$$

The power of this test is given by

$$\Phi\left(\frac{(1.5 f \hat{\sigma}_R - \varepsilon)}{\hat{\sigma}_R \sqrt{\frac{2}{n}}} - z_\alpha\right) + \Phi\left(\frac{(1.5 f \hat{\sigma}_R + \varepsilon)}{\hat{\sigma}_R \sqrt{\frac{2}{n}}} - z_\alpha\right) - 1. \text{ ---- (6)}$$

It can be verified that the power (6) is larger than

$$2\Phi\left(\frac{(1.5 f \hat{\sigma}_R - |\varepsilon|)}{\hat{\sigma}_R \sqrt{\frac{2}{n}}} - z_\alpha\right) - 1. \text{ ------ (7)}$$

There is an approach to find the flexible index f for given sample size $n_R = n_T = n$. The optimalfcan be obtained by selecting the minimum value f whose power exceeds 90% under the following constraint

$$1 < f < \frac{\hat{\sigma}_U}{\hat{\sigma}_R} < \frac{3}{2}$$

where $\hat{\sigma}_U = \sqrt{\frac{n_R - 1}{\chi^2_{\alpha/2, n_R - 1}}}\, \hat{\sigma}_R$

For the cases n_R =6,7,8,9,10 , the constraint 1<f<3/2 is applied for selecting f because

$\frac{\hat{\sigma}_U}{\hat{\sigma}_R} = \sqrt{\frac{n_R - 1}{\chi^2_{\alpha/2, n_R - 1}}}$ is over 3/2 when n≤17 and α= 5%.

The power of equivalence test is calculated based on the formula (7) and the results are summarized in Table 1 and Table 2 for the case where there is no mean shift and for the care where there is allowable mean shift and heterogenicity between test and reference product, respectively. Table 1 shows the power calculation by flexible margin f and sample sizes n = 6, 7, 8, 9, 10 when ε = 0 (i.e $\hat{\mu}_R = \hat{\mu}_T$). Similarly, the power calculation was conducted by f and sample sizes when ε = $1/8\sigma_R$ (i.e $\hat{\mu}_R \approx \hat{\mu}_T$) and 1/2 σ_R (i.e $\hat{\mu}_R \neq \hat{\mu}_T$) to treat allowable and unallowable mean differences. The results show that the power isless than 90% in Table 1 and 80% in Table 2 when the margin is $1.5\sigma_R$. Especially if there is a mean shift and the sample size is too small (n = 6 or 7), the power is less than 50%. It can be concluded even when the mean difference between reference and test biological products is less than $1/8\sigma_R$ (maximum allowable mean difference) is not equivalent. Thus, we suggest to give flexible margin with optimalfto have results of power more than 80% or 90%.The final results of optimal choice of flexible index fare summarized by sample sizes in Table 3 and Table 4 for the case where there is no mean shift and for the case where there is allowable mean shift and heterogenicity between the test and the reference product, respectively.

Sample Size Requirement

For determination of sample size required for performing equivalence test for CQAs that are most relevant to clinical outcomes with high risk ranking, FDA indicated that an appropriate sample size should be able to detect a possible mean shift of 1/8 σ_R under the EAC margin of $1.5\sigma_R$ with a desired power of 1-β, where β=20% or 10%.

The quantity given in (7) is a conservative approximation to the power compared to (6). Hence, the sample size calculated based on (7) is conservative. Chow and Liu [7] provided a different approximation, which leads to the following formula for sample size calculation:

$$n_R = \begin{cases} \frac{2(z_\alpha + z_{\beta/2})^2 \sigma_R^2}{\delta^2} & \text{if } \varepsilon = 0 \\ \frac{2(z_\alpha + z_{\beta/2})^2 \sigma_R^2}{(\delta - |\varepsilon|)^2} & \text{if } \varepsilon \neq 0. \end{cases}$$

Thus, under hypotheses (5), and assuming that $\sigma_R^2 = \hat{\sigma}_R^2$ is known, sample size required can be obtained by evaluating the alternative hypothesis at ε = 1/8 σ_R, the maximum mean shift allowed for demonstration of highly similarity, as follows.

$$n_R = \frac{2(z_\alpha + z_{\beta/2})^2 \hat{\sigma}_R^2}{(\delta - |\frac{1}{8}\hat{\sigma}_R|)^2} \text{ ----- (8)}$$

Based on the formula (8), minimum sample sizes for achieving the desired power (80% or 90%) are calculated and summarized in Table 5. For margin$1.5\sigma_R$, the results show that at least 12 samples are required to obtain 90% power if there is an acceptable mean difference. However, less sample sizes are requiredwhen flexible margins are applied in the equivalence test for assuring 90% power.

