



## INDIVIDUAL VARIATION AND THE ACCEPTANCE OF AVERAGE BIOEQUIVALENCE

LASZLO ENDRENYI

Department of Pharmacology, University of Toronto, Toronto, Ontario, Canada

MIKLOS SCHULZ

SciAn Biostatistics, Toronto, Ontario, Canada

*Regulatory criteria established for the acceptance of average bioequivalence do not fully satisfy the requirements arising for the interchangeability of drug products. The divergence is illustrated by noting that the regulatory criteria are compatible with large inter-individual variation of within-subject differences or ratios of the relevant kinetic parameters. Regulatory criteria of three jurisdictions are considered: the United States, Canada, and the Province of Ontario. For example, in a crossover trial conducted with 24 subjects, in as many as 60% of the individuals, the difference between the relevant kinetic parameters for the two formations could, on average, be outside the range of 70-130% of the reference product, and still satisfy the criterion of the Food and Drug Administration (FDA) for the acceptance of average bioequivalence. It is suggested that individual bioequivalence can be effectively evaluated in three- and four-way crossover trials.*

*Key Words: Individual bioequivalence; Average bioequivalence; Drug interchange- ability; Drug regulatory criteria*

### INTRODUCTION

ANDERSON AND HAUCK (1) recently described a procedure for the evaluation of individual bioequivalence. They saw a need for such an approach because of the following parallel, related, but not identical, considerations.

Regulatory criteria currently use statistical procedures in order to compare generally the averages of a pharmacokinetic parameter observed in test and reference drugs products. Physicians and some regulatory agencies may be more concerned with the interchangeability of drug formulations in patients. They would wish to see that the two products yield closely similar time profiles of plasma concentrations, and ultimately of drug effects in individual.

The assessment of average bioequivalence generally relies, for a given kinetic parameter, either on the standard error of the mean difference between the two formulations, or on related quantities such as confidence intervals. By contrast, the evaluation of individual bioequivalence involves as reference a measure of individual variation such as the standard deviation.

The standard error declines when the number of subjects is increased, whereas the standard deviation is

independent of the number of the subjects. The relationship between the two view points of bioequivalence is of interest. Therefore, the magnitude of individual variation will be illustrated which is compatible with various currently effective regulatory criteria for the acceptance of average bioequivalence.

## PROCEDURE

Criteria for accepting the bioequivalence of two drug products generally have the form equivalent to:

$$m_T - m_R \leq CR$$

Here, m refers to either the median or the mean of concentrations or kinetic parameters. These could be in their original, linear form or transformed, for example, to logarithms. The subscript T indicates the test and R the reference formulation. CR is the confidence range, possibly different in the positive and negative directions; the two ranges cover the confidence interval.  $m_T$  and CR can be expressed as percentage values against the reference product; in this case  $m_R = 100\%$ .

Acceptance criteria for bioequivalence generally require that the confidence interval be within a preset range which is established from regulatory considerations. For example, an expectation of the FDA of the United States (2) can be restated: the conventional 90% confidence interval for the average area under the curve contrasting concentration with time (AUC) should be within 80 and 120% of the average reference value of AUC. This formulation is numerically equivalent to the two one-sided tests procedure (3), executed at the 5% level, which is the explicitly stated criterion of FDA.

The test is most powerful, and the numerically equivalent confidence interval is the shortest, if the investigated quantity has normal distribution (4). Since AUC is a linear combination of concentrations, the condition of its normality is satisfied if the concentration measurements themselves are normally distributed. Simultaneously, a statistical requirement of the analysis of variance (ANOVA), which is to be performed for the crossover trial, is also fulfilled.

For normally distributed observations in a two-way crossover trial, the confidence range is:

$$CR = t\sqrt{2s^2} / N$$

Here N is the number of subjects in the trial,  $s^2$  is the residual variance estimated from the ANOVA, and t is Student's t- statistic with degrees of freedom d.f. Generally, d.f. = N - 2. It will be assumed in the calculations that differing periods and sequences of drug administration have no effect.

