Determination of Bioequivalence for Drugs with Narrow Therapeutic Index: Reduction of the Regulatory Burden

Laszlo Endrenyi¹ and Laszlo Tothfalusi²

¹ University of Toronto, Department of Pharmacology and Toxicology, Toronto, ON, Canada; ² Semmelweis University, Department of Pharmacodynamics, Budapest, Hungary

Received, October 21, 2013; Revised, November 16, 2013; Accepted, November 19, 2013; Published, November 22, 2013

ABSTRACT - The US Food and Drug Administration (FDA) has recently suggested that the bioequivalence (BE) for products of drugs with narrow therapeutic indices (NTI) be assessed by the approach of reference-scaled average BE (SABE). Subsequently, in December, 2012, the FDA issued draft guidances for the comparison of products of warfarin sodium and of tacrolimus. The guidances expect that 4-period studies be performed, that the results be evaluated by SABE, and that the analysis include also unscaled average BE as well as the comparison of the estimated within-subject variations ($s^2_w$) of the test and reference drug products. This communication discusses the new guidances and suggests considerations to reduce the regulatory burden. It is demonstrated that SABE could be applied when the within-subject variation of the reference product is not higher than 21.42%. Beyond this variation, the BE limits would remain 80% to 125%, as usual. No further testing by unscaled average BE is needed. It is also suggested that a comparison of the within-subject variations of the two drug products although interesting for both NTI and other drugs, is not essential for the determination of BE. In addition, when the within-subject variabilities are low then their ratio depends mainly on the non-product dependent factors. Moreover, introduction of an additional test would affect the probabilities involved in the primary comparison of the two means. Therefore, the test of comparing variances is not needed and replicate measurements of the test formulation need not be performed. Alternative considerations and approaches, including the use of partial AUC’s, are suggested for the determination of BE for NTI drugs.

This article is open to POST-PUBLICATION REVIEW. Registered readers (see "For Readers") may comment by clicking on ABSTRACT on the issue's contents page.

INTRODUCTION

The Food and Drug Administration (FDA) published recently a new draft guidance on the determination of bioequivalence of warfarin sodium formulations (1). The new guidance has a number of new and even novel features. This is interesting to potential sponsors of bioequivalence (BE) studies and to those interested in regulatory sciences. Subsequently, another draft guidance was issued on the bioequivalence of tacrolimus formulations (2). It referred for the methodology to be applied to the draft guidance on warfarin (1).

The new features include the application of the approach of scaled average bioequivalence (SABE) to the bioequivalence of warfarin or tacrolimus products and, by implication, to preparations of drugs with narrow therapeutic indices (NTI) (3). The new warfarin draft guidance presents, for the first time, also a (disaggregated) criterion for the comparison of within-subject variations. These features will be described later in greater detail.

There is no strict definition of what is an NTI (or critical-dose) drug. A general definition is, as Yu noted, that with NTI drugs “small differences in dose or blood concentration may lead to serious therapeutic failures and/or adverse drug reactions” (4). Warfarin is an NTI drug for which several investigations compared the effects of different products. Reviews of studies noted that brand-name and most generic products had closely similar bioavailabilities and clinical outcomes (5-7). Nevertheless, caution was often offered in studies.
and editorials against switching among warfarin products. Tacrolimus is also an NTI drug (2, 8, 9). Its various formulations were found to be bioequivalent in healthy subjects (e.g., 10-13). However, concern was expressed about the scarcity of BE data in transplant patients (14).

The question arises if, with the new features of the warfarin draft guidance, the determination of BE for NTI drug products is implemented most advantageously. The purpose of this communication is to explore some answers to this question and to suggest alternative possibilities.

First, relevant aspects of the new warfarin draft guidance will be summarized. This will be followed by an exploration of some of its quantitative features. Comments and recommendations will be presented of alternatives which may have benefits of reduced regulatory burden and enhanced simplicity.

BACKGROUND

Draft Guidance on the Bioequivalence of Warfarin Sodium Products

The procedures proposed in the recent warfarin draft guidance (1) will be briefly described together with relevant background information when necessary.

The draft guidance suggests that fully replicated 4-way crossover in vivo studies be performed on healthy males and nonpregnant females chosen from the general population. The studies are to be undertaken under both fasting and fed conditions. Preparations of 10 mg tablets of warfarin sodium are to be administered orally.

The method of reference-scaled average bioequivalence (SABE) should be applied for the statistical comparison of the relevant pharmacokinetic parameters. The approach has been proposed to evaluate the BE of highly-variable drugs and drug products (15). FDA has adopted and implemented it (16, 17). The procedure modifies the usual requirement for the determination of BE. Typically, the means (µ) of a test (T) and a reference (R) formulation are compared following the logarithmic transformation of the data. It is expected for (unscaled) average BE that the 90% confidence interval around the ratio of the geometric means between the two drug products be within the range of 0.80 to 1.25.

In the approach of scaled average BE, the difference between the logarithmic means is standardized by the within-subject standard deviation of the reference formulation (σWR):

\[-θ_A ≤ (µ_T - µ_R)/σ_{WR} ≤ θ_A\]  [2]

The bioequivalence limits for the scaled, SABE procedure (θA) and the unscaled average BE (BEL) are related by:

\[θ_A = \ln(\text{BEL})/σ_{W0}\]  [3]

Here σW0 is a constant the value of which is set by the regulatory authorities.

