Subject-by-Formulation Interaction in Determinations of Individual Bioequivalence: Bias and Prevalence

Laszlo Endrenyi[1,3] and Laszlo Tothfalusi[2]

Received August 27, 1998; accepted November 11, 1998

Purpose. 1. To determine properties of the estimated variance component for the subject-by-formulation interaction (σ_D^2) in investigations of individual bioequivalence (IBE), and 2. to evaluate the prevalence of interactions in replicate-design studies published by FDA.
Methods. Four-period crossover studies evaluating IBE were simulated repeatedly. Generally, the true bioequivalence of the two formulations, including $\sigma_D^2 = 0$, was assumed. σ_D^2 was then estimated in a linear mixed-effect model by restricted maximum likelihood (REML). The same method was applied for estimating σ_D^2 for the data sets of FDA.
Results. 1. σ_D estimated by REML was positively biased. The bias and dispersion of the estimated σ_D increased approximately linearly with the estimated within-subject standard deviation for the reference formulation (σ_{WR}). Only a small proportion of the estimated σ_D exceeded the estimated σ_{WR}. 2. Distributions of the estimated σ_D were evaluated. At $\sigma_{WR} = 0.30$, a level of estimated $\sigma_D = 0.15$ was exceeded, by random chance, with a probability of about 25%. 3. Importantly, the behaviour of the σ_D^2 values estimated from the FDA data sets was similar to that exhibited by the simulated estimates of σ_D^2 which were generated under the conditions of true bioequivalence.
Conclusions. 1. σ_D estimated by REML is biased; the bias increases proportionately with the estimated σ_{WR}. Consequently, exceeding a fixed level of σ_D (e.g., 0.15) does not indicate substantial interaction. 2. The data sets of FDA are compatible with the hypothesis of $\sigma_D^2 = 0$. Consequently, they do not demonstrate the prevalence of subject-by-formulation interaction. Therefore, it could be sufficient and reasonable to evaluate bioequivalence from 2-period crossover studies.

KEY WORDS: individual bioequivalence; regulatory criterion; intrasubject variation; subject-by-formulation interaction; crossover design; maximum likelihood.

INTRODUCTION

The study of Anderson and Hauck (1) initiated great interest in the development of procedures for the evaluation of individual bioequivalence. The rationale and principles were seen to be attractive (2–4). The Food and Drug Administration (FDA) has demonstrated substantial interest and involvement in the approach. It established a strong Working Group in order to develop a methodology for the determination of individual bioequivalence. As a fruit of its studies and discussions, an interim Draft Guidance was issued (5) for the consideration of and discussion by the pharmaceutical community.

One of the apparent motivations for the promotion of the new approach was a data set of several studies in the files of FDA which was thought to suggest substantial prevalence of subject-by-formulation interactions (4–6). The concern was expressed that various patients might react differently when switched from one formulation to another. The data of FDA were obtained in 3- and 4-period crossover studies which permitted the evaluation of the interaction as well as the inter-and intraindividual variances of the two formulations.

FDA has recently made the data set available to the public on the Internet (6). Preliminary analyses were discussed at an AAPS/FDA workshop held in Crystal City, MD in March, 1998. Some participants noted that the data appeared to suggest a possible relationship between the estimated variance component for the interaction and the estimated intraindividual variabilities.

The present study will explore this relationship. Two approaches will be investigated. First, the FDA data will be analysed. Second, variance terms in the proposed regulatory model will be estimated in computer-simulated bioequivalence trials. Comparison of the results from the two approaches will lead to conclusions about the probable prevalence of the subject-by-formulation interaction.

