3.5.5 Sample Size

In practice, one of the major problems for a biosimilar sponsor is the availability of reference lots for analytical similarity testing. The FDA suggests that an appropriate sample size (the number of lots from the reference product and from the test product) be used to achieve a desired power (say 80%) to establish similarity based on a two-sided test at the 5% level of significance, assuming that the mean response of the test product differs from that of the reference product by $\sigma_R/8$.

Furthermore, since sample size is a function of α (type I error), β (type II error or 1 minus power), δ (treatment effect), and σ^2 (variability), it is a concern that we may have inflated the type I error rate for achieving a desired power to detect a clinically meaningful effect size (adjusted for variability) with a preselected small sample size (i.e., a small number of lots).

3.5.6 Remarks

Different assumptions may lead to different conclusions owing to the difference between mean responses of the various lots and the heterogeneity among lots. It should be noted that the difference between the mean responses of the lots may be offset by the heterogeneity across lots in the FDA's proposed equivalence test. Thus, one of the major criticisms of the FDA's proposed equivalence test procedure is the validity of the primary assumptions, especially the assumption that the difference in the mean responses between the reference product and the proposed biosimilar product is proportional to the variability of the reference product. In addition, for a given CQA, the FDA only requires that a single sample obtained from a lot be tested. In this case, an independent estimate of the variability associated with the test result of the given lot is not available. Similar comments apply to the quality range approach for CQAs from Tier 2.

3.6 RECOMMENDATIONS AND ALTERNATIVE METHODS

3.6.1 Recommendations to Current Approaches for the Assessment of Analytical Similarity

Suppose that there are K reference lots to establish EAC for the equivalence test for Tier 1 CQAs. The FDA suggests that one sample be randomly selected from each lot. The standard deviation of the reference product σ_R can be estimated based on the K test results. Let x_i, $i = 1, 2, \ldots, K$ be the test result of the ith lot. x_i, $i = 1, 2, \ldots, K$ are assumed to be independently and identically distributed with mean μ_R and variance σ_R^2. In other words, we assume that $\mu_{Ri} = \mu_{Rj} = \mu_R$ and $\sigma_{Ri}^2 = \sigma_{Rj}^2 = \sigma_R^2$ for $i \neq j$, $i, j = 1, 2, \ldots, K$. Thus, the expected value of $E(\bar{x}) = \mu_R$ and $\operatorname{var}(\bar{x}) = \sigma_R^2/K$. Under the assumption that $\mu_{Ri} \neq \mu_{Rj}$ and $\sigma_{Ri}^2 \neq \sigma_{Rj}^2$ for $i \neq j$, where μ_{Ri} and σ_{Ri}^2 be the mean and variance of the ith lot of the reference product, we have

$$\frac{\sigma_{(1)}^2}{K} \leq \operatorname{var}(\bar{x}) = \frac{\sigma_R^2}{K} \leq \frac{\sigma_{(K)}^2}{K},$$