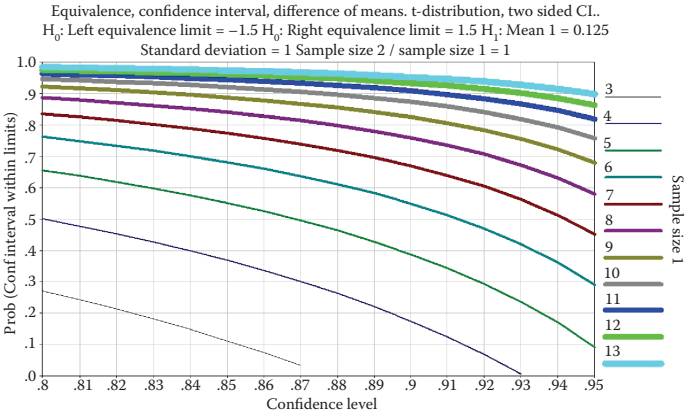

Figure 9.14 Probability that confidence interval is within limits as a function of confidence level and number of lots selected for side-by-side testing.

Figure 9.14 shows the power (probability of confidence interval within limits) as a function of sample size and the study size or the α value expressed as $(1–2\sigma) \times 100$.

When seven lots are used side by side, the power drops to 67%. Generally, 80% or a higher power will be sufficient to convince the FDA of the suitability of the number of lots selected. The number of samples tested is also justified based on the confidence level used; for example, in the case if 81.4% or 85.2% is used, the number of samples required will change, as a minimum.

A significant concern arises when testing Tier 0 and Tier 3 attributes where no numerical calculations are required; in the case of Tier 0, all tested samples must either be identical to the reference product or meet the legacy value (e.g., total mass); for Tier 3, it is simply a display of numbers (where available) and an overall comparison of the shape and details of the graphical output. While one can argue that a smaller number of lots will be sufficient for this exercise, this concern becomes moot due to the fact that the FDA wants you to conduct all tests for the chosen test lots. The developer must present data for Tier 0 and Tier 3 testing for the same lots used for Tier 1 and Tier 2 testing.

9.8.21 Blinding

Prior to conducting the side-by-side analytical similarity testing, the biosimilar developer would create a protocol that will include all acceptance criteria as well as blind the test and reference samples; where a larger number of samples are available, this will be preceded by a random selection of lots. However, when establishing EAC or other acceptance attributes using a separate set of lots, there is no need for blinding the reference samples.