METHODS

Regulatory Criterion

The model of Schall and Luus (7), adopted by the FDA Draft Guidance (5), expects that the following criterion will be satisfied for the assessment of individual bioequivalence:

$$[(\mu_T - \mu_R)^2 + \sigma_D^2 + (\sigma_{WT}^2 - \sigma_{WR}^2)]/\sigma_W^2 \leq \theta_I \qquad (1)$$

Here μ_T and μ_R are the mean kinetic responses (e.g., log AUC) of the test (T) and reference (R) formulations, respectively. σ_D^2 is the variance component for the subject-by-formulation interaction. σ_{WT}^2 and σ_{WR}^2 are intraindividual variances of the two drug products. σ_W^2 takes a constant value of σ_{WO}^2 when $\sigma_{WR}^2 \leq \sigma_{WO}^2$ ("constant-scaled criterion"); alternatively, $\sigma_W^2 = \sigma_{WR}^2$ when $\sigma_{WR}^2 > \sigma_{WO}^2$ ("reference-scaled criterion"). $\sigma_{WO} = 0.20$ was suggested as the value separating the application of the two regulatory criteria (5). θ_I is the preset bioequivalence limit.

The proposed regulatory requirement is based on Eq. 1. According to the suggested approach, first, the parameters μ_T, μ_R, σ_D^2, σ_{WT}^2 and σ_{WR}^2 should be estimated and substituted into Eq. 1 (5). Following this, upper 95% confidence limits of the criterion model are calculated by bootstrap computations (8). Individual bioequivalence of the two formulations may be claimed if the estimated confidence limit does not exceed the preset value of θ_I.

The interim Draft Guidance (5) suggests that the interaction variance component (σ_D^2) be obtained from:

$$\sigma_D^2 = (\sigma_{BT} - \sigma_{BR})^2 + 2(1 - \rho)\sigma_{BT}\sigma_{BR}$$

where σ_{BT}^2 and σ_{BR}^2 are the between-subject variance components for the test and reference formulations, respectively, and ρ is the correlation coefficient between the means within each subject for the two drug products. Consequently, it is necessary to estimate also σ_{BT}^2, σ_{BR}^2 and ρ.

A substantial value of σ_D could indicate significant interaction and give rise to the concern that various patients could

[1] University of Toronto, Department of Pharmacology, Toronto, Ontario M5S 1A8, Canada.
[2] Semmelweis Medical University, Department of Pharmacodynamics, Budapest, Hungary.
[3] To whom correspondence should be addressed. (e-mail: l.endrenyi@utoronto.ca)

0724-8741/99/0200-0186$16.00/0

respond differently to the substitution of the reference formulation by the test product even when, on average, the two products yield the same drug levels. $\sigma_D = 0.15$ was proposed as a limit beyond which interaction could be considered to be important (5,6). It was observed that, in a data set of several 3- and 4-period bioequivalence trials in an FDA depository, $\sigma_D > 0.15$ was calculated in a substantial proportion of the studies (4,6).

Simulation of Trials and Estimation of Parameters

4-period crossover trials were simulated. They assumed that the two formulations were in fact completely bioequivalent, i.e. that $\mu_T = \mu_R$, $\sigma^2_{BT} = \sigma^2_{BR}$, $\sigma^2_{WT} = \sigma^2_{WR}$, and especially $\sigma^2_D = 0$. 24 subjects were assumed in each trial.

200 simulations were used to illustrate the relationship between estimated values of the variance components (or their square roots) for the subject-by-formulation interaction (s^2_D) and the estimated intraindividual variance (or its square root) for the reference formulation (s^2_{WR}). Various true σ_{WR} values were assumed by generating them uniformly between 0.02 and 0.50. In the studies to be illustrated, identical magnitudes of the between- and within-subject variance components were assumed: $\sigma^2_{BT} = \sigma^2_{BR} = \sigma^2_{WT} = \sigma^2_{WR}$. However, additional investigations considered larger values for the between- than for the within-subject variabilities. The (logarithmic) observations were assumed to follow normal distributions with identical means for the two formulations: $\mu_T = \mu_R$.

Parameters of the regulatory model were estimated, within a linear mixed-effect model, by restricted maximum likelihood (REML) from the results of each study in the FDA data set as well as in each simulated trial. The SAS Procedure MIXED (version 6.12, SAS Inc., Cary, NC) was applied for this purpose. As suggested in the FDA Technical Summary (9), the repeated measurements of formulations within subjects were considered to be random effects. The contrasts of formulations, periods, sequences, and a period-by-sequence interaction (within formulations) were fixed effects. The last of these, the interaction, was not included in analyses of 3-period investigations. The heterogeneous compound symmetry (csh) option assured that the variance-covariance matrix (G) for the random-effect parameters would be positive definite.