Table 6 provides relative efficiency of flexible margin with respect to margin$1.5\sigma_R$ for 90% power by optimal flexible index f. The values for relative efficiency are larger than 1 when the flexible margin is applied. For example, the relative efficiency is 1.71 and 2.0 if f is 1.29 and 1.37 respectively. Thus, the tab-

Table 1. Power approximation by *f* and sample sizes n (α = 5%, ε=0).

f	Margin ($1.5 f \sigma_R$)	n = 6	n = 7	n = 8	n = 9	n = 10
1 *	1.5 σ_R	0.579	0.683	0.763	0.825	0.871
1.01	1.52 σ_R	0.593	0.696	0.774	0.834	0.879
1.02	1.53 σ_R	0.607	0.708	0.785	0.843	0.886
1.03	1.55 σ_R	0.620	0.720	0.795	0.852	0.893
1.04	1.56 σ_R	0.633	0.732	0.805	0.860	0.900
1.05	1.58 σ_R	0.646	0.743	0.815	0.868	0.906
1.06	1.59 σ_R	0.659	0.754	0.824	0.875	0.912
1.07	1.61 σ_R	0.671	0.764	0.833	0.882	0.918
1.08	1.62 σ_R	0.683	0.775	0.841	0.889	0.923
1.09	1.64 σ_R	0.695	0.785	0.850	0.896	0.928
1.1	1.65 σ_R	0.706	0.794	0.857	0.902	0.933
1.11	1.67 σ_R	0.717	0.804	0.865	0.908	0.938
1.12	1.68 σ_R	0.728	0.813	0.872	0.914	0.942
1.13	1.7 σ_R	0.739	0.821	0.879	0.919	0.946
1.14	1.71 σ_R	0.749	0.830	0.886	0.924	0.950
1.15	1.73 σ_R	0.759	0.838	0.892	0.929	0.954
1.16	1.74 σ_R	0.769	0.846	0.898	0.934	0.957
1.17	1.76 σ_R	0.778	0.853	0.904	0.938	0.960
1.18	1.77 σ_R	0.787	0.860	0.909	0.942	0.963
1.19	1.79 σ_R	0.796	0.867	0.915	0.946	0.966
1.2	1.8 σ_R	0.805	0.874	0.920	0.949	0.968
1.21	1.82 σ_R	0.813	0.880	0.925	0.953	0.971
1.22	1.83 σ_R	0.821	0.887	0.929	0.956	0.973
1.23	1.85 σ_R	0.829	0.892	0.933	0.959	0.975
1.24	1.86 σ_R	0.836	0.898	0.937	0.962	0.977
1.25	1.88 σ_R	0.844	0.904	0.941	0.965	0.979
1.26	1.89 σ_R	0.851	0.909	0.945	0.967	0.981
1.27	1.91 σ_R	0.857	0.914	0.948	0.970	0.982
1.28	1.92 σ_R	0.864	0.918	0.952	0.972	0.984
1.29	1.94 σ_R	0.870	0.923	0.955	0.974	0.985
1.3	1.95 σ_R	0.876	0.927	0.958	0.976	0.986
1.31	1.97 σ_R	0.882	0.931	0.961	0.978	0.988
1.32	1.98 σ_R	0.888	0.935	0.963	0.979	0.989
1.33**	2 .00σ_R	0.893	0.939	0.966	0.981	0.990
1.34	2.01 σ_R	0.898	0.943	0.968	0.982	0.990
1.35	2.03 σ_R	0.903	0.946	0.970	0.984	0.991
1.36	2.04 σ_R	0.908	0.949	0.972	0.985	0.992
1.37	2.06 σ_R	0.913	0.952	0.974	0.986	0.993
1.38	2.07 σ_R	0.917	0.955	0.976	0.987	0.993
1.39	2.09 σ_R	0.922	0.958	0.978	0.988	0.994
1.4	2.1 σ_R	0.926	0.961	0.979	0.989	0.995
1.41	2.12 σ_R	0.930	0.963	0.981	0.990	0.995
1.42	2.13 σ_R	0.933	0.965	0.982	0.991	0.996
1.43	2.15 σ_R	0.937	0.968	0.984	0.992	0.996
1.44	2.16 σ_R	0.940	0.970	0.985	0.993	0.996
1.45	2.18 σ_R	0.943	0.972	0.986	0.993	0.997
1.46	2.19 σ_R	0.947	0.973	0.987	0.994	0.997
1.47	2.21 σ_R	0.950	0.975	0.988	0.994	0.997
1.48	2.22 σ_R	0.952	0.977	0.989	0.995	0.998
1.49	2.24 σ_R	0.955	0.978	0.990	0.995	0.998
1.5	2.25 σ_R	0.958	0.980	0.991	0.996	0.998