CR is often identical in the positive and negative directions. This is the case, for example, with the normal distribution which is symmetrical. Then, the largest possible variance ( $s_L^2$ ) which is compatible with a symmetrical acceptance criterion for bioequivalence is obtained when the means (or medians) of the two products are identical ( $m_T = m_R$ ) and can be calculated from:

$$s_L^2 = N \cdot CR^2 / (2t^2)$$

For instance, with the FDA criterion, CR = 20 and

$$s_L^2 = 200N/t^2$$

Similar considerations can be applied following a transformation of the data. For example, the Health Protection Branch (HPB) of Canada requires (5) that 95% confidence limits for differences of the average, logarithmically transformed AUCs be within  $\pm \log 1.25$ . This is equivalent to the expectation that the confidence limits for the ratios of geometric means of AUCs be between 0.80 and 1.25.

The criterion used by the Province of Ontario is similar to that of the FDA, except that the application of 95% and not 90% confidence intervals is expected (6). Having computed the largest variance,  $s_L^2$ , compatible with a regulatory criterion, features for the distribution of individuals can be evaluated. For example, if AUCs are normally distributed, then so are the differences in their values which are recorded for the two drug products. The largest variance of the differences compatible with the regulatory criterion is  $2s_L$ .

Therefore, from features of the normal distribution, the fraction of subjects can be calculated who have AUC differences or ratios either between or outside preset values.

Such properties will be presented for three jurisdictions, the United States, Canada, and Ontario.

## **RESULTS**

Table 1 illustrates individual variability which is compatible with the acceptance criterion for bioequivalence established by the FDA. The proportion of subjects (in %) who can have differences in AUCs (in %) outside the given ranges while participating in a crossover trial in which the average bioequivalence is judged to satisfy, barely, the regulatory criterion is shown.

For example, in a two-way crossover trial with 24 subjects,  $N = 24$ ,  $t$  with d.f. = 22 is, at a two-sided significance level of 10% (that is, a one-sided level of 5%), equal to 1.717. Consequently, with the FDA criterion of CR = 20,

$$s_L^2 = 24 \times 20^2 / (2 \times 1.717^2) = 1628$$

is calculated. With  $\sqrt{2}s_L = 57.1$ , and considering individual differences between the AUCs of the two drug products within 70 and 130% of the AUC of the reference formulation, the standard normal variate is  $30.0/57.1 = 0.526$ , provided that AUCs can be thought to be normally distributed. Therefore, under this limiting condition, 40.1% of the AUC differences will be, on the average, within the stated range of 70% and 130% and 59.9% will be outside. The latter percentage is entered in Table 1 at the indicated range between 70 and 130% and for a trial to be conducted with 24 subjects.

Not surprisingly, a larger number of subjects will be contained within, and fewer outside, a wider range of AUC differences. It could be less obvious that the proportion of individuals outside a given

TABLE 1  
Individual Variability of Difference in AUC Which is Compatible with the Acceptance Criterion  
For Bioequivalence Established by the FDA

Number of Subjects	Percentage of Subjects Outside Range of			
	80-120% <sup>a</sup>	70-130%	60-140%	50-150%
8	49.2	30.3	16.9	8.6
12	60.1	43.3	29.5	19.1
16	66.0	50.9	37.9	27.1
20	69.8	56.1	43.8	33.2
24	72.6	59.9	48.3	38.1
28	74.7	62.9	51.9	42.0
32	76.4	65.3	54.8	45.3
36	77.8	67.2	57.3	48.1
40	79.0	68.9	59.4	50.5

<sup>a</sup>Percentage differences of AUCs measured for the two drug products

range of AUC differences rises with an increasing number of subjects. But a given regulatory criterion allows in a crossover trial a larger intersubject variance ( $s_L^2$ ) when the number of participating individuals is raised. The resulting distribution of AUC differences is wider and, consequently, a larger proportion of the subjects is at its defined edges.

