***Supplemental Examples For Illustrating Statistical Concepts Described in the VICH In Vivo Bioequivalence Draft Guidance GL52***

## EXAMPLE: SAMPLE SIZE ESTIMATION:

Scenario: The study will be conducted as a two-period, two-treatment, two-sequence crossover design where subjects receive a single dose of the test and reference products. For the sake of this example, animals are sorted by identification (ID) number and assigned to sequence 1 or 2 completely at random. The study design is as follows:

Sequence 1: Period 1 = Test; Period 2 = Reference
Sequence 2: Period 1 = Reference; Period 2 = Test

To estimate the number of subjects needed in the study, a pilot crossover study was conducted. The expected ratio of the test and reference means = 1.05. The anticipated within subject error = 15% Coefficient of variation (%CV). The iterative equation used for estimating the number of subjects is as follows:

If $\theta = 1$, then: $n \geq [t(\alpha, 2n\text{-}2) + t(\beta/2, 2n\text{-}2)]^2 [CV/\ln 1.25]^2$
If $1 < \theta < 1.25$, then: $n \geq [t(\alpha, 2n\text{-}2) + t(\beta, 2n\text{-}2)]^2 [CV/(\ln 1.25 - \ln \theta)]^2$
If $0.8 < \theta < 1$, then: $n \geq [t(\alpha, 2n\text{-}2) + t(\beta, 2n\text{-}2)]^2 [CV/(\ln 0.8 - \ln \theta)]^2$

Where $1\text{-}\beta$ = the power of the study (80%); $\alpha$ = the Type 1 error for each side of the 90% confidence interval (= 0.05), n = the number of subjects per sequence (and the total number of subjects = N = 2n), and $\theta$ = the anticipated ratio of the test/reference mean.

Based upon this equation and the results of the pilot study, sample size estimation procedure was as follows:

If we begin our sample size estimation with n = 5 (N=10), the equation would be:

$5 \geq [1.860+0.889]^2 * [0.15/(\text{Ln } 1.25 - \text{Ln } 1.05)]^2 = 5.59.$

Since this is not a true statement, we need to try the next highest value, n = 6 (N=12). In this case, the calculation is as follows:

$6 \geq [1.812+0.879]^2 * [0.15/(\text{Ln } 1.25 - \text{Ln } 1.05)]^2 = 5.40.$

When we use n=6, the conditional statement is now correct. Therefore, our sample size estimate (the number of subjects per sequence) = 6 and the total number of subjects included in this study should be no less than 12.

With N=12, the results of the simulated bioequivalence trial are as follows:

## EXAMPLE: BIOEQUIVALENCE DATA STATISTICAL ANALYSIS:

Scenario: The study will be conducted as a two-period, two-treatment, two-sequence crossover design where subjects receive a single dose of the test and reference products. For the sake of this example, subjects are sorted by identification ID number and assigned to sequence 1 or 2 completely at random. The study design is as follows:

Sequence 1: Period 1 = Test; Period 2 = Reference
Sequence 2: Period 1 = Reference; Period 2 = Test

With N=12, the results of the simulated trial were as follow (Table 1):

Table 1: Dataset from a simulated bioequivalence trial

| Animal | Sequence | Period | Treatment | Value |
|---|---|---|---|---|
| 1 | 1 | 2 | Reference | 86.76 |
| 2 | 1 | 2 | Reference | 72.23 |
| 3 | 1 | 2 | Reference | 102.10 |
| 4 | 1 | 2 | Reference | 138.42 |
| 5 | 1 | 2 | Reference | 120.67 |
| 6 | 1 | 2 | Reference | 81.83 |
| 7 | 2 | 1 | Reference | 84.91 |
| 8 | 2 | 1 | Reference | 92.84 |
| 9 | 2 | 1 | Reference | 114.42 |
| 10 | 2 | 1 | Reference | 119.48 |
| 11 | 2 | 1 | Reference | 95.32 |
| 12 | 2 | 1 | Reference | 105.77 |
| 1 | 1 | 1 | Test | 93.38 |
| 2 | 1 | 1 | Test | 78.81 |
| 3 | 1 | 1 | Test | 108.81 |
| 4 | 1 | 1 | Test | 154.68 |
| 5 | 1 | 1 | Test | 131.96 |
| 6 | 1 | 1 | Test | 71.30 |
| 7 | 2 | 2 | Test | 75.80 |
| 8 | 2 | 2 | Test | 96.98 |
| 9 | 2 | 2 | Test | 129.46 |
| 10 | 2 | 2 | Test | 131.24 |
| 11 | 2 | 2 | Test | 91.27 |
| 12 | 2 | 2 | Test | 90.47 |

Prior to analysis, all data were transformed to the natural logarithm. The statistical model used in the analysis included sequence, period and treatment as fixed effects and animal-nested-within-sequence as a random effect. There are numerous statistical programs and program specifications that can be used. All correctly programmed analyzed data should give the following results (Tables 2 and 3).

Table 2: Test of fixed effects

| Effect | Numerator Degrees of Freedom | Denominator Degrees of Freedom | F-Value | Probability of a Greater F |
|---|---|---|---|---|
| Sequence | 1 | 10 | <0.01 | 0.9527 |
| Period | 1 | 10 | 0.89 | 0.3667 |
| Treatment | 1 | 10 | 0.43 | 0.5274 |

Table 3: Difference and confidence interval

| Difference | Standard Error | Lower 90% Limit | Upper 90% Limit |
|---|---|---|---|
| 0.1958 | 0.02991 | –0.0346 | 0.0738 |

Using the statistical output information, we calculated the confidence bounds:

Lower BE Bound = Exp(-0.0346) = 0.97
Upper BE Bound = Exp(0.0738) = 1.08

If both the lower and upper BE bounds are between 0.80 and 1.25, bioequivalence is established for that value.

