[Note: The document below was previously provided to the Committee as background material for the April 2004 ACPS meeting]

Background Information for Advisory Committee Meeting On April 14, 2004

Bioequivalence Requirements for Highly Variable Drugs and Drug Products

Introduction

Bioequivalence studies are generally conducted by comparing the *in vivo* rate and extent of drug absorption of a test and a reference drug product in healthy subjects. In a standard *in vivo* bioequivalence study design, participants receive a single dose of test and reference products on separate occasions with random assignment to the two possible sequences of product administration. Samples of an accessible biologic fluid such as blood or urine are analyzed for drug concentrations, and pharmacokinetic measures such as area under the curve (AUC) and peak concentration (C_{max}), are obtained from the resulting concentration-time profiles. To evaluate bioequivalence, the U.S. Food and Drug Administration (FDA) has employed a testing procedure termed *the two one-sided tests procedure* to determine whether the average values for the pharmacokinetic measures from the test and reference products are comparable[1]. This procedure involves the calculation of a confidence interval for the ratio between the average values of the test and reference product[2]. In the U.S., a test product is considered to be bioequivalent to a reference product if the 90% confidence interval of the geometric mean ratio of AUC and C_{max} between the test and reference fall within 80-125%. Currently, the bioequivalence limits of 80-125% have been applied to almost all drug products by the FDA[3].

Concerns have been expressed at times regarding the difficulty of meeting the standard bioequivalence criteria for highly variable drugs and/or drug products[4,5]. To date, there is no regulatory definition for these drugs or drug products. In the context of bioequivalence, however, drugs and drug products exhibiting intra-subject variability greater than 30% C.V. (coefficient of variation) in the pharmacokinetic measures, AUC and/or C_{max} are considered highly variable[4,5]. To pass the conventional "goalposts", the number of subjects required for a study of these drugs or drug products can be much greater than normally needed for a typical bioequivalence study. Thus, the resource implications coupled with the ethical concern of exposing a large number of healthy subjects to a test drug further challenges the appropriateness of the conventional bioequivalence criteria for highly variable drugs/products. Examples exist of a highly variable reference product failing to demonstrate bioequivalence with itself using the standard design/sample size for a bioequivalence study[6].

The issue of highly variable drugs/products in bioequivalence has been discussed in many conferences and meetings, nationally and internationally. However, there is no universal

consensus or solution at this time. The objectives of this paper are to: a) explore the need for applying alternative bioequivalence limits for highly variable drugs/products; b) review the available proposals for alternative criteria and/or limits; and c) discuss possible regulatory approaches for resolution of the issue for these drugs/products.

Background

Although global harmonization is a general goal, to date, bioequivalence has not been accepted as a topic by the International Conference on Harmonization (ICH). Nonetheless, the resource and ethical concerns for highly variable drugs/products in bioequivalence are generally recognized by international regulatory agencies. It is thus useful to review the differing regulatory approaches before an informed recommendation is made on the topic. The following outlines the bioequivalence standards used in different regions:

In Canada, for drugs with uncomplicated characteristics, a 90% confidence limit of 80-125% is required for AUC. However, a limit is placed only on the means (or point estimate) for C_{max}[7]. As a result of random variation or a larger than expected relative difference, the sponsor may add more subjects. If this option is chosen, it must be stated in the study protocol. In addition, two criteria must be met before combining is acceptable:

1) The same protocol must be used; and

2) Consistency tests must be met at an alpha error rate of five percent.

The European Agency for the Evaluation of Medicinal Products (EMEA) has similar bioequivalence standards to those in the FDA, *i.e.*, 90% confidence limits of 80-125% on AUC and C_{max}, with the qualification that these limits may be expanded in certain cases for C_{max} (*e.g.*, 75-133%) provided that there are no safety or efficacy concerns[8].

In Japan, the bioequivalence standards also rely on the 90% confidence limits of 80-125% for both AUC and C_{max}, although wider limits are allowed for less potent drugs. Additionally, if the confidence limits are outside of 80-125%, bioequivalence may be claimed on the grounds that the study meets all three conditions listed below[9].

1) The total number of subjects in the initial bioequivalence study is no less than 20 (n=10/group), or pooled sample size of the initial and add-on studies is no less than 30;

2) The differences in average values of logarithmic AUC and C_{max} between two products are between log(0.9) – log(1.11); and

3) Dissolution rates of test and reference products are determined to be equivalent under all dissolution testing conditions specified.

Japan allows the addition of subjects to increase the power of a failed bioequivalence

study. However, the add-on subjects can not be less than half the number in the original study.
South Africa accepts an acceptance interval of 75-133% for C_{max}, except for narrow therapeutic range drugs, when an acceptance interval of 80-125% applies[10]. For highly variable drugs, a wider interval or other appropriate measure may be acceptable, but should be stated a priori and justified in the protocol.