* *f* value for margin $1.5\sigma_R$; ** *f* value for flexible margin $2.0\sigma_R$

Table 2. Power approximation by *f* and sample sizes n (α = 5%, ε= 1/8 σ_R, 1/2 σ_R).

f	Margin ($1.5 f\ \sigma_R$)	n = 6		n = 7		n = 8		n = 9		n = 10	
		$1/8\sigma_R$	$1/2\sigma_R$	$1/8\sigma_R$	$1/2\sigma_R$	$1/8\sigma_R$	$1/2\sigma_R$	$1/8\sigma_R$	$1/2\sigma_R$	$1/8\sigma_R$	$1/2\sigma_R$
1 *	1.5 σ_R	0.382	0.000	0.497	0.000	0.593	0.000	0.673	0.000	0.738	0.000
1.01	1.52 σ_R	0.399	0.000	0.514	0.000	0.609	0.000	0.688	0.000	0.751	0.000
1.02	1.53 σ_R	0.416	0.000	0.530	0.000	0.625	0.000	0.702	0.000	0.764	0.000
1.03	1.55 σ_R	0.433	0.000	0.546	0.000	0.640	0.000	0.715	0.000	0.776	0.025
1.04	1.56 σ_R	0.449	0.000	0.562	0.000	0.654	0.000	0.729	0.000	0.788	0.050
1.05	1.58 σ_R	0.465	0.000	0.578	0.000	0.669	0.000	0.742	0.004	0.800	0.075
1.06	1.59 σ_R	0.481	0.000	0.593	0.000	0.683	0.000	0.754	0.028	0.810	0.100
1.07	1.61 σ_R	0.497	0.000	0.608	0.000	0.696	0.000	0.766	0.052	0.821	0.125
1.08	1.62 σ_R	0.513	0.000	0.622	0.000	0.709	0.000	0.778	0.076	0.831	0.150
1.09	1.64 σ_R	0.528	0.000	0.636	0.000	0.722	0.019	0.789	0.100	0.840	0.175
1.1	1.65 σ_R	0.543	0.000	0.650	0.000	0.734	0.042	0.800	0.123	0.850	0.199
1.11	1.67 σ_R	0.557	0.000	0.664	0.000	0.746	0.064	0.810	0.147	0.858	0.224
1.12	1.68 σ_R	0.572	0.000	0.677	0.000	0.758	0.087	0.820	0.170	0.867	0.248
1.13	1.7 σ_R	0.586	0.000	0.689	0.018	0.769	0.109	0.829	0.194	0.875	0.272
1.14	1.71 σ_R	0.600	0.000	0.702	0.039	0.780	0.131	0.839	0.217	0.882	0.295
1.15	1.73 σ_R	0.613	0.000	0.714	0.060	0.790	0.154	0.847	0.240	0.890	0.319
1.16	1.74 σ_R	0.627	0.000	0.726	0.081	0.800	0.176	0.856	0.262	0.896	0.342
1.17	1.76 σ_R	0.640	0.000	0.737	0.102	0.810	0.198	0.864	0.285	0.903	0.364
1.18	1.77 σ_R	0.652	0.017	0.748	0.123	0.819	0.219	0.872	0.307	0.909	0.387
1.19	1.79 σ_R	0.665	0.036	0.759	0.143	0.829	0.241	0.879	0.329	0.915	0.409
1.2	1.80 σ_R	0.677	0.056	0.770	0.164	0.837	0.262	0.886	0.351	0.921	0.430
1.21	1.82 σ_R	0.689	0.075	0.780	0.185	0.846	0.284	0.893	0.372	0.926	0.452
1.22	1.83 σ_R	0.700	0.095	0.790	0.205	0.854	0.305	0.899	0.394	0.931	0.473
1.23	1.85 σ_R	0.712	0.114	0.799	0.226	0.861	0.325	0.905	0.414	0.936	0.493