Among the data published by FDA (6), the results for ln AUC_T and ln C_{max} were analysed. Individual kinetic parameters of 34 analyses for 12 drugs were available and evaluated. However, in the data set No. 4, the extrapolated ln AUC_{inf} was evaluated since AUC_T values were not available. In the data set 16b, negative variance estimate was obtained for ln AUC_T, possibly due to overparameterization, and therefore only the result for ln C_{max} is presented.

RESULTS

Figure 1A illustrates the relationship between the square-rooted estimates of σ^2_D and σ^2_{WR} obtained in simulated bioequivalence trials. The square-rooted estimates of σ^2_D generally increase with rising square-rooted estimates of σ^2_{WR}; the relationship between the two quantities is approximately linear. The diagram includes a line drawn though the origin with a slope of 1.0. Most points are below this line but a few exceed it. The behaviour of the data demonstrates the bias of σ^2_D when it is estimated by REML; the estimated σ^2_D is biased since its values, on the average, deviate from the assumed true magnitude of zero. The bias increases along with the intrasubject variance and its estimate.

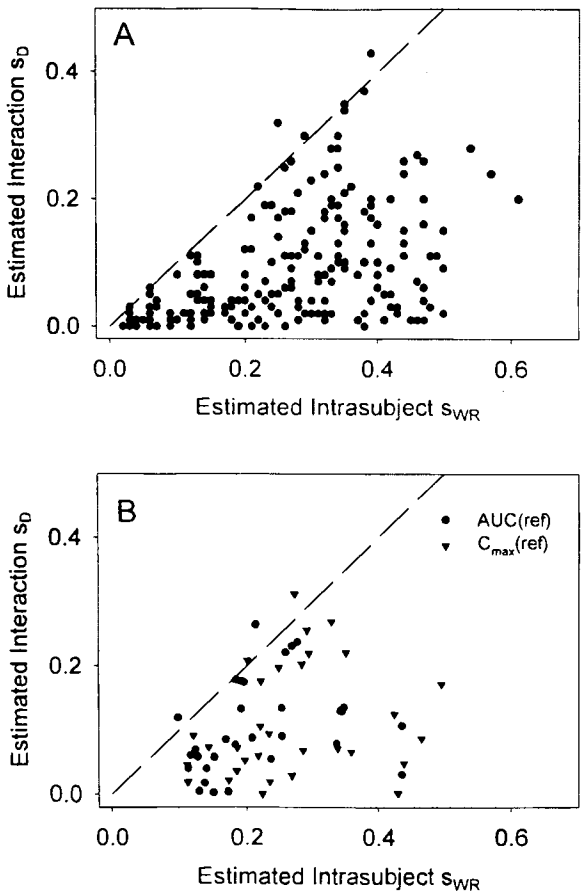

Fig. 1. Relationship between estimated, square-rooted variance components for subject-by-formulation interaction (s_D) and within-subject variability for the reference formulation (s_{WR}). The equality between the two estimates, $s_D = s_{WR}$, is indicated by a dashed line. (A) 200 simulated data. (B) Data set published by FDA. Values estimated for AUC are marked by circles, those for C_{max} by triangles.

Figure 1B presents the same contrast from results of analyses of the data set published by FDA. The simulated and observed data follow very similar patterns.

This impression is confirmed when both simulated and observed data are replotted in a more sensitive contrast of s_D/s_{WR} against s_{WR} (Fig. 2). The simulated data show essentially a horizontal pattern (Fig. 2A). Most ratios are below the level of 1.0 but a few exceed this value. The studies in the FDA data set exhibit a similar pattern (Fig. 2B).

The characteristics of the estimated σ_D are illuminated further by the results of 200 simulations which were performed at each of the fixed levels of $\sigma_{WR} = 0.10$, 0.30, and 0.50. Boxplots of s_D are shown (Fig. 3A). They confirm the rising trend of s_D with increasing σ_{WR}. The contrast of the estimates s_D and s_{WR} indicates that the dispersion of both quantities increases with their rising average levels (Fig. 3B). The diagram illustrates also how frequently the level of $s_D = 0.15$ can be exceeded, by chance, under various conditions.

