TABLE 10.1 Degree of Certainty.

| Statistical Decision | Truth | |
|---|---|---|
| | $H_0$ Is True | $H_1$ Is True |
| Fail to reject $H_0$ | Correct decision | Type II error (beta, $\beta$) |
| Reject $H_0$ (accept $H_1$) | Type I error (alpha, $\alpha$) | Correct decision |

## V. DEGREE OF CERTAINTY: ALPHA ($\alpha$) AND BETA ($\beta$)

In reality (truth), the null hypothesis is either true or false. The statistical decision process is set up so that it either rejects the null hypothesis or fails to reject the null hypothesis. This leaves four possibilities of the outcome of our statistical testing. In two of the four possibilities, we will make the correct decision. However, there are two possibilities where we will make an error in our decision process. The type I error, commonly referred to as the alpha ($\alpha$) value, is the probability of incorrectly rejecting the null hypothesis or the false-positive rate. From a society perspective, this type of error is to be minimized, as it would not be acceptable to have products on the market that do not work. Commonly, the type I error rate is set to 5% or less. However, the actual type I error rate has to be determined based on the design of the study. Studies with more than one objective being tested (ie, more than one hypothesis test constructed) may need to adjust the $\alpha$ level to preserve the true probability of observing a significant result. Additional factors that could impact the alpha ($\alpha$) level include a study that has more than two treatment arms or that includes interim analyses of the data. These items lead to what are commonly referred to as multiplicity problems. If we test a single set of hypotheses then the type I error rate is maintained at the level we set. However, if we test more than one set of hypotheses, the type I error rate is inflated, and for at least one test is larger than we set.

The second type of error, type II error or the beta ($\beta$) value, is the probability of failing to reject the null hypothesis when the alternative hypothesis is true. It is a false-negative result and represents risk from the researcher's standpoint. Here, while the alternative hypothesis is true, it is not confirmed by mistake. As such, a truly effective treatment may be overlooked or discounted due to the finding. Here, researchers want to keep this value small, typically in the 10−20% range. A related concept to type II error is power, which is calculated as $1-\beta$ in proportion or $100-\beta$ if we are using percentages. Hence, if $\beta = 10\%$, then power is 90%.

## VI. STATISTICAL METHODS FOR DETERMINING TRIAL SIZE

A variety of statistical analysis techniques can be applied to the statistical problems encountered in clinical trials. For each of these techniques, there is an appropriate method for determining sample size appropriate for use in that setting. This section provides a starting point for calculating sample sizes appropriate when utilizing the most commonly encountered data analysis techniques.

## VII. CONTINUOUS VARIABLES: TESTING THE DIFFERENCE BETWEEN TWO MEANS

Clinical trials commonly have endpoints which can be measured on a continuous numeric scale, such as changes from baseline in HIV-1 RNA for studies of HIV subjects or changes in $FEV_1$ for studies in subjects with chronic obstructive pulmonary disease. While