Commentary

Individual Bioequivalence: Attractive in Principle, Difficult in Practice

Laszlo Endrenyi,1,5 Gordon L. Amidon,2 Kamal K. Midha,3 and Jerome P. Skelly4

Received February 23, 1998; accepted May 29, 1998

KEY WORDS: individual bioequivalence; average bioequivalence; drug regulation; generic drugs.

The Food and Drug Administration (FDA) has recently stated its intention to promulgate new bioequivalence requirements. It has published a Draft Guidance on the introduction of individual and population bioequivalence (1). FDA has invited comments on the draft. Public discussion has considered almost exclusively individual bioequivalence. Therefore, the present commentary will focus on aspects of individual bioequivalence proposed in the draft.

The new approach of individual bioequivalence is intended to supersede the present procedure based on the evaluation of average bioequivalence. The principles and methods of the present and proposed approaches are considerably different (2–6). Therefore it is important to investigate the properties, methodology, and computation of FDA’s proposed new approach and establish its appropriateness and plausibility before implementation. A major problem is that the results of only a few investigations are available which shed light on characteristics of the procedures suggested for the evaluation of individual bioequivalence.

The rationale, principles and procedures for the assessment of individual bioequivalence have been described repeatedly (2–7). Therefore, these will be presented only briefly. Questions about some of the properties of the suggested regulatory requirements, and their need, will then be presented.

OUTLINE OF THE RATIONALE AND PRINCIPLES FOR THE DETERMINATION OF INDIVIDUAL BIOEQUIVALENCE

The primary reason for introducing the approach of individual bioequivalence was that it was suggested to deal with the issue of switchability (or interchangeability) of drug formulations in patients (2,3,5–9). The issue arises when patients are stabilized on one formulation and are then switched to another.

By contrast, prescribability of a drug was said to concern patients who have not yet taken the drug in any of its approved marketed formulations. It has customarily relied on average properties of the drug products.

The new approach would be able to assess whether the responses in various subjects would be similar or changed following the substitution of one formulation by a different formulation. Formally, this is equivalent to the evaluation of the subject-by-formulation interaction.

Another feature of the suggested methodology is that it enables the estimation not only of the two means of a given metric (such as log AUC), but also their inter- and intraindividual variances.

If the new formulation is “better” than the previous reference product in the sense that the test formulation has the smaller variation, then a premium or “reward” is provided in that a wider difference between average kinetic parameters could still allow the test formulation to meet the regulatory criteria. This suggestion is a new departure in the assessment of bioequivalence.

The narrowing or widening of bioequivalence limits is also proposed for drugs exhibiting either a narrow therapeutic index or large intraindividual variation, respectively. The adjustment of bioequivalence limits is accomplished by standardizing (scaling) the relevant regulatory criterion by the intrasubject variance of the reference formulation.

OUTLINE OF THE PROCEDURE PROPOSED FOR THE EVALUATION OF INDIVIDUAL BIOEQUIVALENCE

FDA has suggested (1) a model for the evaluation of individual bioequivalence which had been originally described by Schall and Luus (10). The model has three components:

\[(\text{Difference of means})^2 + \text{Interaction} + \text{Difference of variances} \leq \theta_u^2\]

The first component is the squared difference between the means of the two formulations. The second term is the variance component for the so-called subject-by-formulation interaction; it measures quantitatively the similarity or dissimilarity of the
kinetic responses observed in various subjects when they switch from one drug product to another.

The third component is the difference between intrindividual variances which are recorded following the replicate administrations of the test and reference formulations, respectively.

The unscaled regulatory criterion for individual bioequivalence requires that the sum of the three terms should not exceed a preset limit, \( \theta_0 \). In the Draft Guidance, both sides of the above expression are divided by \( \sigma_{WO}^2 \), where \( \sigma_{WO}^2 \) is a constant, fixed within-subject variance, the value of which is set at 0.20. Thereby a constant-scaled regulatory criterion is obtained which applies a regulatory limit of

\[
0_i = \theta_i / \sigma_{WO}^2
\]

In a corresponding, reference-scaled criterion, the three terms of the unscaled model are standardized by the intrasubject variance of the reference formulation:

\[
\frac{[(\text{Difference of means})^2 + \text{Interaction} + \text{Difference of variances}]}{\text{(Reference intrasubject variance)}} = \theta_i^2
\]

It is expected that the terms in the reference scaled model should not exceed the preset regulatory limit (\( \theta_i^2 \)).