These impressions are confirmed by the results of 3,000 additional simulations at each of $\sigma_{WR} = 0.1, 0.2, \ldots, 0.5$ which

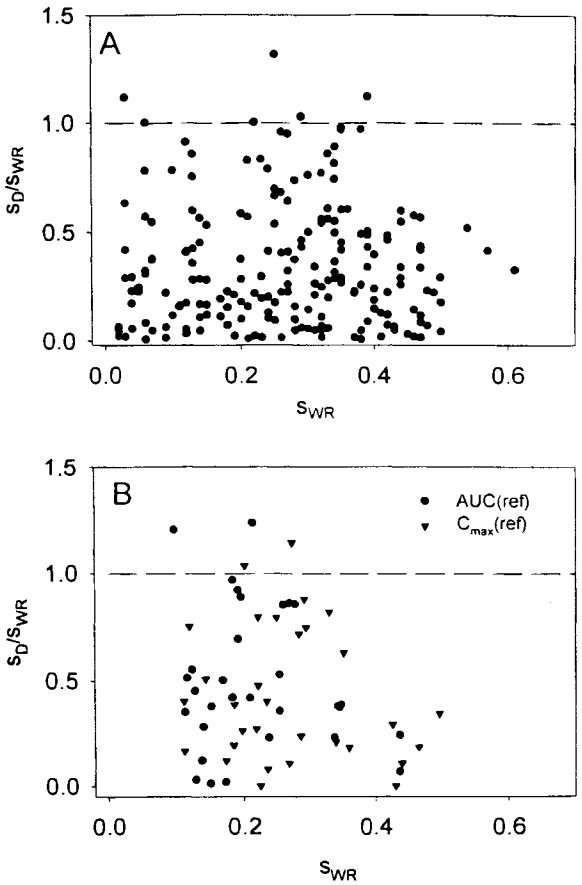

Fig. 2. Relationship between the ratio of estimates s_D/s_{WR} and the estimated s_{WR}. The equality between the two estimates, $s_D = s_{WR}$, is indicated by a horizontal dashed line. (A) 200 simulated data. (B) Data set published by FDA. Values estimated for AUC are marked by circles, those for C_{max} by triangles.

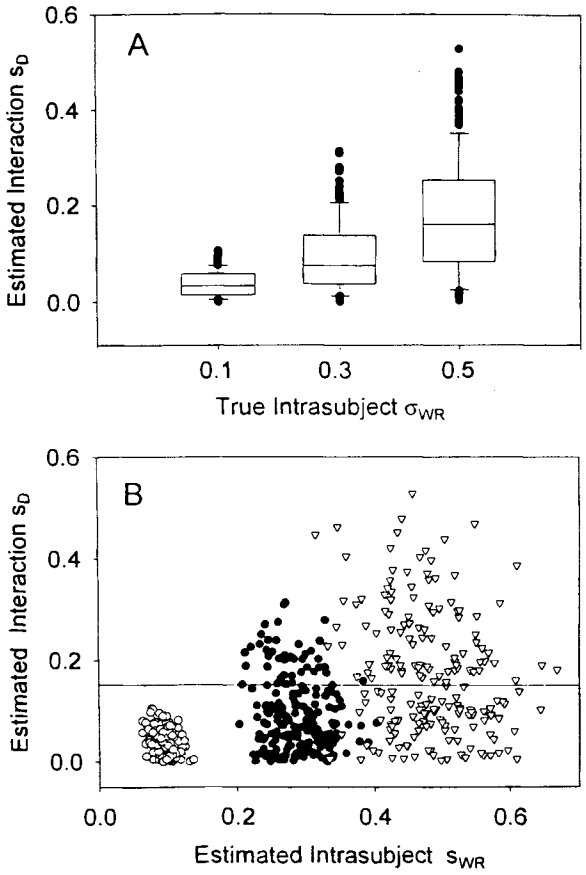

Fig. 3. Estimated square-rooted variance components for the subject-by-formulation interaction (s_D) which were generated at three levels of intrasubject variability of the reference formulation, $\sigma_{WR} = 0.10$, 0.30, and 0.50. (A) Contrasts with true values of σ_{WR}. Boxplots containing the 10, 25, 50 (the median), 75 and 90th percentiles of the distributions are shown. (B) Contrasts with the estimated values, s_{WR}. The level of $s_D = 0.15$ is shown by a horizontal line.

