Background Information for Advisory Committee Meeting
On April 14, 2004

Criteria for Establishing Bio-inequivalence between Two Drug Products

Introduction

Bioequivalence is defined as "the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study...". To evaluate bioequivalence, the U.S. Food and Drug Administration (FDA) has employed a testing procedure termed *the two one-sided tests procedure* (i) to determine whether the average values for the pharmacokinetic measures from the test and reference products are comparable. This procedure involves the calculation of a confidence interval for the ratio between the average values of the test and reference product. FDA considers a test product to be bioequivalent to a reference product if the 90% confidence interval of the geometric mean ratio of AUC and C_{max} between the test and reference fall within 80-125% (ii).

Recently, the FDA has received several studies intended to show bio-inequivalence between two drug products, for example an innovator company might conduct a study to challenge FDA's approval of generic versions of its drug product. Although there has not been a formal definition of the concept of bio-inequivalence in the regulation, intuitively, the concept of bio-inequivalence is not hard to perceive, given the well-defined concept of bioequivalence. However, there are no clear criteria to guide sponsors in conducting bio-inequivalence studies and FDA reviewers in assessing the validity of such bio-inequivalence studies. Because of a lack of a clear definition of bio-inequivalence, there has been some confusion and misunderstanding by the public.

Many questions arise when evaluating a bio-inequivalence claim. A typical question is if it is appropriate to claim bio-inequivalence when the two-sided 90% confidence intervals for the ratios of the PK parameters do not fall inside the bioequivalence interval? There are numerous literature reports that claim bio-inequivalence based on a failed bioequivalence study without identification of the causes of the study failure. There are many ways that a bioequivalence study can fail, including an insufficient number of subjects. Many products that were claimed to be bio-inequvialent in the literature might well be bioequivalent if the studies were conducted appropriately. Therefore, it is imperative to develop and establish a bio-inequivalence criterion to clarify confusion and misunderstanding in the public.

In these presentations, we first introduce the concepts of bio-inequivalence and present a statistical explanation for the proposed criterion to assess bio-inequivalence. We then discuss several statistical strategies to assess bio-inequivalence studies with three pharmacokinetic parameters (Cmax, AUCt and AUC∞). The goal is to propose a set of

criteria that are scientifically sound, statistically valid, and easy to use and to provide sufficient information to stimulate discussion on the evaluation of bio-inequivalence.

**The concept of bio-inequivalence and test criteria**

FDA's bioequivalence criteria require the 90% confidence interval of the ratio of the geometric means of the test and reference drug products to be within the bioequivalence interval [80%, 125%]. The definition of the bio-inequivalence region then is simply the region that lies outside the bioequivalence interval, i.e., (0, 80%) or (125%, 8). Now the question is why a study failing to show bioequivalence cannot be used to claim bio-inequivalence. Once this question is answered, it will be a little easier to understand the statistical criteria proposed for bio-inequivalence claims.

To answer this question, we need to understand statistically how the criteria for bioequivalence are formed. To test bioequivalence, the null hypothesis is set to be the bio-inequivalence region and the alternative hypothesis to be the bioequivalence interval. The goal is to see if bio-inequivalence can be rejected so that we may conclude that bioequivalence is true. For this purpose, it is important for the probability of an error that wrongfully rejects bio-inequivalence, and therefore falsely concludes bioequivalence, to be small. This error is usually controlled at the level of 0.05, which is the so-called significance level or the type I error rate. To reject the bio-inequivalence region, we need to perform two one-sided tests, each controlling the type I error rate at the level of 0.05. The maximum error rate in the two tests are actually controlled at the level of 0.05. The statistical criteria for rejecting bio-inequivalence and claiming bioequivalence are to have two-sided 90% confidence intervals (for the geometric mean ratio for each of the three PK parameters) that are each within the bioequivalence interval. This procedure based on 90% confidence intervals is identical to carrying out the two one-sided tests described above.

To address whether failing to show bioequivalence demonstrates bio-inequivalence, we need to understand that in a bioequivalence test we usually do not control the error of wrongfully failing to conclude bioequivalence. If this error were controlled at a very low level, this would be equivalent to having very high power in a bioequivalence test. In order for both the significance level and power to be controlled at high level, a large sample size will generally be required, which will increase the cost of the study. For example, if we set the power to be 85%, and assuming the variance is 0.04, the sample size required is about 22, given the ratio of the two geometric means deviates from 1 by no more than 5%. In this case, the test could have about a 15% chance to fail to show bioequivalence even when the two drugs are truly equivalent. If the variance is larger than 0.04 and the ratio of the two geometric means deviates from 1 by more than 5% but still within the bioequivalence interval, the power could be much lower than 85% for the given sample size of 22. That is, the chance of failing to show bioequivalence would be much higher than 15% even when the two drugs are equivalent. Therefore, because there is less control over the probability of failing to show bioequivalence, it is inappropriate to use a study that fails to show bioequivalence to claim bio-inequivalence.