The most important conclusion from Table 1 is that a large fraction of individuals can have substantially differing AUCs for the two drug products while a test for average bioequivalence still satisfies the regulatory acceptance criterion. The same impression is gained from Figure 1 which shows individual variation compatible with the acceptance criteria of FDA, HPB, and Ontario. It also reflects the increasingly tight requirements of these jurisdictions.

Table 1 and Figure 1 consider conditions which evaluate the largest possible interindividual variance ( $s_L^2$ ), the widest distribution of AUC differences, and the most subjects with AUC differences outside the preset limits. Such conditions can be noted when the acceptance criterion is just barely satisfied. However, in a given trial, the estimated confidence limits are well within the region of acceptance. Consequently, in most trials, smaller intersubject variance is actually observed and the proportion of subjects outside the regions given in Table 1 and Figure 1 is lower than the limiting, maximal fraction indicated there.

It is noted also that the theoretical proportions of individuals shown in Table 1 and Figure 1 can be viewed as averages of observations. In any single trial, even if the interindividual variance of within-subject differences reaches its "regulatory limit" of  $s_L^2$ , the observed proportions could be larger or smaller than their theoretical values.

## DISCUSSION

The results illustrate that very substantial variation among individuals is compatible with regulatory criteria, which are applied for the determination of average bio-equivalence. The large variability may be distressing some investigators and raise questions about the actual interchangeability of drug products.

Others will not be surprised, particularly those with clinical experience. The calculated dispersions are similar in magnitude to variations of concentrations seen in subjects when various drugs are repeatedly administered to them.

For the assessment of drug interchangeability, the subject-to-subject variation of intra-individual kinetic differences is relevant. Therefore, it is distressing that little information is available on intra-subject variability, at least in comparison with that gained from studies conducted on inter-individual variation. Examples for systematic observations of within-subject kinetic variability include digoxin (7), beta-blockers such as metoprolol, propranolol, and oxprenolol (8), furosemide (9), labetalol (10), ranitidine (11), theophylline (12, 13), cyclosporine (14), nitrendipine (15) and CGS 16617, an angiotensin converting enzyme inhibitor (16). Upton et al. (17) described a procedure for the evaluation of intrasubject variability of parameters in the simple, one-exponential kinetic model.

Unfortunately, the subject-to-subject of intra-individual variation are changes generally not presented even when the average within-subject variations, based on individual data, are shown.

It is particularly important to consider and assess interchangeability for drugs which have highly variable kinetics. The variations of concentrations and kinetic parameters within individuals can at times be larger than those observed between subjects. If switching to a different drug product were to contribute an additional component to the variation within individuals, then they could be exposed to risk, particularly when the therapeutic index of the drug is small.

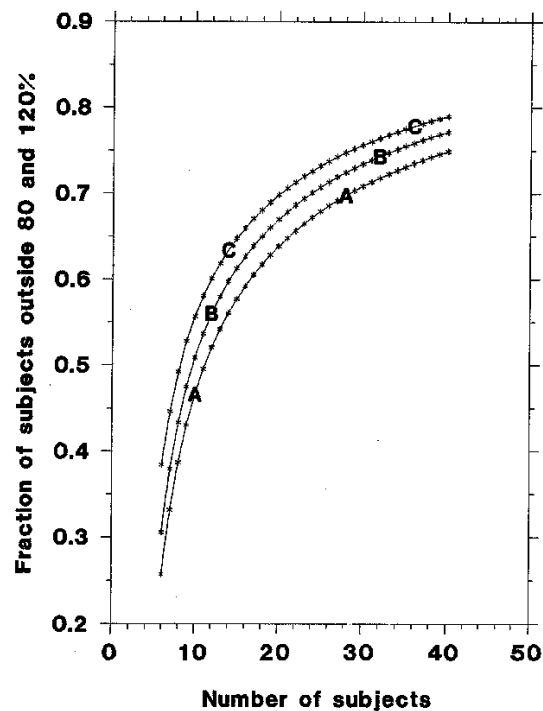


FIGURE 1. Mean proportion of subjects in whom the ratio (in %) of kinetic parameters is outside the range of 80 and 120% while regulatory criteria for the acceptance of average bioequivalence are still satisfied. The calculated mean proportions assume that relative bioavailability is evaluated in a two-way crossover trial with the indicated number of subjects. The regulatory criteria are those of (A) the Province of Ontario, (B) the Federal Health Protection Branch in Ottawa, and (C) the Food and Drug Administration of the United States.