are presented in Table 1. It lists the average, standard deviation, and the 5, 25, 50 (the median), 75 and 95% cutoffs of s_D estimated at the various levels of the intrasubject variabilities. Each illustrated statistic increases with rising values of σ_{WR}. It is interesting to observe the frequency of estimated s_D values exceeding a given level. For example, there is about 5% probability that $s_D = 0.15$ is exceeded, by random chance, when $\sigma_{WR} = 0.20$. The probability rises to about 25% when $\sigma_{WR} = 0.30$, and to over 50% when $\sigma_{WR} = 0.50$.

Figure 4 illustrates that the average and standard deviation change approximately linearly with σ_{WR} in the investigated range. Similar behaviour was observed for the corresponding nonparametric statistics such as the median and interquartile range (the difference of the 75th and 25th percentiles) as well as for other quantiles, including that of 95%.

DISCUSSION

One of the main results of the present work is that the variance component for the subject-by-formulation interaction (σ_D^2) is positively biased when it is estimated by restricted maximum likelihood (REML). This method is suggested by the FDA Draft Guidance and its technical summary (5,9). The parameters (including the variance component σ_D^2) of the regulatory model can be estimated e.g. by applying the SAS MIXED procedure.

The positive bias of the estimated σ_D^2 is, in retrospect, not surprising. All variance components, estimated by the suggested REML procedure, are constrained to have positive values. Consequently, the average of the estimated σ_D^2 values is also positive

Table 1. Descriptive Statistics for the Distribution of the Estimated s_D

σ_{WR}	0.1	0.2	0.3	0.4	0.5
Average	0.0341	0.0683	0.1028	0.137	0.166
St. deviation	0.0247	0.0489	0.0730	0.101	0.120
Percentiles					
5%	0.0029	0.0054	0.0091	0.0102	0.0146
25%	0.0138	0.0281	0.0422	0.0534	0.0678
50%	0.0291	0.0595	0.0884	0.116	0.142
75%	0.0509	0.101	0.155	0.208	0.248
95%	0.0804	0.160	0.238	0.326	0.389

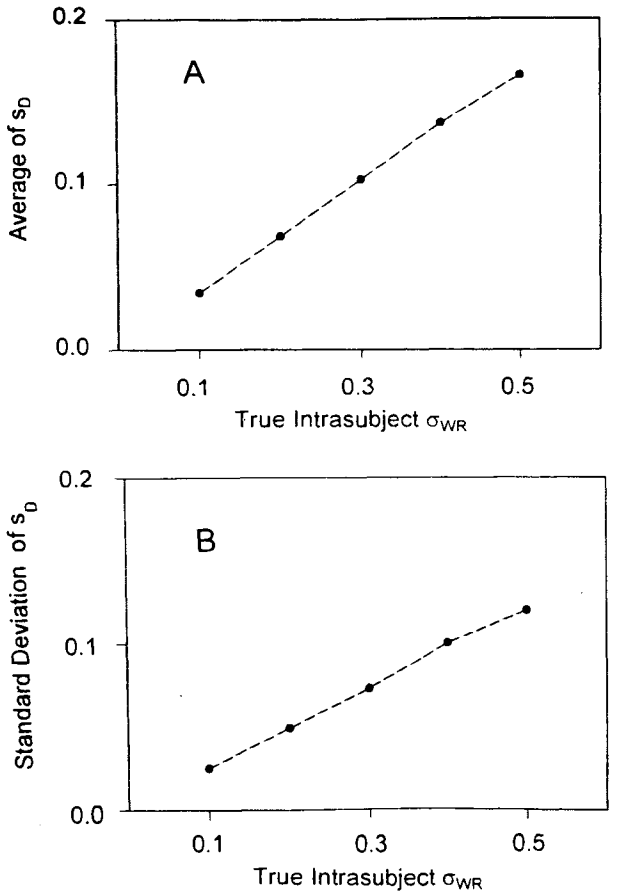

Fig. 4. Dependence of average and standard deviation of the estimated square-rooted variance components for the subject-by-formulation interaction (s_D) on the intrasubject variability of the reference formulation (σ_{WR}). (A) Average; (B) standard deviation.

even when the true magnitude of the parameter, on which the simulations are based, is $\sigma_D^2 = 0$.

