TABLE 3.4
Assessment of Analytical Similarity for CQAs from Tier 1

Number of RP Lots	Number of TP Lots	Selection of *c*	Test Size (Confidence Interval)	Statistical Power at (1/8) × RP SD (%)
6	6	1.5	9% (82% CI)	74
7	7	1.5	8% (84% CI)	79
8	8	1.5	7% (86% CI)	83
9	9	1.5	6% (88% CI)	86
10	10	1.5	5% (90% CI)	87

According to Table 3.4, there is 79% power for 84% CI of $\hat{\Delta} = \hat{\mu}_T - \hat{\mu}_R$ to fall within ±EAC, assuming that the number of lots is 7 and true difference between TP and RP is $\sigma_R/8$.

This approach has inflated alpha from 5% to 16%. Note that for a fixed regulatory standard *c*, the sponsor may appropriately select sample size (the number of lots) for achieving a desired power (for detecting a $\sigma_R/8$ difference) and significance level for analytical similarity assessment. As can be seen from the above, if one wishes to reduce the test size (i.e., α level) from 8% to 5%, 10 TP lots need to be tested. Testing 10 TP lots will give an 87% power for detecting a $\sigma_R/8$ difference.

3.4.2 Quality Range Approach for Tier 2

For Tier 2, the FDA suggests that analytical similarity be assessed on the basis of the concept of quality ranges, that is, $\pm x\sigma$, where σ is the standard deviation of the reference product and *x* should be appropriately justified. Thus, the quality range of the reference product for a specific quality attribute is defined as $\left(\hat{\mu}_R - x\hat{\sigma}_R,\ \hat{\mu}_R + x\hat{\sigma}_R\right)$. Analytical similarity would be accepted for the quality attribute if a sufficient percentage of test lot values (e.g., 90%) falls within the quality range.

For a given critical attribute, the quality range is set based on test results of available reference lots. If $x = 1.645$, we would expect 90% of the test results from reference lots to lie within the quality range. If *x* is chosen to be 1.96, we would expect that about 95% test results of reference lots will fall within the quality range. As a result, the selection of *x* could impact the quality range and consequently the percentage of test lot values that will fall within the quality range. Thus, the FDA indicates that the standard deviation multiplier (*x*) should be appropriately justified.

The quality range approach for comparing populations between a proposed biosimilar product and a reference product is a reasonable approach under the assumption that $\mu_T = \mu_R$ and $\sigma_T = \sigma_R$. Under this assumption, we expect that a high percentage (say 90%) of test values of the test product will fall within the quality range obtained based on the test values of the reference product. Thus, one of the major criticisms of the quality range approach is that it ignores the fact that there are differences in population mean and population standard deviation