comparative assessment of quality attributes in drug development (EMA, 2013b). Without being exhaustive, some considerations are presented here.

The use of an inappropriate statistical approach is not uncommon. A common intuitive but inappropriate statistical method consists of establishing equivalence by setting tolerance intervals (TIs) for the reference product. Under this approach, a TI calculated on historical data for the reference product is compared with a TI calculated on the test product data. If the test product TI fits inside the reference product TI, the test product is considered similar to the reference product. From a statistical point of view, the main objection against this approach is that it is based on an erroneous hypothesis. Actually, the question of significance is not even addressed in the TI approach since it is not based on probabilistic inference. Because the TI approach is not a hypothesis testing method, its hypothesis is implicit: if the test product is not significantly different from the reference product, then it is significantly similar. This is a wrong assumption. Various seemingly paradoxical situations arise in connection with the wrong hypothesis. For example, the reference TI is wider with smaller reference sample size, higher reference variability, and inconsistency of the production process over time. All these undesirable characteristics of the reference product, in a rather contradictory manner, make it easier to establish the similarity with the test product.

The correct approach to be used in establishing biosimilarity is equivalence testing, which is based on a correct hypothesis. In this approach, a 90% confidence interval (CI) is established about the difference between reference and test sample means (mean difference). If the mean difference and the CI fit inside some equivalence limits (ELs) established prior to the test, the reference and test samples are significantly equivalent (similar) (Figure 5.1). Otherwise, they are not significantly similar.

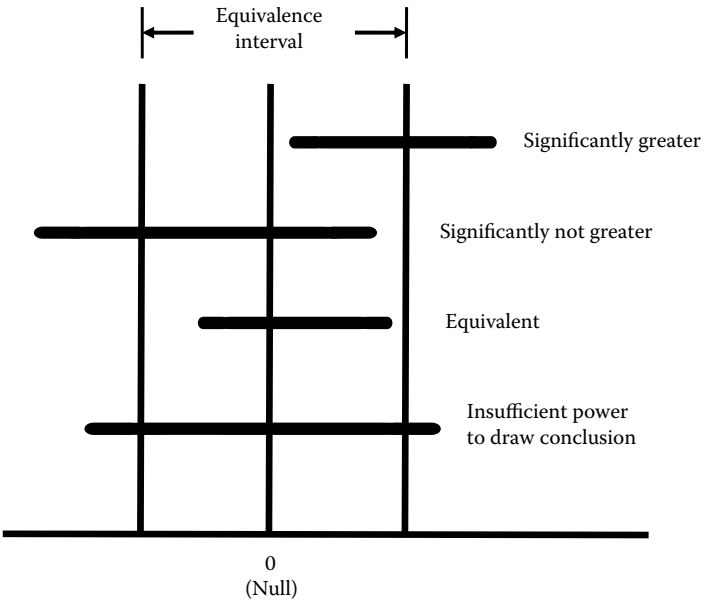

FIGURE 5.1 Equivalence tests, equivalent limits, and types of significance.