It is interesting that the bias of the estimated σ_D^2 increases proportionately with the true intrasubject variance of the reference formulation (σ_{WR}^2). It appears that higher levels of σ_{WR}^2 raise not only the bias of the estimated interaction but also the uncertainties of both estimates of variance components, s_D^2 and s_{WR}^2 (Fig. 3B). As a consequence, at higher levels of intraindividual variability, the recorded s_D value can exceed 0.15 quite frequently (Fig. 3B, Table 1).

Therefore, in the actual absence of interaction ($\sigma_D^2 = 0$), the upper limit of its value, estimated by REML, is dependent on the intrasubject variances. Consequently, fixed levels of s_D (such as $s_D = 0.15$ suggested by the interim Draft Guidance (5)) are not useful for testing the probable presence or absence of the subject-by-formulation interaction.

In most simulations of the present study, the inter-and intrasubject variabilities were assumed to be identical. However, in a few additional simulations, the true intersubject standard deviations (and coefficients of variations) were 3 times as large as the corresponding intrasubject variabilities. The patterns between s_D and s_{WR} remained entirely unchanged. Consequently, intersubject variabilities may have little if any effect on the behaviour of the estimated interaction.

The presented investigations assumed also the true bioequivalence, in all respects, of the two formulations. It would be interesting to explore the effects of bioinequivalence, i.e. of unequal means and inter-and intraindividual variabilities. Such investigations will be undertaken in the future.

The second main result of the present study is the close similarity of the relationship between the variance components σ_D^2 and σ_{WR}^2 estimated for the FDA data set, on the one hand, and in simulations assuming complete bioequivalence, including $\sigma_D^2 = 0$, on the other hand (Figs. 1 and 2).

The similarity strongly suggests that this assumption can be applied to the FDA data. In other words, the FDA data set is compatible with the hypothesis of the absence of subject-by-formulation interactions. In both analyses, most s_D estimates are under the corresponding s_{WR} values. Only few s_D estimates exceed the level of s_{WR}, and they can be considered to occur randomly.

M. Spino and Y.C. Tsang reported at the AAPS/FDA workshop in March 1998 a statistical analysis of the studies within the FDA data set. They found that, for the analyses of AUC_T and C_{max}, the subject-by-formulation interaction was never significantly larger (at $P = 0.05$) than the intrasubject variance of the reference product. This conclusion is in accordance with the results of the present study.

Anecdotal statements suggest differing responses by some subjects when one formulation is substituted by another. However, documented evidence has not been published to date (10). Results of the present study indicate that the data set published by FDA (5) does not provide evidence for subject-by-formulation interactions with the limited sample sizes of the investigations.

Consequently, the concern for the apparently high prevalence of interaction (4,6) does not seem to have been warranted. Therefore, the studies in the FDA data set do not indicate, or even suggest, a possibility of potential danger to public health.

The concern is particularly unnecessary when *small variations* are observed. For example, if a residual variation of less than 15 or 20% is recorded in an analysis of variance for a conventional 2-period, 2-sequence bioequivalence trial then there is very little reason to be concerned about the interaction (10).

Consequently, investigations of individual bioequivalence should not be expected when the residual variation is small in a 2-way bioequivalence study. Thus, the suggested constant-scaled (unscaled) criterion for the evaluation of individual bioequivalence should generally not be applied.

