*Pharmaceutical Research, Vol. 15, No. 9, 1998*

*Commentary*

# Individual Bioequivalence: Attractive in Principle, Difficult in Practice

**Laszlo Endrenyi,[1,5] Gordon L. Amidon,[2] Kamal K. Midha,[3] and Jerome P. Skelly[4]**

*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 It has published a Draft Guidance on the introduction of individual and population bioequivalence (1).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

[1] University of Toronto, Department of Pharmacology, Toronto, Ontario M5S 1A8, Canada.
[2] University of Michigan.
[3] University of Saskatchewan.
[4] University of Cincinnati.
[5] To whom correspondence should be addressed. (e-mail: l.endrenyi@utoronto.ca)



0724-8741/98/0900-1321$15.00/0 

kinetic responses observed in various subjects when they switch from one drug product to another.

The third component is the difference between intraindividual 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, $\theta_u^2$. 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 $0.20^2$. Thereby a *constant-scaled regulatory criterion* is obtained which applies a regulatory limit of

$$\theta_l = \theta_u/\sigma_{WO}$$

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

$$[(\text{Difference of means})^2 + \text{Interaction} + \text{Difference of variances}] / (\text{Reference intrasubject variance}) \leq \theta_l^2$$

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

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:

$$-\theta_u \leq \text{Difference of means} \leq \theta_u$$

Here the customary limit for log AUC is $\theta_u = \log 1.25$. In the context of individual bioequivalence, this corresponds to $\theta_l = 1.25$. In the Draft Guidance, this value is substantially enhanced, to a range between $\theta_l = 2.25$ 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 the 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

bioequivalent even when the difference between their means exceeded 25%. Liberal judgements by the scaled criterion also can be deduced from the results of two additional studies (14,15). For example, Figure 1 presents rearranged results of Kimanani and Potvin (14); these were obtained from simulations of 4-period crossover trials. Figure 1 illustrates the percentage of trials in which bioequivalence would be accepted. A high percentage of stated bioequivalence can be observed, particularly with large intraindividual variations, when the scaled criterion is assumed to be useful.

The scaled criterion of bioequivalence is also very insensitive to deviations between the means of the two formulations, i.e. to passing or failing the criterion of average bioequivalence, especially when the intrasubject variation is large (14,15). This is illustrated in Figure 1 by a rearrangement of Kimanani and Potvin's results (14).

## UNCERTAINTIES OF COMPUTATIONS

The models for individual bioequivalence are substantially more complicated and have more parameters than those applied for average bioequivalence. It is therefore expected that the estimated parameters of these models have comparatively large uncertainties and undesirable correlations under certain conditions. As a further consequence, the estimated probability of accepting a given bioequivalence study, such as the results illustrated in Figure 1, is also more uncertain. In other words, with a given bioequivalence study, it could be to an extent a matter of chance whether the results satisfy the regulatory criteria. These effects require extensive further investigations.

Additional uncertainties could arise from the bootstrap computations which are to be applied for the evaluation of the upper one-sided confidence limit (1). 2,000 computations give some assurance of reliability. Nevertheless, the result is different in each set of calculations. This could conceivably give rise to manipulations in borderline cases.

## IS THERE A PROBLEM, IS THERE A NEED?

As observed earlier, the proposal for the determination of individual bioequivalence has an interesting rationale. Still, it is not clear that the approach would be reasonable in practice, or that its acceptance, and especially its speedy adoption, would be justified. Therefore, pertinent questions need to be asked.