It is emphasized that evaluations of both the individual and average bioequivalence do not imply their acceptance or rejection. Such decisions should involve regulatory considerations which are based on features of clinical responses to the drug. As a result, regulatory acceptance criteria could be either tightened or relaxed depending on the pharmacodynamics of the drug.

Regulatory considerations are not identical for the assessment and acceptance of average and individual bio-equivalence. Approval of a drug product by agencies such as the FDA in the United States and HPB in Canada, enables physicians to prescribe it. The decision is based on the general consideration of efficacy and safety. Therefore, these agencies are concerned primarily with average bioequivalence.

On the other hand, direct maintenance of and payment for health care is generally the responsibility of the states and provinces in the United States and Canada, respectively. Their regulatory bodies often wish to ascertain the effect of substituting one drug product for another in individuals. This parallels the concerns of each patient and physician.

It is likely that the quantitative evaluation of drug interchangeability will be the subject of extensive discussions in the near future. The proposal of Anderson and Hauck (1) for the calculation of individual bioequivalence has already been mentioned. Their nonparametric procedure appears to be strict in its conditions and will probably lead to the acceptance of individual bioequivalence fairly infrequently. Ekbohm and Melander (18) suggested that drug interchangeability be evaluated from the analysis of subject-by-formulation inter-action in crossover trials.

Assessment of individual bioequivalence by population kinetic analysis can be anticipated in analogy to calculations of average bioavailability (19). It is also expected that evaluation of the distributional features of individual bioequivalence will be explored. Mallet's (20) nonparametric approach to population kinetic analysis could be useful in this regard.

Still, the acceptance of individual bioequivalence could be based on a readily available standard. Consider a four-way crossover trial in which each subject receives twice both the reference and test formulations. The repeated observations obtained on the two applications of each drug product provide a measure not only of the within-subject variation but also of its inter-individual scatter. This should be similar, for the acceptance of individual bioequivalence, to the scatter of differences between responses to the two drug products which are measured within individuals.

Actually, simplified experimental designs can yield the information needed for the assessment of individual bioequivalence. Three-way crossover trials would contain only a single set of replications. This provides a more limited but still sufficient, basis for the contrast with differences between individual responses to the two drug products. It may well be that this will turn out to be the most effective approach to the evaluation of drug inter-changeability.

In summary, it has been demonstrated that large variability of within-subject differences is compatible with current regulatory criteria for the acceptance of average bioequivalence. Therefore, the reminder of Anderson and Hauck (1) that separate standards be developed and established for the assessment of drug interchange- ability is repeated and reinforced. It is quite possible that this can be accomplished effectively and rather simply in three- and four-way crossover trials.

## NOTE

Since this paper was presented and written, efforts toward international harmonization of criteria applied to adopt the equivalence of drug products (21) have been very successful. Thus, both HPB and FDA have agreed to use the standard proposed by the European Community (22): the 90% confidence limits around the ratio of the geometric means AUCs for the test and reference products should (based on logarithmic AUC calculations) be within 0.80 and 1.25.

The changed regulatory criteria would modify some details of the quantitative statements presented in the paper. The resulting modifications would be very small for the criterion of the FDA (satisfying the new criterion, the relative AUCs of the two drug products could be outside the range of 0.7 and 1.3, on the average, in 62.8% of 24 subjects). In any case, the essential conclusions about the relationship between average equivalence and individual variation remain unchanged.

