Statistical Information from the June 1999 Draft Guidance and Statistical Information for In Vitro Bioequivalence Data Posted on August 18, 1999

Statistical analysis method recommendations for in vitro nonprofile bioequivalence data, to accompany the draft guidance for industry entitled *Bioavailability and Bioequivalence Studies for Nasal Aerosols and Nasal Sprays for Local Action* (April 2, 2003), are under development. At a later time, the analysis methods will be posted. Until these methods are prepared, two documents, both in need of updating, are being made available in the present document. These documents are the original statistical information taken from the earlier June 1999 draft of this guidance, and also from the added *Statistical Information for In Vitro Bioequivalence Data* material posted on August 18, 1999. The subsequent implementation will include the estimation of within canister (between life stage) component of variance.

FROM THE DRAFT JUNE 1999 GUIDANCE

IX. STATISTICAL ANALYSES

B. IN VITRO BE DATA: NONPROFILE ANALYSES USING A CONFIDENCE INTERVAL APPROACH

Nonprofile analyses should be applied to the following tests: (1) dose or spray content uniformity through container life; (2) droplet size distribution; (3) spray pattern; and (4) priming and/or repriming, when this information is specified in the labeling.

1. Study Protocol

Data for the BE criterion should be based on testing a suitable number of bottles or canisters from each of three batches of the T and R drug products. Each bottle or canister should be tested for the measure (parameter) of interest at beginning and end, or beginning, middle, and end of unit life, as indicated in section V and Table 1. Rather than evaluate performance at each life stage separately, a criterion is recommended that combines the multiple life stages. In doing so, the multiple life stages are considered as providing measures of the same underlying quantity. The recommended criterion considers deviations from uniformity across bottle or canister life stages; results are ideally uniform. Lack of uniformity between life stages should be treated as another variance component in the criterion.

For suspension formulation nasal sprays and solution formulation and suspension formulation nasal aerosols, the number of canisters or bottles (units) of product to be studied should not be fewer than 30 for each of the test and reference products (i.e., no fewer than 10 from each of three batches). For solution formulation nasal sprays, no fewer than 10 units from each of the three batches or three sublots should be studied. The number of units is a function of T to R product means and variances. Estimates of these mean differences and variances will necessitate pilot studies.

2. Criterion for Comparisons, Confidence Interval, and Bioequivalence Limit

The equivalence approach for nonprofile tests relies on (1) a criterion to allow the comparison, (2) a confidence interval for the criterion, and (3) a BE limit for the criterion.

a. Criterion for comparison

The in vitro population BE criterion and BE limit are:

$$\frac{(\mu_T - \mu_R)^2 + (\sigma_T^2 - \sigma_R^2)}{\sigma_R^2} \leq \theta$$

where:

μ_T, μ_R	=	T and R means (log scale)
σ_{BT}, σ_{BR}	=	between batch T and R standard deviations (log scale)
σ_{CT}, σ_{CR}	=	between canister T and R standard deviations (log scale)
σ^2_R	=	$\sigma^2_{BR} + \sigma^2_{CR} + \sigma^2_{LR}$
σ^2_T	=	$\sigma^2_{BT} + \sigma^2_{CT} + \sigma^2_{LT}$
σ_{LT}, σ_{LR}	=	within T and R canister between life stage standard deviation
θ	=	in vitro BE (upper) limit

The overall means for the two formulations should be averaged over all bottles or canisters, life stages (except for priming and repriming evaluations), and batches.

The general approach should be to calculate a 95 percent upper bound for the criterion. If this upper bound is less than or equal to the upper limit, θ, the test product may be judged to be bioequivalent to the reference product at the 5 percent level. The criterion will be further discussed in the guidance for industry on *In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches* (draft December 1997), when finalized. A population, rather than average, bioequivalence criterion is recommended in order to estimate whether the test product may be more variable than the reference product. The test product should be as or more consistent in the delivery of drug than is the reference product. An individual BE approach is not appropriate for in vitro data because there are no subjects, thus no subject-by-formulation interaction.

b. Determining a 95 percent upper bound

CDER recommends that a method of moments approach be used for estimating the means and variances needed to determine the population bioequivalence criterion. Approaches based on restricted maximum likelihood (REML) may be used in special cases. For determining the 95 percent upper bound, CDER recommends using a method analogous to one proposed for individual bioequivalence (Hyslop, Hsuan and Holder 1998).

c. Specification of the population BE upper limit

The general form of the upper limit, θ, is analogous to the form of the population BE criterion, which is

$$\frac{(\textit{mean difference in natural log scale})^2 + \textit{variance terms}}{\textit{comparison variance}}$$

The corresponding form for the upper limit is then

$$\frac{(\textit{average BE limit in natural log scale})^2 + \textit{variance terms offset}}{\textit{scaling variance}}$$

This formula contains three values to be specified: (1) average BE limit, (2) variance terms offset, (3) and scaling variance. These values will be specified when this guidance is finalized based on simulation work now in progress.