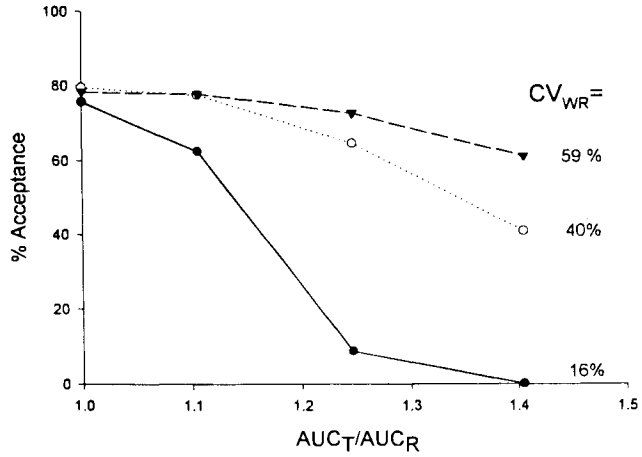


**Fig. 1.** Percentage of simulated 4-period crossover trials which accept individual bioequivalence by the reference-scaled criterion at various true ratios of the AUCs of the two formulations and with different $CV_{WR}$ values. Rearranged from the results of Kimanani and Potvin (14), with the permission of the authors and the publisher.

1. Has average bioequivalence failed, i.e. have there been documented problems observed following the substitution of one formulation by another based on the acceptance of average bioequivalence? Anecdotal reports have suggested that some patients have detected changes in response when they were switched from one formulation to another even though the products have met the acceptance criteria for average bioequivalence. Such reports may involve switching between the innovator product and a generic, or between two generics each being average bioequivalent with the same reference formulation, or between two or more divergent lots of the innovator due to changes in, e.g., colour, inert ingredients or shape. Unfortunately, no objective, adequately documented reports on this subject have been published. This lack of documentation was expressed by the FDA Advisory Committee on Pharmaceutical Sciences in August, 1996 which did not perceive a justification for the introduction of individual bioequivalence. Moreover, there is no evidence that the new approach would solve the problem if it actually exists.

2. Is there evidence that subject-by-formulation interactions are important? FDA has studied and presented the results of 34 data sets (drugs and metabolites) from twelve 4-period crossover investigations. The results were interpreted to demonstrate that about one-third of the data sets exhibited a defined level of substantial interaction. It was noted, however, during the discussion at the recent AAPS annual meeting that only a small number of replicate design studies had been available for exploratory investigations and that none of them was designed with individual bioequivalence in mind. In addition, technical questions were raised identifying possible problems with their analysis. It was understood that the data would be reanalysed by an external panel. Very recently (in March, 1998), the data set has been made available on the Internet. At present, however, the data do not appear to provide convincing evidence on the prevalence of subject-by-formulation interactions to support the introduction of individual bioequivalence. Furthermore, we do not know that a possible interaction would have clinical consequences of either overdosing or underdosing patients, i.e. resulting either in toxicity or lack of efficacy. Underlying these questions is the need for a mechanistic analysis of the basis for the interaction.

Are the suggested comparisons *clinically* relevant? Two subquestions can be considered:

3. What populations, e.g. healthy volunteers or patients, are appropriate for the assessment of individual bioequivalence and, in particular, for the study of the subject-by-formulation interaction? The draft Guidance is permissive and recommends that "restrictions to entry to the study should be based solely on safety considerations". Levy (16,17) argued that bioequivalence should be assessed under clinically relevant conditions. This stand is supported by a study of Carter et al. (18) which observed a verapamil product to be bioequivalent with two others in healthy young subjects but bioinequivalent in elderly hypertensive patients. Levy (16,17) suggested that trials be conducted in subjects who are as similar as possible to the population of patients who are targets for the drug, and that the studies should be performed in the steady state when this is relevant based on the therapeutic use of the drug. The contrast between out-

comes of investigations performed in target populations and in young healthy volunteers is likely to be more conspicuous in studies of individual bioequivalence than those of average bioequivalence. This is expected because intraindividual variations and especially subject-by-formulation interactions are probably often meaningfully larger in a target population of patients than among the fairly homogeneous volunteers.

4. Are comparisons of and with intrasubject, intraformulation variations relevant? The suggested procedures utilize duplicate observations within given batches of the two formulations. But switching a patient from one drug product to another involves a change of formulations as well as batches. Therefore the relevant basis for the individual comparison of the test and reference products should actually be the batch-to-batch variations observed within each of the two formulations in each subject. In the presence of intrabatch variation within subjects, the intraindividual variance terms and, probably, the subject-by-formulation interaction are underestimated. As a result, estimation by the constant-scaled criterion could be too liberal, whereas the reference-scaled criterion would be generally too constraining. Midha et al. (19) demonstrated the relevance of batch-to-batch variation in the context of determining average bioequivalence. The interbatch variations take even greater importance for the assessment of individual bioequivalence.