Proposals from the Literature

As indicated, the bioequivalence criteria in the U.S. recommend that the 90% confidence interval of the geometric mean ratio between the test and reference products fall within 80-125%. Over the years, various suggestions have been made in an attempt to alleviate the difficulty of meeting the bioequivalence limits for highly variable drugs and drug products.

Various authors have explored the use of replicate designs or group-sequential designs. If a subject-by-formulation interaction is negligible, the sample size required for a replicate design study can be reduced up to 50% of that for a non-replicate design study[11]. The number of study periods is the same since approximately half the usual number of subjects is used but they are each studied for twice as many periods. Therefore, it takes a longer time to complete a replicate design study, resulting in an increased chance of subject dropout from the trial. A group-sequential design may be useful in cases where there is uncertainty about the estimates of variability. Nonetheless, the total number of subjects employed with this design may be the same as that used for a study without the group-sequential design if the interim analysis does not indicate bioequivalence[11]. Also, to preserve the overall Type I error rate of 5%, a higher level of confidence interval has to be used at each stage of the interim analysis[11].

Several proposals are available in the literature to modify the existing bioequivalence criteria for highly variable drugs and drug products[12,13]. In general, these various criteria are based on either the reduction of the level of the confidence interval or an increase of the width of the equivalence limits, or both.

The level of confidence interval reflects the degree of consumer risk (Type I error in statistical terms) that can be tolerated by the regulatory agencies. A reduction in the level of confidence interval, for example, from 90% to 85%, implies a possible increase in the consumer risk, which would not be in the best interests of public health. In contrast, the width of equivalence limits represents the allowable boundary for the ratio (or difference) of the means between products in comparison. Any adjustment of these limits should be based on consideration of the statistical properties of the data as well as on the clinical characteristics of the individual drug. Statistically, widening the bioequivalence limits can be accomplished through expansion of the allowable boundary or by scaling the criteria based on the high variability of the reference product.

Discussed below are the proposals available to date for widening the acceptable limits of pharmacokinetic measures in bioequivalence studies.

A. *Direct Expansion of Bioequivalence Limits*

Sample size in bioequivalence studies is determined in large part by the bioavailability parameter with the highest variability. In most cases, C_{max} has higher variability than AUC. Thus, widening of the bioequivalence limits for C_{max} has been proposed to reduce the sample size needed in the evaluation of bioequivalence for highly variable drugs/products. The greater variability observed with C_{max} may result from the fact that this parameter is a single point measurement, which is highly dependent on the sampling time/frequency and elimination rate of the drug.

The EMEA currently allows for expanded limits (*e.g.*, 75-133%) for C_{max} in certain cases where no safety or efficacy concern arises, based on the consideration of higher variability for this measure as compared to AUC[8].

B. *Expansion of Bioequivalence Limits Based on Fixed Sample Size*

This method was proposed based on the notion that only a reasonable number of subjects should be required for a bioequivalence study[4,12].

The number of subjects is fixed by a standard two-period, crossover study comparing the reference product with itself where the study fails to meet the 80-125% limit. The confidence interval obtained from the reference product in this study would become the "goalposts" for the subsequent studies comparing the test with reference product, using the same number of subjects.

C. *Widening of Bioequivalence Limits Based on Reference Variability*

The bioequivalence limits for these methods are not determined by the sample size. Rather, they will be scaled based on the within-subject variability of the reference product. For both Methods 2 and 3 below, a side condition to constrain the mean difference between the test and reference products has also been proposed (see **Discussion**).

Method 1:

The rationale for this approach is that a mean difference of 25% is considered small relative to the range of values an individual may experience when the within-subject variability is high, *e.g.*, 40%. Therefore, the acceptable limits may be scaled in relation to the size of within-subject variability as follows[12]:

$$[U, L] = \text{Exp}\,[\pm k\sigma_{WR}] \qquad \text{(Eq. 1)}$$

where U and L are the upper and lower limits, respectively; k represents the pth percentile of the standard normal distribution, Zp; and σ_{WR} is the estimated

within-subject standard deviation (obtained from the ANOVA on the log scale) for the reference. When k = 1, ~ 67% of the pharmacokinetic measures (such as AUC) experienced by an individual will be within the range of [U, L]. Table I lists the choices of limits at k = 1.

Table I

CV (%)	SD (σ_{WR})	Lower Limit	Upper Limit
30	0.294	0.75	1.34
35	0.340	0.71	1.40
40	0.385	0.68	1.47
45	0.429	0.65	1.54
50	0.472	0.62	1.60

Different k values could be chosen for different drugs depending on their therapeutic windows.