Average BE Limit

Due to the low variability of in vitro measurements, at the present time CDER recommends that the limit not be larger than 90/111 (i.e., the ratio of geometric means would fall within 0.90 and 1.11). A value of 0.90 is tentatively recommended as the average BE limit. This value should be used in calculating the population BE limit (refer to θ in the equation in section IX.B.2.a, above).

Variance Terms Offset

This value arises to allow some difference among the total variances that may be inconsequential. In this regard, the variance terms offset is analogous to the average BE limit. The variance terms offset also helps correct for the effect on power and sample size for the need to estimate the variances. Because of the low variability of in vitro measurements, the variance terms offset, denoted ϵ_P in the draft guidance on *In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches* (December 1997), when finalized, should be taken as 0.0. CDER is also considering ϵ_P equal to 0.01.

Scaling Variance

This value adjusts the BE criterion depending on the reference product variance. When this variance is greater than the scaling variance, σ_{T0}^2, the limit is widened. When this variance is less than the scaling variance, the limit is narrowed.

Mixed scaling should be employed for in vitro studies, as described in the draft guidance on *In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches* (December 1997), when finalized. With mixed scaling, when the reference variance **in the study** is less than the scaling variance, the population BE criterion should be modified to its *constant-scaled* form:

$$\frac{(\mu_T - \mu_R)^2 + (\sigma_T^2 - \sigma_R^2)}{\sigma_{T0}^2}$$

Mixed scaling is used to avoid penalizing test products for cases with very low reference variance. It is CDER's current intent to select σ_{T0} for in vitro studies so that most studies will use constant scaling and thus, that σ_{T0} will be at least 0.10.

The upper limit may be interpreted by reference to a population distance ratio (PDR). The PDR is the ratio of the test-reference distance (in the log scale) to the reference-reference distance. In contrast to individual BE, the distances for population BE are based on administration to separate individuals (further details will be provided in the guidance for industry on *In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches* (draft December 1997), when finalized. The population BE criterion, denoted by PBC, is related to the PDR by

$$PDR = \left(1 + \frac{PBC}{2}\right)^{\frac{1}{2}}$$

Substituting the BE limit θ for PBC expresses the upper limit in the PDR scale. The specification of 0.90 for the average limit, 0.0 for the variance offset, and 0.10 for the scaling standard deviation corresponds to an upper limit for PDR of 1.25.

STATISTICAL INFORMATION FOR IN VITRO BIOEQUIVALENCE DATA

(Originally posted August 18, 1999)

BIOAVAILABILITY AND BIOEQUIVALENCE STUDIES FOR NASAL AEROSOLS AND NASAL SPRAYS FOR LOCAL ACTION

IN VITRO NONPROFILE BIOEQUIVALENCE DATA: POPULATION BIOEQUIVALENCE - PARALLEL DESIGNS

Method for Statistical Test of Population Bioequivalence Criterion

Since three batches are not sufficient to reliably estimate the between batch component of variance, the total variances are estimated as the between canister variance of the "super-batch" consisting of the three batches combined. In addition, this initial implementation does not include the estimation of within canister (between life stage) component of variance.

Criterion:

$$((\mu_T - \mu_R)^2 + \sigma_T^2 - \sigma_R^2) / \sigma_R^2 \leq \theta_p$$

Following the method developed by Hyslop, Hsuan, and Holder (1999) for the individual bioequivalence criterion, we propose the following method for testing this criterion. The procedure involves the computation of a test statistic which is either positive (does not conclude population bioequivalence) or negative (concludes population bioequivalence). This method is based on the work of Howe (1974) and Ting et al. (1990). The method outlined below assumes equal numbers of canisters per batch, and that three batches for each product will be combined as one "super-batch" for each product for analysis.

