stratification variables, use of more than two treatment groups, missing data, differential follow-up times in studies utilizing a time-to-event variable as the primary endpoint. Methods are available to address each of these issues as well as countless other issues encountered in clinical trials. However, many of these methods may require someone formally trained in statistics to successfully determine and implement the proper method for a given situation.

It is worth stating that even in the statistical community, sample size calculations are viewed as an art backed by science. Given the sensitivity of calculations to seemingly modest changes in assumptions, it is very important to investigate the sample sizes necessary for a reasonable range of values for the expected treatment effect or variance. This exercise will provide a review of the risk/benefit of choosing a particular sample size over another. While this review can sometimes be painful, it is most often enlightening, as it is best to set the trial size armed with the most information available rather than for it to be arbitrarily picked using only scant information. In the case that there are little to no good data on which to base sample size calculations, it is possible to include procedures to reestimate the sample size during the conduct of the study. These methods, called sample size reestimation techniques, use interim data to recalculate the sample size, while preserving the type I error rate of the study. The FDA issued draft guidance on adaptive designs which included its current thinking on sample size reestimation (14). Methods which are focused on reestimation using blinded data are available and have been successfully utilized in clinical studies.

After the conduct of a study and review of the results, statisticians are sometimes asked "What was the power of the study based on the treatment differences that were observed in the study?" In essence, the investigator is asking for a post-hoc analysis of power. As stated previously, power is the probability of correctly rejecting the null hypothesis. From a statistical perspective, once the study has been conducted, the power as related to that study is either 100% (ie, the null hypothesis was correctly rejected) or 0% (ie, the null hypothesis was not correctly rejected). With that said, evaluating power for a future study based on the observed effects from the given study may be a valuable tool for researchers as they evaluate the future development of the drug candidate. For these post-hoc analyses of power, the formulas as presented in this chapter may be applied using the observed treatment differences or treatment effect, as appropriate. If it is found that the revised assumptions would have required a larger sample size, then these estimates should be considered when planning future studies.

XV. WRITING A SAMPLE SIZE SECTION OF A CLINICAL STUDY PROTOCOL

Every Clinical Study Protocol should include some justification of the sample size utilized in the study. The justification should include the following information: the primary endpoint, the null and alternative hypotheses being tested, the test statistic, the type I and type II error rates, estimates of treatment effect and/or variability, and any other information that would be helpful should someone need to reproduce the calculation. An example of a sample size justification which could be included in a Protocol is as follows.

[14]U.S. Department of Health and Human Services. Food and Drug Administration. Guidance for industry. Adaptive design clinical trials for drugs and biologics; February 2010 (50 pp.).