Then why should the bio-inequivalence criterion be that the upper (lower) limit of the two-sided 90% CI should be less (greater) than 80% (125%)? As mentioned before, usually it is not realistic to control both types of errors, i.e., wrongfully rejecting bio-inequivalence and bioequivalence. A reasonable study only tightly controls one type of error. Therefore when testing for bio-inequivalence, we would like to control the error of wrongfully rejecting bioequivalence to be small. To be consistent with the bioequivalence testing, the error rate is also chosen at the level of 0.05. To reject bioequivalence, we also need to perform two one-sided tests, however, the level of each test may need to be 0.05. For one of the two tests to be significant at the 0.05 level, either the upper limit of the two-sided 90% CI has to be less than 80% or the lower limit to be above 125%.

Theoretically, it is possible for the type I error to reach 0.10 when a two-sided 90% CI is used to assess bio-inequivalence. However, this is true only when the variance of the estimated treatment difference (the ratio of geometric means) is very large. For typical crossover bio-inequivalence trials, such a large variance may not be a realistic possibility. Therefore, the type I error rate should be maintained at the level of 0.05 when two-sided 90% CI is used.
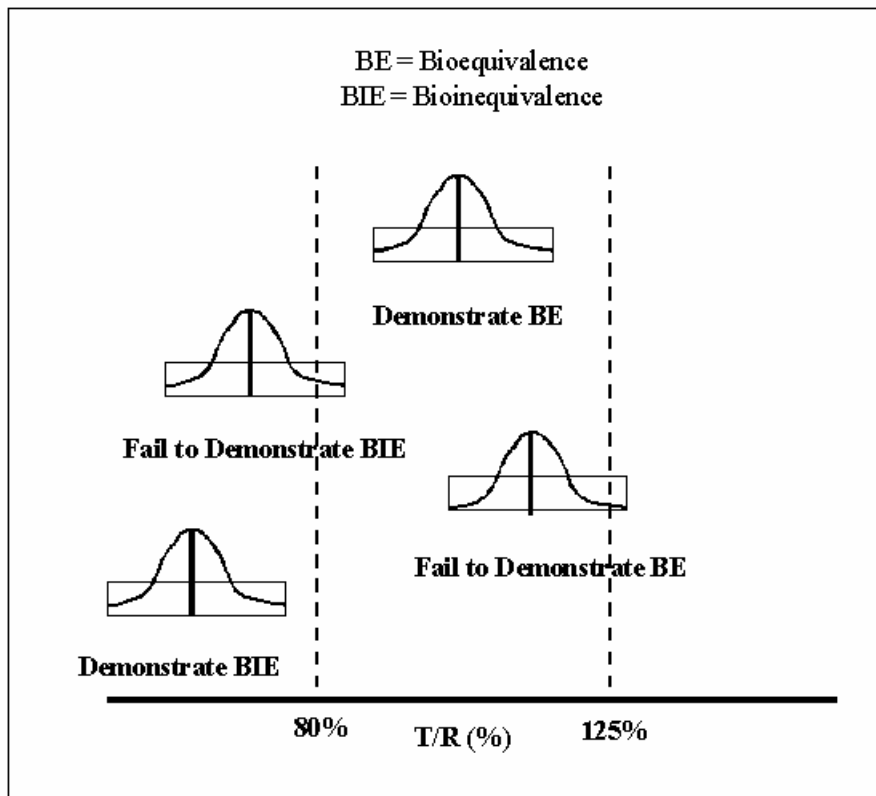
The above figure illustrates the different possible outcomes. A study with the two-sided 90% confidence interval completely between 80-125% demonstrates bioequivalence and allows market access. A study with the two-sided 90% confidence interval completely outside 80-125% demonstrates bio-inequivalence and may be grounds for market exclusion. A study with the point estimate within 80-125% but the two-sided 90% confidence interval outside of 80-125% fails to demonstrate bioequivalence. A study with the point estimate outside 80-125% but the two-sided 90% confidence interval overlapping 80-125% fails to demonstrate bio-inequivalence. Both of the failing cases would require studies with larger sample sizes to draw a definitive regulatory conclusion.

Evaluating the three PK parameters collectively:

As mentioned earlier, based on the interpretation of regulation, FDA usually requires three pharmacokinetic parameters (Cmax, AUCt, and AUC∞) to show bioequivalence. All the two-sided 90% confidence intervals for the ratios of the geometric means for the three pharmacokinetic parameters must be within the bioequivalence interval to demonstate bioequivalence. If the 90% confidence interval for just one of the three pharmacokinetic parameters does not fall completely within the bioequivalence interval, the study has not demonstrated that the two drugs are bioequivalent. However, the statistical criteria for testing bio-inequivalence using all the three pharmacokinetic parameters will not be as simple. Here we discuss several strategies that potentially can be used for assessing bio-inequivalence using three pharmacokinetic parameters. The evaluation of the strategies is based on both the error rate of wrongfully rejecting bioequivalence and power for detecting bio-inequivalence under various correlation structures.