Notation:

n_T, n_R:	Number of canisters per batch, for T and R products
ℓ_T, ℓ_R:	Number of batches of T and R products
$\Delta = \mu_T - \mu_R$:	Mean difference of T and R products
σ_T^2, σ_R^2:	Total variance of T and R products
σ_{T0}, θ_p:	Regulatory constants

Linearized Criteria:

$$\eta_1 = (\mu_T - \mu_R)^2 + (\sigma_T{}^2 - \sigma_R{}^2) - \theta_P \cdot \sigma_R^2 < 0, \quad \text{for } \sigma_R > \sigma_{T0}$$

$$\eta_2 = (\mu_T - \mu_R)^2 + (\sigma_T{}^2 - \sigma_R{}^2) - \theta_P \cdot \sigma_{T0}^2 < 0, \quad \text{for } \sigma_R \leq \sigma_{T0}$$

Estimating the Linearized Criteria:

Begin by computing the separate means and variances for the log of the measure of each product. Since three batches is not sufficient to reliably estimate the between batch component, the total variances are estimated as the between canister variance of the "super-batch" consisting of the three batches combined. Compute the total sum of squares for each product and denote them as SST_T and SST_R. Compute: $MST_T = SST_T / (n_T \cdot \ell_T - 1)$ and $MST_R = SST_R / (n_R \cdot \ell_R - 1)$ (Searle). Estimate the overall means of each product and compute:

$$\hat{\Delta} = \overline{X}_T - \overline{X}_R = \hat{\mu}_T - \mu_R.$$

$$\hat{\eta}_1 = (\overline{X}_T - \overline{X}_R)^2 + MST_T - \left(1 + \theta_p\right) MST_R$$

$$\hat{\eta}_2 = (\overline{X}_T - \overline{X}_R)^2 + MST_T - MST_R - \theta_p \sigma_{T0}{}^2$$

To test for population bioequivalence, compute the 95% upper confidence bound of either the reference-scaled or constant-scaled linearized criterion. The procedure for computing this is described in the next paragraphs. If this upper bound is negative, conclude population bioequivalence. If the upper bound is positive, do not conclude population bioequivalence.

95% Upper Confidence Bounds of Components:

Use the estimated total variance for T and for R based on $\ell_T \cdot n_T - 1$ and $\ell_R \cdot n_R - 1$ degrees of freedom where n_T and n_R are the number of canisters in each of the T and R batches and ℓ_T, ℓ_R are the number of batches of each product.

Using methods developed by Lee and Gurland for the Behrens-Fisher problem and the estimation method provided by Lee and Fineberg, compute two-sided confidence interval for $\hat{\Delta}$ based on the total variances.

Let $E0 = (\overline{X}_T - \overline{X}_R)^2$, $H0 = \max\left\{LCL^2, UCL^2\right\}$ using the two-sided interval obtained for $\hat{\Delta} = \overline{X}_T - \overline{X}_R$ which is described above (Hsu et al, 1994).

Let $E1 = MST_T$, compute $H1 = \dfrac{(\ell_T n_T - 1)E1}{\chi^2_{\ell_T n_T - 1, \alpha}}$

Let E2rs= $-\left(1+\theta_p\right) MST_R$, compute $H2rs = \dfrac{(\ell_R n_R - 1)E2rs}{\chi^2_{\ell_R n_R - 1, 1-\alpha}}$

Let E2cs= $-MST_R$, compute $H2cs = \dfrac{(\ell_R n_R - 1)E2cs}{\chi^2_{\ell_R n_R - 1, 1-\alpha}}$

For each component above, also compute $Ui = \left(Hi - Ei\right)^2$.

95% Upper Confidence Bounds for Linearized Criteria:

$$H_{\eta_1} = (E0 + E1 + E2rs) + (U0 + U1 + U2rs)^{1/2}$$

$$H_{\eta_2} = (E0 + E1 + E2cs - \theta_p \sigma_{T0}{}^2) + (U0 + U1 + U2cs)^{1/2}$$

References

Howe W.G., "Approximate confidence limits on the mean of X+Y where X and Y are two tabled independent random variables". *J Amer Statist Assoc,* 69:789-794, 1974.

Hsu J.C., J.T.G.Hwang, H.-K. Liu, and S.J. Ruberg, "Confidence intervals associated with tests for bioequivalence". *Biometrika,* 81: 103-114, 1994.

Hyslop T., F. Hsuan, D.J. Holder, "A small-sample confidence interval approach to assess individual bioequivalence," Submitted, 1999.

Lee A.F.S., J. Gurland, "Size and power of test for equality of means of two normal populations with unequal variances," *J Amer Statist Assoc,* 70, 933-941, 1975.

Lee A.F.S., N.S. Fineberg, "A fitted test for the Behrens-Fisher problem," *Comm Statist- Theory Meth*, 20, 653-666, 1991.

Searle S.R., *Linear Models*. John Wiley and Sons, New York, 1997.

Ting N., R.K. Burdick, F.A. Graybill, S. Jeyaratnam, T.-F.C.Lu, "Confidence intervals on linear combinations of variance components that are unrestricted in sign". *J Statist Computation and Simulation*, 35, 135-143, 1990.