3.5.3 Estimate of Σ_R

The FDA proposed that the equivalence test using available lot values be based mainly on the assumptions that (1) there is no lot-to-lot variability within the reference product and the test product and (2) the difference in mean responses is proportional to the variability of reference product. In practice, however, it is recognized that $\mu_{Ri} \neq \mu_{Rj}$ and $\sigma^2_{Ri} \neq \sigma^2_{Rj}$ for $i \neq j$. The differences between lots and heterogeneity among lots are major challenges to the *validity* of the FDA's proposed approaches for both equivalence testing for CQAs in Tier 1 and the concept of quality range CQAs from Tier 2. Under the assumptions that $\mu_{Ri} \neq \mu_{Rj}$ and $\sigma^2_{Ri} \neq \sigma^2_{Rj}$ for $i \neq j$, it is *not* clear what are the statistical properties/finite sample performances and corresponding impact on the assessment of analytical similarity and consequently on providing totality of the evidence to demonstrate similarity.

3.5.4 Heterogeneity within and between Test and Reference Products

Let σ^2_R and σ^2_T be the variabilities associated with the reference product and the test product, respectively. Also, let n_R and n_T be the number of lots for analytical similarity assessment for the reference product and the test product, respectively. Thus, we have

$$\sigma^2_R = \sigma^2_{WR} + \sigma^2_{BR} \text{ and } \sigma^2_T = \sigma^2_{WT} + \sigma^2_{BT},$$

where σ^2_{WR}, σ^2_{BR} and σ^2_{WT}, σ^2_{BT} are the within-lot variability and between-lot (lot-to-lot) variability for the reference product and the test product, respectively. In practice, it is very likely that $\sigma^2_R \neq \sigma^2_T$ and often $\sigma^2_{WR} \neq \sigma^2_{WT}$ and $\sigma^2_{BR} \neq \sigma^2_{BT}$ even $\sigma^2_R \approx \sigma^2_T$. This has posed a major challenge to the FDA's proposed approaches for the assessment of analytical similarity for CQAs from both Tier 1 and Tier 2, especially when there is only one test sample from each lot from the reference product and the test product. The FDA's proposal ignores lot-to-lot (between lot) variability, that is, when $\sigma^2_{BR} = 0$ or $\sigma^2_{BR} = \sigma^2_{WR}$. In other words, sample variance based on x_i, $i = 1,\ldots, K$ from the reference product may underestimate the true σ^2_R, and consequently may not provide a fair and reliable assessment of analytical similarity for a given quality attribute.

In practice, it is well recognized that $\mu_{Ri} \neq \mu_{Rj}$ and $\sigma^2_{Ri} \neq \sigma^2_{Rj}$ for $i \neq j$, where μ_{Ri} and σ^2_{Ri} are the mean and variance of the ith lot of the reference product. A similar argument is applied to the proposed biosimilar (test) product. As a result, the selection of reference lots for the estimation of σ_R is critical for the proposed approach. The selection of reference lots has an impact on the estimation of σ_R and consequently on the EAC. Suppose there are K reference lots available and n lots will be tested for analytical similarity. The FDA suggests using the remaining $K - n$ lots to establish EAC to avoid selection bias. It sounds a reasonable approach if $K \gg n$. In practice, however, few lots are available. In this case, the FDA's proposed approach may not be feasible.