## REFERENCES

1. Anderson S, Hauck WW. Consideration of individual bioequivalence. *J Pharmacokinet Bio-pharm*. 1990; 18 : 259 - 273 .
2. Report by the bioequivalence task force on recommendations from the bioequivalence hearing (September 29 to October 1, 1986). Rockville, Md: US Food and Drug Administration; 1988.
3. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Bio-pharm*. 1987;15:567- 680).
4. Steinijs VW, Hauschke D. Update on the statistical analysis of bio-equivalence studies. *Int J Clin Pharmacol Ther Toxicol*. 1990;28 : 105- 110.
5. Expert Advisory Committee on Bio-availability. Report on the Bioavailability of Oral Dosage Formulations of Drugs Used for Systemic Effects. A. Drugs with Uncomplicated Characteristics. Canada: Health Protection Branch; 1990.
6. Interpretative Guidelines for Manufacturers Making Drug Submissions. Ontario Ministry of Health; 1986.
7. Yacobi A, Stoll RG, Chao GC, Schwartz R A, Weidler DJ, Ayers DJ, Sakmar E, Hallmark M, Wagner JG. The assessment of the intra-subject variability in digoxin absorption in man from two oral dosage forms. *J Clin Pharmacol*. 1981; 21:301-310.
8. Jack DB, Quarterman CP, Zaman R, Kendall MJ. Variability of beta-blocker pharmacokinetics in young volunteers. *Eur J Clin Pharmacol*. 1982;23:37-42.
9. Grahnén A, Hammarlund M, Lundqvist T. Implications of intraindividual variability in bioavailability studies of furosemide. *Eur J Clin Pharmacol*. 1984;27:595-602.
10. Awni WM, Skaar DJ, Schwenk MH, Plaschek JR, Matzke GR. Interindividual and intraindividual variability in labetalol pharmacokinetics. *J Clin Pharmacol*. 1988;28:344-349.
11. Shim CK, Hong JS. Inter- and intrasubject variations of ranitidine pharmacokinetics after oral administration to normal male subjects. *J Pharm Sci*. 1989;78:990-994.
12. Upton RA, Thiercelin J-F, Guentert TW, Wallace SM, Powell JR, Sansom L, Riegelman S. Intra-individual variability in theophylline pharmacokinetics: statistical verification in 39 of 60 healthy young adults. *J Pharmacokinet Biopharm* 1982;10:123-134.
13. Dederich RA, Szeffler SJ, Green ER. Intrasubject variation in sustained-release theophylline absorption. *J Allergy Clin Immunol*. 1981 ;67 :465- 471.
14. Lemaire M, Fah r A, Maurer G. Pharmacokinetics of cyclosporine: inter- and intra-individual variations and metabolic pathways. *Transplant Proc*. 1990;22: 1 110-1 112.
15. Lobo J, Jack DB, Kendall MJ. Inter- and intra- subject variability of nitrendipine and the effects of food. *Eur J Clin Pharmacol*. 1987;32:357- 360.
16. Kochack GM, Smith RA, Choi RL, John V, Tipnis V, Honc F, deSilva JK, Weidler DJ. Intra- and intersubject variability: mixed-effects statistical analysis of repeated doses of an angiotensin converting enzyme

- inhibitor, CGS 16617. *Pharm Res.* 1989;6:328-331.
17. Upton RA, Thiercelin J-F, Moore JK, Riegelman S. A method for estimating within-individual variability in clearance and in volume of distribution from standard bioavailability studies. *Pharmacokinetics Biopharm.* 1982; 10:135-146.
  18. Ekbohm G, Melander H. The subject-by-formulation interaction as a criterion of inter-changeability of drugs. *Biometrics.* 1989;45: 1249-1254.
  19. Graves DA, Chang I. Application of NONMEM to routine bioavailability data. *J Pharmacokinetics Biopharm.* 1990; 18:145-160.
  20. Mallet A. A maximum likelihood estimation method for random coefficient regression models. *Biometrika.* 1986;73:645-656.
  21. Cartwright AC, Gundert-Remy U, Rauws G, McGilveray I, Salmond T, Walters S. International harmonization and consensus DIA meeting on bioavailability and bioequivalence testing requirements and standards. *Drug Inf J.* 1991;25:471-482.
  22. CPMP (Committee for Proprietary Medicinal Products). Working Party on the Efficacy of Medicinal Products. Notes for guidance: Investigation of bioavailability and bioequivalence, 1991.