Method 2:

A scaled average bioequivalence criterion has been proposed[13,14,15]:

$$(\mu_T - \mu_R)^2 / \sigma^2_{WR} \leq \theta \qquad \text{(Eq. 2)}$$

where μ_T and μ_R are the averages of the log-transformed measure for the test and reference products, respectively; and θ is the bioequivalence limit. Comparing Methods 1 and 2, it can be seen that $k = \theta^{-1/2} = (\ln 1.25)/\sigma_{W0}$ where σ_{W0} is the cutoff within-subject standard deviation for scaling. Table II shows the relationship of k and σ_{W0}.

Table II

σ_{W0}	k
0.20	1.116
0.223	1.0
0.25	0.893
0.294	0.759

Method 3:

Derived from the comparison of the distance measure between the test and reference products, the following individual bioequivalence criterion has a reference variance in the denominator, and thus is scaled to the reference variability (10, 12):

$$[(\mu_T - \mu_R)^2 + (\sigma^2_{WT} - \sigma^2_{WR}) + \sigma^2_D] / \sigma^2_{WR} \leq \theta_I \qquad \text{(Eq. 3)}$$

where σ_{WT} is the estimated within-subject standard deviation for the test product; σ^2_D is the subject-by-formulation interaction variance component; and θ_I is individual bioequivalence limit.

Although theoretically sound, the individual bioequivalence criterion requires replicate designs and inclusion of target population in the study. Because of these resource implications, the FDA has recommended the continued use of an average criterion to compare bioavailability measures[3].

D. *Expansion of Bioequivalence Limits Based on Sample Size and Scaling*

In addition to fixing the sample size, this method takes into consideration the producer's risk (Type II error) and reference variability[12]. The equation for the allowable limits is:

$$[U, L] = \text{Exp} [\pm (t_\alpha + t_{\beta/2}) n^{-1/2} \sigma_{WR}] \qquad \text{(Eq. 4)}$$

where α and β are the consumer and producer risks, respectively; 2n is the number of subjects desired in the study; and t is the percentile of the t-distribution with 2n-2 degrees of freedom.

The current regulatory standard has kept the consumer risk at a level of no more than 5% while allowing the drug applicant or sponsor to control its own producer risk. Based on Eq. 4, for example, assuming a 5% consumer risk and 10% producer risk, the proposed bioequivalence limits for a typical sample size of 24 subjects will be

(0.74, 1.35) at $\sigma_{WR} = 0.3$

(0.61, 1.65) at $\sigma_{WR} = 0.5$

Discussion

The impact of C_{max} variability on the determination of bioequivalence, as well as the possible approaches to resolving this issue, has been discussed extensively in the published literature. Major regulatory agencies have provisions in their regulations which can accommodate the effect of higher variability associated with C_{max} on the design of bioequivalence studies. For example, *Health Canada* does not require any limits on the confidence interval for C_{max}, although limits are placed on the point estimates for this parameter. The EMEA and Medicines Control Council of South Africa both allow for expanded limits for C_{max} in certain cases provided that there are no safety or efficacy concerns[8]. The expanded limits are not defined, although they cite 75-133% as an example. Similarly, the Japanese Division of Drugs accepts limits greater than 80 –

125%, "for drugs with pharmacologically mild actions"[9]. Additionally, a failed bioequivalence study can utilize additional subjects to increase power and the likelihood of meeting BE criteria, provided other conditions are met.

Tothfalusi *et al*., compared scaling bioequivalence limits for highly variable drugs to widening of the limits around C_{max}[15]. Scaling may involve widening of the confidence interval limits as a function of the variability of the reference drug product. The authors' conclusion was that scaling would significantly reduce the sample size needed for bioequivalence studies of highly variable drugs. Additionally, they concluded that the same result could be achieved by simple expansion of the regulatory limits to 75-133% or even 70-143% for C_{max}.

Simple expansion of the regulatory limits may lead to acceptance of BE for drug products with mean ratios for C_{max} exceeding 125%. This possibility was discussed by Hauck *et al*.[16]. The authors reported that widening the confidence limits to 70-143% could allow acceptance of C_{max} ratios of 128%. A difference of this magnitude in the point estimate may not be acceptable for many drugs. This possibility, however, may be eliminated by placing an additional regulatory constraint on the point estimate for C_{max}, which would accompany any expanded limits of the confidence interval.