In the mixed strategy originally suggested by Schall and Williams (5) and recommended by the Draft Guidance (1), the reference-scaled criterion is applied when \( \sigma_{WR}^2 > \sigma_{WO}^2 \), and the constant-scaled criterion is implemented when \( \sigma_{WR}^2 \leq \sigma_{WO}^2 \).

The suggested unscaled (or constant-scaled) criterion has some notable properties.

First, if the two formulations have the same intraindividual variances and there is no subject-by-formulation interaction then the criterion reduces to that of average bioequivalence and can be written in the form:

\[
-\theta_a \leq \text{Difference of means} \leq \theta_a
\]

Here the customary limit for log AUC is \( \theta_a = \log 1.25 \). In the context of individual bioequivalence, this corresponds to \( \theta_i = 1.25 \). In the Draft Guidance, this value is substantially enhanced, to a range between \( \theta_i = 2.25 \) and 2.50, for the evaluation of individual bioequivalence.

Second, the expressions for individual and average bioequivalence imply, in practice, that estimated values of the terms in the regulatory criteria, together with their confidence limits, should not exceed the bioequivalence limits.

Third, there is a tradeoff between the difference of intraindividual variances of the two formulations, and the difference between the means of the two products (11). If the intrasubject variation of the test formulation is smaller than that of the reference product, then the third term in the regulatory criterion, the difference of variances, is negative. With a fixed interaction, the difference of means in the first term can then expand for the declaration of bioequivalence. This was considered to be an attractive feature of the recommended approach since it rewards a better product which exhibits a smaller variation (5–7,11).

The second and third properties apply also to the reference-scaled analysis of bioequivalence.

INDIVIDUAL BIOEQUIVALENCE: SCIENTIFIC FLAWS OF THE PROPOSED PROCEDURES

Only two papers evaluating the properties of the suggested procedure have been published to date. To our knowledge, two additional investigations which are in various stages of the publication process and one Master’s thesis are relevant to this subject.

Results of each of the five studies raise significant questions about the characteristics of the proposed methodology and place in doubt the current appropriateness of adopting the Draft Guidance. Two scientific issues will be summarized. Both problems can be extracted from at least two of these five investigations.

1. The numerical tradeoff of the deviations between the intraindividual variances and the means of the two formulations is strongly asymmetric. Notably, changes in the difference between the means of the two formulations are very sensitive to changes in the difference between the intraindividual variances. This can be deduced from the study of Hauck et al. (11), notably from its Figures 2, 3, and 4. For example, when the coefficient of variation of the reference formulation is 40% (\( CV_{WR} = 0.40 \)), with a 5% deviation the coefficient of variation of the test product would be either 38 or 42% (\( CV_{WT} = 0.38 \) or 0.42). When \( CV_{WT} = 0.38 \), the estimated variation of the test formulation is 5% lower than that of the reference product and the allowable difference between the means, compatible with the declaration of bioequivalence, can expand by 12%. This sizeable benefit, or “reward”, is provided in recognition of the apparent improvement shown by the test formulation. On the other hand, when the estimated variation of the test product is 5% higher than that of the reference formulation (\( CV_{WT} = 0.42 \)), the difference between the two means must be contracted by 11% in order to declare bioequivalence (12). This implies the imposition of a “penalty” which arises when the test formulation is apparently “poorer” than the reference product.

As a result of the high sensitivity of the allowable difference between means, large rewards or penalties can arise for its estimated value, by random chance, with very high probabilities. For example, with a 40% intrasubject variation of the reference formulation (\( CV_{WR} = 0.40 \)), either an expansion or contraction by 10% or more of the allowable difference between means can occur with a probability of 84% (12). About half of this probability is allocated to “reward” and the other half to “penalty”.

Based on the computational method of Schall and Williams (5), Midha et al. (13) observed similarly strong sensitivity of the allowable difference between means to small changes in the estimated intrasubject variations. These authors recorded also sensitivity to period effects.

The high sensitivity and the associated large probabilities are unreasonable. Rewarding an improved formulation which exhibits reduced intrasubject variation, is appealing in principle. However, the rewards and penalties which can arise from the model by random chance, are quantitatively excessive. The characteristics of the mean-variability tradeoff are unattractive even if the range for the change of the allowable difference between means were truncated.

2. The scaled criterion of bioequivalence declares the equivalence of two formulations very liberally. Midha et al. (13) demonstrated that two formulations could be judged as