## NEED TO KNOW, NICE TO KNOW?

It has been stressed that the determination of the intraindividual variations of the two formulations and of the subject-by-formulation interaction is important in investigations of individual bioequivalence (5–7). However, this is unproven. Moreover, it is doubtful that this information would be useful for evaluating the equivalence of formulations of *all* drugs. For example, the residual error in the analysis of variance for a typical two-period crossover trial combines contributions from the intrasubject variations and the subject-by-formulation interaction. But if the residual error is small, say less than the equivalent of a coefficient of variation of 15%, then the corresponding interaction will also be lower than 15% and would not give rise to concern for the assessment of individual bioequivalence. Consequently, the evaluation of average bioequivalence based on two-period trials should be sufficient in these cases. Therefore, separate information on the interaction and intraindividual variations is "nice to know" but not a necessity.

However, it has been suggested in discussions of public sessions that the conduct of 4- (or 3-) period crossover trials should be encouraged, and perhaps expected, even if the regulatory requirements for individual bioequivalence would not be implemented. This would still enable the estimation of intrasubject variances and subject-by-formulation interaction, and would permit the accumulation of information on various drugs and drug products. But is this information needed for the regulatory approval of drug products or is it "nice to know"?

## PROCEDURE FOR THE DEVELOPMENT OF THE DRAFT GUIDANCE

Although the concept of individual bioequivalence has been discussed for some time, the framework of the procedure has become available in various public sessions only in the past 1 1/2 years. The plans for the Guidance were presented to the FDA Advisory Committee on Pharmaceutical Sciences on August 15, 1996. The views of the Committee were summarized by its chairman, Dr. J. Swarbrick:

> A sense and lack of enthusiasm which is disturbing but is reflective of an element based on a lack of significance of the problem. In theory, the concept makes sense and has very attractive aspects. On the other hand, there is concern that we are throwing out the baby with the bathwater, the way we are assessing bioequivalence right now. There is almost universal agreement on the need for a trial period with some sort of scheduled series of studies that would generate some data that would answer some of these questions and would encourage companies to use replicate studies.

Individual bioequivalence and the proposed Guidance were discussed extensively at full sessions of the annual meetings of the Drug Information Association in June, 1997 and of AAPS in November, 1997. On both occasions, all speakers and discussants from the floor, with the exception of representatives of FDA, questioned the need for the new approach and its appropriateness as well as its usefulness.

The widely held views of the pharmaceutical community need to be seriously considered in developing regulatory policy as has been the case for other regulatory guidances. It is gratifying that at a recent AAPS/FDA Workshop on Scientific and Regulatory Issues in Product Quality (Arlington, VA, March 16–18, 1998), issues involving individual bioequivalence were discussed extensively and constructively, and the timetable for its consideration was extended.

## CONCLUSIONS

The rationale and principles of the procedure proposed for the adoption of individual bioequivalence are attractive and merit further scientific discussion. However, there has been insufficient time to evaluate carefully the properties of this completely new approach (20–22). Moreover, results of the few studies conducted to date suggest serious questions about the practical plausibility of these characteristics. Consequently, the proposed procedure should not be implemented in its present form without further research and analysis to establish its validity and utility.

## ACKNOWLEDGMENTS

We are grateful to Drs. John W. Hubbard (University of Saskatchewan) and Iain J. McGilveray (McGilveray Pharmacon Inc.) for their very helpful and insightful comments and suggestions on earlier versions of this paper.