Another approach, using reference scaling for all drugs, will effectively widen the equivalence limits for highly variable drugs/products. The method, however, should not be used for drugs exhibiting low intra-subject variability in the reference product since it may unnecessarily narrow the equivalence limit beyond the public-health need. The choice of cutoff for reference scaling (σ_{W0}) will have to be made by the regulatory agency. This approach, however, has a discontinuity at the changeover point (σ_{W0}) from no scaling to reference scaling. For example, if the estimate of the within-subject standard deviation of the reference is just above the changeover point, the confidence interval will be wider than just below. In this context, the confidence interval could pass the predetermined bioequivalence limit if the estimate is just below the boundary and could fail if just above. Another question that may be raised for using the scaling method is the reliability of the estimate for the reference variability although this may be achieved by increasing the sample size or setting minimum requirements for the precision of the estimate.

With the exception of direct expansion of bioequivalence limits, it appears that all the reference-scaling approaches either require a study with replicate design or need more than one study to allow determination of reference variability.

Discussion Topics for the ACPS Meeting

April 14, 2004

1) Highly variable drugs or drug products may be defined as those exhibiting intra-subject variability of 30% CV or greater in AUC or Cmax.

   Does the committee concur?

2) The Advisory Committee is asked to comment on the following approaches and whether there is promise in developing one or both of these approaches to improve the bioequivalence assessment of HVDs.

   a) Direct Expansion of Bioequivalence Limits: Change from 80-125%, and restrict the mean T/R difference, *e.g.*, ± 20? What information is necessary to properly set these new confidence interval limits?

   b) Reference Scaling: Scale current bioequivalence criterion based on the reference variability in each study and restrict the mean T/R difference as above.

References

[1] D.J. Schuirmann. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokinet. Biopharm.* 15: 657-680 (1987).

[2] W.J. Westlake. Response to Kirkwood, TBL.: Bioequivalence testing – a need to rethink. *Biometrics*. 37:589-594 (1981).

[3] U.S. Food and Drug Administration, Center for Drug Evaluation and Research. *Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products--General Considerations*, Office of Training and Communications, Division of Communications Management, Drug Information Branch, HFD-210, Rockville MD 20857, March 2003.

[4] H. Blume and K.K. Midha. Report of Consensus Meeting: Bio-international '92. Conference on Bioavailability, Bioequivalence and Pharmacokinetic Studies. Bad Homburg, Germany, 20-22 May, 1992. *Eur. J. Pharm. Sci.* 1:165-171 (1993).

[5] Bio-international '94, Munich Germany. 14-18 June, 1994.

[6] Health Canada, Therapeutic Products Directorate (TPD). Discussion paper on 'Bioequivalence requirements - highly variable drugs and highly variable drug products: issues and options'. Expert Advisory Committee on Bioavailability and Bioequivalence (EAC-BB) Meeting, June 26-27, 2003.

[7] Health Canada, Ministry of Health, *Guidance for Industry: Conduct and Analysis of Bioavailability and Bioequivalence Studies – Part A: Oral Dosage Formulations Used for Systemic Effects.* 1992.

[8] Committee for Proprietary Medicinal Products (CPMP), the European Agency for the Evaluation of Medicinal Products (EMEA). *Note for Guidance on the Investigation of Bioavailability and Bioequivalence*. 2001.

[9] Japan National Institute of Health, Division of Drugs. *Guideline for Bioequivalence Studies of Generic Drug Products*. 1997.

[10] Medicines Control Council, Department of Health, Republic of South Africa. *Registration of Medicines: Biostudies*. 2003.

[11] S.D. Patterson, N.M.-D. Zariffa, T.H. Montague, and K. Howland. Non-traditional study designs to demonstrate average bioequivalence for highly variable drug products. *Eur. J. Clin. Pharmacol.* 57:663-670 (2001).

[12] A.W. Boddy, F.C. Snikeris, R.O. Kringle, G.C.G. Wei, J.A. Oppermann, and K.K. Midha. An approach for widening the bioequivalence acceptance limits in the case of highly variable drugs. *Pharm. Res.* 12:1865-1868 (1995).

[13] L. Tothfalusi, L. Endrenyi, K.K. Midha, M. J. Rawson, and J.W. Hubbard. Evaluation of the bioequivalence of highly variable drugs and drug products. *Pharm. Res.* 18:728-733 (2001).

[14] L. Tothfalusi and L. Endrenyi. Limits for the scaled average bioequivalence of highly variable drugs and drug products. *Pharm. Res.* 20:382-389 (2003).

[15] L. Tothfalusi, L. Endrenyi, and K.K. Midha. Scaling or wider bioequivalence limits for highly variable drugs and for the special case of C_{max}. *Int. J. Clin. Pharmacol. Ther.* 41:217-225 (2003).

[16] W.W. Hauck, A. Parekh, L.J. Lesko, M.-L. Chen, and R.L. Williams. Limits of 80-125% for AUC and 70-143% for C_{max}. *Int. J. Clin. Pharmacol. Ther.* 39:350-355 (2001).