1.24	1.86 σ_R	0.723	0.133	0.808	0.246	0.869	0.346	0.911	0.435	0.940	0.513
1.25	1.88 σ_R	0.733	0.152	0.817	0.266	0.876	0.366	0.916	0.455	0.944	0.533
1.26	1.89 σ_R	0.744	0.172	0.826	0.286	0.883	0.386	0.922	0.475	0.948	0.552
1.27	1.91 σ_R	0.754	0.191	0.834	0.305	0.889	0.406	0.927	0.494	0.952	0.571
1.28	1.92 σ_R	0.764	0.210	0.842	0.325	0.895	0.425	0.931	0.513	0.955	0.589
1.29	1.94 σ_R	0.773	0.228	0.849	0.344	0.901	0.445	0.936	0.532	0.959	0.607
1.3	1.95 σ_R	0.783	0.247	0.857	0.363	0.907	0.463	0.940	0.550	0.962	0.624
1.31	1.97 σ_R	0.792	0.266	0.864	0.382	0.912	0.482	0.944	0.568	0.964	0.641
1.32	1.98 σ_R	0.800	0.284	0.871	0.400	0.917	0.500	0.948	0.585	0.967	0.657
1.33**	2.00 σ_R	0.809	0.302	0.877	0.418	0.922	0.518	0.951	0.602	0.970	0.673
1.34	2.01 σ_R	0.817	0.320	0.884	0.436	0.927	0.535	0.955	0.619	0.972	0.688
1.35	2.03 σ_R	0.825	0.338	0.890	0.454	0.931	0.552	0.958	0.635	0.974	0.703
1.36	2.04 σ_R	0.832	0.356	0.895	0.472	0.935	0.569	0.961	0.650	0.976	0.717
1.37	2.06 σ_R	0.840	0.373	0.901	0.489	0.939	0.585	0.963	0.666	0.978	0.731
1.38	2.07 σ_R	0.847	0.390	0.906	0.506	0.943	0.601	0.966	0.680	0.980	0.745
1.39	2.09 σ_R	0.854	0.407	0.911	0.522	0.947	0.617	0.968	0.695	0.981	0.758
1.4	2.10 σ_R	0.861	0.424	0.916	0.538	0.950	0.632	0.971	0.709	0.983	0.770
1.41	2.12 σ_R	0.867	0.441	0.921	0.554	0.953	0.647	0.973	0.722	0.984	0.782
1.42	2.13 σ_R	0.873	0.457	0.925	0.570	0.956	0.662	0.975	0.735	0.986	0.794
1.43	2.15 σ_R	0.879	0.473	0.929	0.585	0.959	0.676	0.977	0.748	0.987	0.805
1.44	2.16 σ_R	0.885	0.489	0.933	0.600	0.962	0.689	0.979	0.760	0.988	0.816
1.45	2.18 σ_R	0.891	0.505	0.937	0.615	0.965	0.703	0.980	0.772	0.989	0.826
1.46	2.19 σ_R	0.896	0.520	0.941	0.629	0.967	0.716	0.982	0.783	0.990	0.836
1.47	2.21 σ_R	0.901	0.535	0.944	0.643	0.969	0.728	0.983	0.794	0.991	0.845
1.48	2.22 σ_R	0.906	0.550	0.948	0.657	0.971	0.740	0.984	0.805	0.992	0.854
1.49	2.24 σ_R	0.911	0.565	0.951	0.670	0.973	0.752	0.986	0.815	0.992	0.863
1.5	2.25 σ_R	0.915	0.579	0.954	0.683	0.975	0.763	0.987	0.825	0.993	0.871

**f* value for margin $1.5\sigma_R$; ***f* value for flexible margin $2.0\sigma_R$

Table 3. Optimal flexible index *f* by *n* and power (α = 5% , ε=0).