## REFERENCES

1. FDA. In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches—Draft Guidance for Industry. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Rockville, MD, October, 1997.
2. S. Anderson. Individual bioequivalence: a problem of switchability (with discussion). *Biopharm. Rep.* **2**:1–11 (1993).
3. W. W. Hauck and S. Anderson. Measuring switchability and prescribability: when is average bioequivalence sufficient? *J. Pharmacokin. Biopharm.* **22**:551–564 (1994).
4. R. Schall. A unified view of individual, population, and average bioequivalence. In Blume, H. H. and Midha, K. K. (Eds.), Bio-International 2: Bioavailability, Bioequivalence and Pharmacokinetic Studies. Medpharm, Stuttgart, pp. 91–106 (1995).

5. R. Schall and R. L. Williams. Towards a practical strategy for assessing individual bioequivalence. *J. Pharmacokin. Biopharm.* **24**:133–149 (1996).
6. R. N. Patnaik, L. J. Lesko, M.-L. Chen, and R. L. Williams. Individual bioequivalence—new concepts in the statistical assessment of bioequivalence metrics. *Clin. Pharmacokin.* **33**:1–6 (1997).
7. M. L. Chen. Individual bioequivalence—A regulatory update. *J. Biopharm. Stat.* **7**:5–11 (1997).
8. S. Anderson and W. W. Hauck. Considerations of individual bioequivalence. *J. Pharmacokin. Biopharm.* **18**:259–273 (1990).
9. L. B. Sheiner. Bioequivalence revisited. *Stat. Med.* **11**:1777–1788 (1992).
10. R. Schall and H. E. Luus. On population and individual bioequivalence. *Stat. Med.* **12**:1109–1124 (1993).
11. W. W. Hauck, M.-L. Chen, T. Hyslop, R. Patnaik, D. Schuirmann, and R. Williams. Mean difference vs. variability reduction: tradeoffs in aggregate measures for individual bioequivalence. *Int. J. Clin. Pharmacol. Ther.* **34**:535–541 (1996).
12. L. Endrenyi and Y. Hao. Asymmetry of the mean-variability tradeoff raises questions about the model in investigations of individual bioequivalence. *Int. J. Clin. Pharmacol. Ther., accepted* for publication.
13. K. K. Midha, M. J. Rawson, and J. W. Hubbard. Individual and average bioequivalence of highly variable drugs and drug products. *J. Pharm. Sci.* **86**:1193–1197 (1997).
14. E. K. Kimanani and D. Potvin. A parametric confidence interval for a moment-based scaled criterion for individual bioequivalence. *J. Pharmacokin. Biopharm.*, in press.
15. R. M. Goosen. A Simulation Study to Determine Sample Sizes for the Assessment of Individual Bioequivalence. M.M.Sc. thesis, University of Orange Free State, Bloemfontein, South Africa, pp. 117 (1996).
16. G. Levy. The clay feet of bioequivalence. *J. Pharm. Pharmacol.* **47**:975–977 (1995).
17. G. Levy. Bioequivalence assessment: irrationality within a rational science. In K. K. Midha and T. Nagai (eds.), Bioavailability, Bioequivalence and Pharmacokinetic Studies, F.I.P. Bio-International'96, Academic Societies of Japan, Tokyo, 1996, pp. 29–32.
18. B. L. Carter, M. A. Noyes, and R. W. Demmier. Differences in serum concentrations of and responses to generic verapamil in the elderly. *Pharmacotherapy* **13**:359–368 (1993).
19. K. K. Midha, B. S. Chakraborty, R. Schwede, E. M. Hawes, J. W. Hubbard, and G. McKay. Comparative bioavailability of a new commercial tablet formulation and two lots of a reference formulation of haloperidol. *J. Pharm. Sci.* **78**:443–44 (1989).
20. L. Endrenyi. Some issues for the consideration of individual bioequivalence. *J. Biopharm. Stat.* **7**:35–39 (1997).
21. A. L. Gould. Discussion of individual bioequivalence by M.-L. Chen. *J. Biopharm. Stat.* **7**:23–29 (1997).
22. L. Endrenyi and K. K. Midha. Individual bioequivalence—has its time come? *Eur. J. Pharm. Sci.*, in press.