**EXAMPLE: SEQUENTIAL ANALYSIS:**

The use of a sequential analysis allows for an opportunity to recalculate sample size based upon the observed study variance by altering α to accommodate an interim analysis of the dataset. However, it is important to note that to avoid inflation of the Type I error, the sequential analysis does not allow for sample size adjustments based upon incorrect assumptions of the ratio of treatment means.

There are several types of sequential designs that can be used for bioequivalence studies. The following is only one possible example of how this analysis may be executed. The primary reference used in the writing of this example is Potvin et al., *Pharm Stat*, 7:245-262.

**Method B**: Test for bioequivalence at the $\alpha = 0.0294$ level, regardless of power. At Stage 1, re-calculate necessary sample size for the entire study. Confidence intervals can be estimated at both Stages 1 and 2. A schematic diagram of the steps in the Method B version of the sequential analysis (based upon Potvin et al., 2008) is provided in Figure 1.

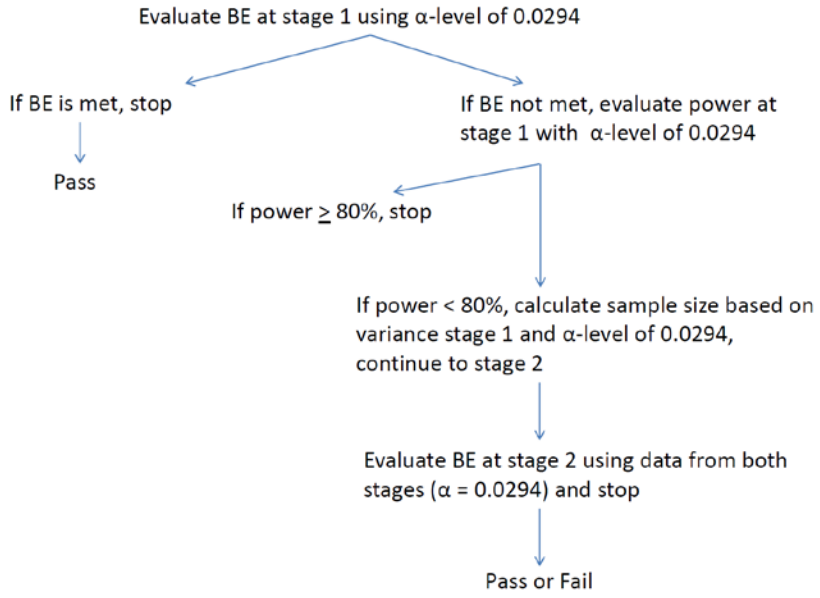


**Figure 1: Schematic diagram of the steps involved in Method B version of the sequential analysis.**

Evaluate BE at Stage 1 using an α level of 0.0294, regardless of the power achieved. If the BE criteria are met or if the study power is equal to or greater than 80%, then no additional subjects should be tested. Conversely, if the BE criteria are not met, the sample size necessary to achieve 80% power should be calculated based on the information derived at Stage 1. At Stage 2, the confidence intervals are re-calculated at a level of $\alpha = 0.0294$ using data generated at Stages 1 and 2.

Bioequivalence evaluation does not extend beyond Stage 2, regardless of the outcome.

*SCENARIO*: For the sake of this example, we will use the following Stage 1 assumptions:

- We estimate a residual error of 15% CV and a ratio of the test/reference means of 0.90.
- We will conduct Stage 1 with 20 subjects (10 per sequence).

At Stage 1, we did not meet the BE criteria of 0.80 to 1.25. By having pre-planned the use of a sequential analysis (using Method B), we can combine the data generated with the original 20 subjects with that obtained from the additional subjects at Stage 2. So, the first question is how many total subjects will be needed to demonstrate product BE in this situation and how many additional subjects will need to be included in our trial at Stage 2?

To answer that question, we need to plug these observed values into the equation for sample size:

SAMPLE SIZE EQUATION:

If $0.8 < \theta < 1$, then: $n \geq [t(\alpha, 2n\text{-}2) + t(\beta, 2n\text{-}2)]^2 [CV/ (\ln 0.8 - \ln \theta)]^2$

Our first estimate is that N = 40 (20 per sequence). In that case, our calculation is as follows (keeping in mind that $\alpha$ is now set at 0.0294);

$[t(\alpha, 2n\text{-}2) = 1.948, t(\beta, 2n\text{-}2) = 0.851,$
$[t(\alpha, 2n\text{-}2) + t(\beta, 2n\text{-}2)]^2 = 7.834$
$[CV/ (\ln 0.8 - \ln \theta)]^2 = [0.2/(\ln 0.8 - \ln 0.90)]^2 = 2.883$
$n \geq [t(\alpha, 2n\text{-}2) + t(\beta, 2n\text{-}2)]^2 [CV/ (\ln 0.8 - \ln \theta)]^2 = 22.587$

We repeat our sample size calculation, this time using n = 23. If we plug the corresponding t values into the equation, the results again indicate an n of at least 23 subjects per sequence (i.e., convergence is achieved). Based upon this outcome, we conclude that the total number of study subjects (N) needed to meet the BE criteria with 20% CV and $\theta = 0.90$ (at $\alpha = 0.0294$) will be 46. Since we already have Stage 1 data generated with N=20 (i.e., 10 subjects per sequence), Stage 2 will require an additional 13 subjects per sequence (N=26) based upon our revised estimates.

For the sake of comparison, if we had instead opted to do a pilot study prior to executing the pivotal BE trial, the total number of subjects (N) that would have been needed (at $\alpha = 0.05$ rather than $\alpha$=0.0294) with the current estimates of $\theta = 0.90$ and CV = 0.20 would have been 38 rather than 46.