Power (1-β)	flexible index *f*	n = 6	n = 7	n = 8	n = 9	n = 10
80%	f (EAC=1.5 $f\sigma_R$)	1.2 (1.8 σ_R)	1.11 (1.665 σ_R)	1.04 (1.56 σ_R)	1 (1.5 σ_R)	1 (1.5 σ_R)
90%	f (EAC=1.5 $f\sigma_R$)	1.35 (2.025 σ_R)	1.25 (1.875 σ_R)	1.17 (1.755 σ_R)	1.1 (1.65 σ_R)	1.05 (1.575 σ_R)

Table 4. Optimal flexible index *f* by n and power (α = 5%, $\varepsilon = 1/8\ \sigma_R$).

Power (1-β)	flexible index *f*	n = 6	n = 7	n = 8	n = 9	n = 10
80%	f (EAC=1.5 $f\sigma_R$)	1.33 (1.995 σ_R)	1.24 (1.86 σ_R)	1.17 (1.755 σ_R)	1.11 (1.665 σ_R)	1.06 (1.59 σ_R)
90%	f (EAC=1.5 $f\sigma_R$)	1.47 (2.205 σ_R)	1.37 (2.055 σ_R)	1.29 (1.935 σ_R)	1.23 (1.845 σ_R)	1.17 (1.755 σ_R)

Table 5. Minimum sample size by *f* and ε (α = 5% , $\varepsilon = 0$, $1/8\ \sigma_R$ and $1/2\ \sigma_R$).

f	Margin (1.5 f σ_R)	1 - β = 80 %			1 - β = 90 %		
		0	1/8 σ_R	1/2 σ_R	0	1/8 σ_R	1/2 σ_R
1 *	1.5 σ_R	8	10	18	10	12	22
1.01	1.52 σ_R	8	9	17	10	12	22
1.02	1.53 σ_R	8	9	17	10	11	21
1.03	1.55 σ_R	8	9	16	10	11	20
1.04	1.56 σ_R	8	9	16	9	11	20
1.05	1.58 σ_R	7	9	15	9	11	19
1.06	1.59 σ_R	7	8	15	9	11	19
1.07	1.61 σ_R	7	8	15	9	10	18
1.08	1.62 σ_R	7	8	14	9	10	18
1.09	1.64 σ_R	7	8	14	9	10	17
1.1	1.65 σ_R	7	8	13	8	10	17
1.11	1.67 σ_R	7	8	13	8	10	16
1.12	1.68 σ_R	7	8	13	8	9	16
1.13	1.70 σ_R	6	7	12	8	9	16
1.14	1.71 σ_R	6	7	12	8	9	15
1.15	1.73 σ_R	6	7	12	8	9	15
1.16	1.74 σ_R	6	7	12	8	9	15
1.17	1.76 σ_R	6	7	11	8	9	14
1.18	1.77 σ_R	6	7	11	7	8	14
1.19	1.79 σ_R	6	7	11	7	8	14
1.2	1.80 σ_R	6	7	11	7	8	13
1.21	1.82 σ_R	6	6	10	7	8	13
1.22	1.83 σ_R	6	6	10	7	8	13
1.23	1.85 σ_R	6	6	10	7	8	12
1.24	1.86 σ_R	5	6	10	7	8	12
1.25	1.88 σ_R	5	6	10	7	8	12
1.26	1.89 σ_R	5	6	9	7	7	12
1.27	1.91 σ_R	5	6	9	6	7	11
1.28	1.92 σ_R	5	6	9	6	7	11
1.29	1.94 σ_R	5	6	9	6	7	11
1.3	1.95 σ_R	5	6	9	6	7	11
1.31	1.97 σ_R	5	6	8	6	7	11
1.32	1.98 σ_R	5	5	8	6	7	10
1.33**	2.00 σ_R	5	5	8	6	7	10
1.34	2.01 σ_R	5	5	8	6	7	10
1.35	2.03 σ_R	5	5	8	6	6	10
1.36	2.04 σ_R	5	5	8	6	6	10
1.37	2.06 σ_R	5	5	8	6	6	9
1.38	2.07 σ_R	4	5	7	6	6	9