Even with *highly variable drugs*, results of the present investigation suggest that little concern is generally needed for the prevalence of subject-by-formulation interactions. This conclusion raises questions about the design and evaluation of bioequivalence studies involving formulations of such drugs. Application of reference-scaled analyses or their equivalent has been recommended (2–5,11). Should these be performed within the framework of individual bioequivalence which requires the conduct of 3-or 4-period crossover trials, or are 2-period studies sufficient? A strong argument can be made in favour of the latter, more simple design and the corresponding analysis. If the interaction is not important then the residual mean squares, calculated in the related analysis of variance, estimate a pooled intrasubject variance. This could then be usefully applied for scaling.

Multiperiod studies can evaluate also whether intrasubject variations of the two formulations are identical or at least

similar. It is not clear, in view of the large uncertainty of this comparison (12), whether investigation of this question is generally useful. Therefore, it should be normally sufficient to perform a reference-scaled analysis based on 2-period investigations. When, however, a sponsor believes that a new formulation would have more favourable properties, by having smaller intrasubject variability than the marketed one, then this should be demonstrated in a multiperiod trial.

It is conceivable that postmarketing surveillance of new formulations, conducted in many patients could, at times, identify unusual responses in some subjects and with some drugs. However, studies of individual bioequivalence, with their comparatively small sample sizes and use of healthy subjects, may generally not be able to detect subject-by-formulation interactions. Therefore, a regulatory expectation or requirement to undertake such investigations would not be justified.

CONCLUSIONS

The variance component for subject-by-formulation interaction (σ_D^2), estimated by restricted maximum likelihood, was found to have positive bias. The bias rose with increasing intraindividual variations of the reference formulation (σ_{WR}^2). Consequently, a preset value for the interaction (e.g., $\sigma_D = 0.15$) is not suitable to identify the presence or absence of interaction.

Features of the estimated interaction were very similar in a data set published by FDA and in simulations which assumed the actual absence of interaction ($\sigma_D^2 = 0$). Consequently, the FDA data are fully compatible with this assumption, and they do not demonstrate the high prevalence of the interaction. Therefore, investigations of individual bioequivalence and applications of replicate-design studies are not needed when intraindividual variations are small; their expectation is questionable even in the presence of large variations.

ACKNOWLEDGMENTS

We are grateful for the perceptive and helpful comments of Drs. J. W. Hubbard, K. K. Midha, and M. J. Rawson, University of Saskatchewan, on an earlier version of the manuscript.

REFERENCES

1. S. Anderson and W. W. Hauck. Considerations of individual bioequivalence. *J. Pharmacokin. Biopharm.* **18**:259–273 (1993).
2. R. Schall and R. L. Williams. Towards a practical strategy for assessing individual bioequivalence. *J. Pharmacokin. Biopharm.* **24**:133–149 (1996).
3. M. L. Chen. Individual bioequivalence - a regulatory update. *J. Biopharm. Stat.* **7**:5–11 (1997).
4. 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).
5. 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.
6. FDA. Bioequivalence studies: bioequivalence data. Published on Internet: www.fda.gov/cder/bioequivdata/index.htm. (1998).
7. R. Schall and H. E. Luus. On population and individual bioequivalence. *Stat. Med.* **12**:1109–1124 (1993).
8. B. Efron and R. J. Tibshirani. An Introduction to the Bootstrap. Chapman and Hall, New York (1993).
9. FDA. Statistical methods for obtaining confidence intervals for individual and population bioequivalence criteria: technical summary. Published on Internet: www.fda.gov/cder/bioequivdata/statproc.htm. (1998).
10. L. Endrenyi, G. L. Amidon, K. K. Midha, and J. P. Skelly. Individual bioequivalence: attractive in principle, difficult in practice. *Pharm. Res.* **15**:1321–1325 (1998).
11. A. W. Boddy, F. C. Snikeris, R. O. Kringle, G. C. G. Wei, J. A. Opperman, and K. K. Midha. An approach for widening the bioequivalence limits in the case of highly variable drugs. *Pharm. Res.* **12**:1865–1868 (1995).
12. L. Endrenyi and Y. Hao. Asymmetry of the mean-variability tradeoff raises questions about the model in investigations of bioequivalence. *Int. J. Clin. Pharmacol. Ther.* **36**:450–457 (1998).