One strategy that seems intuitive is to have at least one of the three pharmacokinetic parameters satisfy the statistical criteria for bio-inequivalence, i.e., the upper (lower) limit of the two-sided 90% CI to be less (greater) than 80% (125%). However, this strategy could potentially inflate the error rate of wrongfully rejecting bioequivalence above the level of 0.05 if the three pharmacokinetic parameters are not highly correlated.

The second strategy that is just the opposite of the first one discussed above is to require all the three pharmacokinetic parameters to satisfy the statistical criteria for bio-inequivalence. This strategy can certainly control the error rate of wrongfully rejecting bioequivalence under all correlation structures. However, it may not always provide adequate power under alternatives that are of interest.

The third strategy that could protect the error rate of wrongfully rejecting the bioequivalence is to pre-specify one pharmacokinetic parameter for bio-inequivalence testing. For example, one could pre-specify AUCt and completely ignore the results of the other two pharmacokinetic parameters. However, this strategy only has good power when AUCt is the parameter most likely to demonstrate bio-inequivalence. If only Cmax of the two drugs were bioinequivalent, then pre-specifying AUCt would give the test zero power to detect bio-inequivalence.

It is possible to develop a compromise approach. Instead of requiring all the three pharmacokinetic parameters to satisfy the statistical criteria for bio-inequivalence with two-sided 90% confidence intervals as the measurement, we could have flexible width of the one-sided confidence intervals, while controlling the error rate at the level of 0.05 under all correlation structures. For example, it is possible to have one pharmacokinetic parameter use a two-sided 91% confidence interval (slightly wider than 90% confidence interval) to show bio-inequivalence, while the second pharmacokinetic parameter uses a two-sided 87% confidence interval (narrower than 90% confidence interval) and the third pharmacokinetic parameter uses two-sided 80% confidence interval (much narrower than 90% confidence interval). For this strategy, it does not matter which pharmacokinetic parameters uses which confidence interval. The advantage of this strategy is to use narrower confidence intervals to increase power to show bio-inequivalence, although at the cost of slightly widening one pharmacokinetic parameter's confidence interval. Notice this strategy is developed using the assumption of a normal distribution. If the normal assumption is inadequate, it is possible to derive slightly different widths of confidence intervals under other distributions.

We would like to note here that it might not be necessary to control all the correlation structures, as it may be very unlikely for the three pharmacokinetic parameters to be highly correlated (the correlation coefficient is above 0.99). For the strategy with flexible confidence intervals discussed above, the error inflation occurs at correlation structures that are highly correlated. If it is possible to show that the three pharmacokinetic parameters are unlikely to have correlation higher than 0.99, the strategy can be further relaxed.

Summary

In summary, this meeting will introduce and clarify the concepts of bioequivalence, bio-inequivalence, failing to demonstrate bioequivalence, and failing to demonstrate bio-inequivalence. We will explain the statistical criteria used to claim bio-inequivalence for one pharmacokinetic parameter. We will present the pros and cons of several strategies to collectively evaluate the three pharmacokinetic parameters. Our main focus for the discussion of bio-inequivalence criteria is on statistical issues related to power and error. Other statistical issues, such as an inadequate statistical model, study design, as well as conduct of the studies, may also impact bioequivalence and bio-inequivalence testing.

Discussion Topic for ACPS Meeting

April 14, 2004

Do you agree with the distinction between demonstrating bio-inequivalence and failure to demonstrate bioequivalence?

What is your preferred method for evaluating the three pharmacokinetic parameters for bio-inequivalence?

- If bio-inequivalence is demonstrated for any one pharmacokinetic parameter, then bio-inequivalence is demonstrated for the products.

- bio-inequivalence must be demonstrated for all three pharmacokinetic parameters for bio-inequivalence to be demonstrated for the products.

- There should be one preselected pharmacokinetic parameter used for bio-inequivalence testing. If so, which one?

- The three pharmacokinetic parameters should be evaluated for bio-inequivalence with statistical corrections to the level of significance for each parameter in order to maintain an overall significance level of 0.05.

References

i   D.J. Schuirmann. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J. Pharmacokinet. Biopharm. 15: 657-680 (1987).

ii U.S. Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products--General Considerations, Office of Training and Communications, Division of Communications Management, Drug Information Branch, HFD-210, Rockville MD 20857, March 2003.