1.39	2.09 σ_R	4	5	7	5	6	9
1.4	2.10 σ_R	4	5	7	5	6	9
1.41	2.12 σ_R	4	5	7	5	6	9
1.42	2.13 σ_R	4	5	7	5	6	9
1.43	2.15 σ_R	4	5	7	5	6	8
1.44	2.16 σ_R	4	5	7	5	6	8
1.45	2.18 σ_R	4	5	7	5	6	8
1.46	2.19 σ_R	4	5	6	5	6	8
1.47	2.21 σ_R	4	4	6	5	6	8
1.48	2.22 σ_R	4	4	6	5	5	8
1.49	2.24 σ_R	4	4	6	5	5	8
1.5	2.25 σ_R	4	4	6	5	5	8

Table 6. Relative efficiency of flexible margin with respect to margin $1.5\sigma_R$.

Margin (1.5 f σ_R)	Optimal *f*	1-β = 90 %	
		ε = 0	ε = 1/8 σ_R
1.5 σ_R	1 *	1	1
1.58 σ_R	1.05	1.11	1.09
1.65 σ_R	1.1	1.25	1.20
1.76 σ_R	1.17	1.25	1.33
1.85 σ_R	1.23	1.43	1.50
1.88 σ_R	1.25	1.43	1.50
1.94 σ_R	1.29	1.67	1.71
2.00 σ_R	1.33 **	1.67	1.71
2.03 σ_R	1.35	1.67	2.0
2.06 σ_R	1.37	1.67	2.0
2.21 σ_R	1.47	2.0	2.0

* f value for margin 1.5 σ_R; ** f value for flexible margin $2.0\sigma_R$

ulated results support that the flexible margin is more efficient than margin$1.5\sigma_R$.

Concluding Remarks

For analytical similarity assessment of a given CQA, FDA recommended an equivalence test with an EAC of $1.5\sigma_R$ be used. However, FDA's recommended EAC margin has been criticized not flexible enough for demonstration of highly similarity between a proposed biosimilar product and an innovative biological product. This is because current equivalence test ignores the variability associated with the response.In addition, the EAC margin is data-dependent and usually determined based on a point estimate of σ_R. Thus,for the equivalence test with margin $1.5\sigma_R$, a larger sample is required to achieve a desired power for establishing similarity of the data with variability (mean shift). Alternatively, we propose the use of different flexible margin with respect to different sample sizes from 6 to 10. The flexible index f is selected in the range of a conservative margin and most aggressive margin ($1 < f < 1.5$). For a given sample size, the optimal flexible margin achievesdesired power 90% even thoughmargin $1.5\sigma_R$ does not. Additionally, the result indicates our flexible margins is more efficient thanmargin$1.5\sigma_R$ for demonstrating biosimilarity between a test and reference product.

References

[1]. Guidance for industry: Scientific Considerations in Demonstrating Biosimilarity to a Reference Product. United States Food and Drug Administration, Silver Spring, Maryland. FDA. 2015.

[2]. Draft Guidance for Industry: Statistical Approaches to Evaluate Analytical Similarity Guidance for Industry. United States Food and Drug Administration, Silver Spring, Maryland. (Circulated for public comments on September 27, 2017 but was officially withdrawn in June, 2018). FDA 2017.

[3]. Chow SC, Song F, Bai H. Analytical similarity assessment in biosimilar studies. The AAPS journal. 2016 May 1;18(3):670-7.

[4]. Shutter S. Biosimilar sponsors seek statistical flexibility when reference product change. FDA Pink Sheet; December 21, 2017.

[5]. Chow SC. Analytical Similarity Assessment in Biosimilar Product Development. Chapman and Hall/CRC; 2018 Sep 3.

[6]. Development of Therapeutic Protein Biosimilars: Comparative Analytical Assessment and Other Quality-Related Considerations. United States Food and Drug Administration, Silver Spring, Maryland. FDA 2019.

[7]. Chow SC, Liu JP. Design and analysis of bioavailability and bioequivalence studies. CRC press; 1999 Nov 24.