The draft guidance on the BE of warfarin products recommends BEL = 1/0.9 = 1.11111 and σW0 = 0.10 (1). Consequently, the scaled average BE limit is θA = 1.054.

Steps of Statistical Analysis Recommended by the Draft Guidance

The steps presented by the draft guidance will be briefly summarized.

Step 1. Calculate the estimated within-subject standard deviation of the reference product (σWR), from the results of the completed study, by the usual procedure applied to a (balanced) replicate design.

Step 2. Apply the procedure of reference-scaled average BE (SABE) to each pharmacokinetic parameter. Evaluate the upper 95% confidence bound to the linearized, squared and rearranged form of SABE:

\[(µ_T - µ_R)^2 - θ_A^2 * σ_{WR}^2 ≤ 0\]  [4]

By inserting estimated values of the means and the variance, and their distributions, the approximate confidence bound can be calculated (18, 19) following the approach of Howe (20).

Step 3. By using unscaled average BE, evaluate, for each pharmacokinetic parameter, the 90% confidence interval around the ratio of geometric means of the two drug products. The 90% confidence limits should be, as usual, between

\[-\text{BEL} ≤ µ_T - µ_R ≤ \text{BEL}\]  [1]
Step 4. Evaluate the upper limit of the 90% confidence interval for the ratio of within-subject standard deviations of the two drug products. The ratio of the variances is characterized by an F-distribution. The calculated 90% upper limit should not be higher than 6.25, i.e., the estimated upper limit for the ratio of the standard deviations should not exceed 2.5.

POSSIBLE REDUCTIONS OF THE REGULATORY BURDEN

Confidence Limits for Unscaled Average BE

It will be useful to consider scaled average BE in a corresponding but different form. Multiplying the defining equation [2] by $\sigma_{WR}$ yields:

$$- \theta_A^* \sigma_{WR} \leq \mu_T - \mu_R \leq \theta_A^* \sigma_{WR} \quad [5]$$

The expression characterizes unscaled average bioequivalence, as in Eq. 1, but with expanding limits (ABEL) (19, 21). The limits are proportional to the within-subject standard deviation ($\sigma_{WR}$).

The approach of ABEL has the advantage over the procedure of SABE that it can be easily assessed by Schuirmann’s two one-sided tests method (22). The expanding limits can be calculated by using the values for $BEL = 1.11111$ and $\sigma_{W0} = 0.10$ suggested by the draft guidance for warfarin. The limits are illustrated against increasing within-subject variation in Figure 1. (The diagram displays the variation in terms of the intra-subject coefficient of variation.) Similar figures were shown at a meeting of the FDA Advisory Committee but without the interpretation given here (23-26).

As expected, the implied BE limits are seen to expand with increasing variation. However, it is not desirable that the BE limits be wider than the customary range of 0.80 to 1.25. This is the intent of Step 3 of the draft guidance. Thus, above a certain variation, the widening of the BE limits should stop. This is also illustrated in Figure 1. The figure demonstrates that the widening stops at a coefficient of variation of 21.42% beyond which the BE limits remain constant.

It can thus be concluded that the approaches of SABE and ABEL may be applied if, and only if, the within-subject variation of the reference product does not exceed 21.42%. Therefore it is sufficient to estimate the value of $\sigma_{WR}$ and see if it is below or above 0.2118 (corresponding to CV = 21.42%). No additional test is needed to demonstrate the separation of regions where either scaled or unscaled average BE should be evaluated. Consequently, Step 3 of suggested testing procedure is superfluous and unnecessary.

Determination of the Within-Subject Variation of Both Drug Products

The warfarin draft guidance suggests that the within-subject variations of not only the reference formulation but also of the test product be estimated ($\sigma_{WR}$ and $\sigma_{WT}$, respectively). This can be accomplished, as the draft guidance recommends (1), if both formulations are measured twice in each subject.

As noted earlier, the draft guidance expects that the two variations be compared, in Step 4, by an F-test. This is a new requirement in the area of evaluating BE studies. It is not clear that it is needed any more for NTI drugs than for drugs not in this category.

The within-subject variations of drugs and drug products are interesting. These data are always “nice to know”. However, it is not clear that they are essential in determinations of BE for either NTI or other drugs. Similarly, placing a cap on the deviation between the variations observed for the two formulations is potentially attractive. However, it is not more important for NTI drugs, notably for warfarin and tacrolimus, than for other drugs.
Therefore it is suggested that it is not necessary to replicate measurements of the test product within individuals. Also, Step 4 suggested in the warfarin draft guidance (1) is not needed. Two additional comments are offered in regard to the comparison of variances in BE studies. First, the proposal was offered and discussed extensively earlier in connection with the evaluation of individual and population bioequivalence (27, 28). At that time, aggregated regulatory criteria were considered. The warfarin draft guidance presents a disaggregated regulatory expectation. Secondly, introduction of a second regulatory criterion, that for the comparison of variances, would affect and alter the features for the errors of the principal criterion, that for the comparison of the means. This aspect is of concern and should be further investigated.

DISCUSSION

Are Separate Estimates of Within-Subject Variations Needed for NTI Drug Products?
It is questionable that the comparison of \( \sigma_{WR} \) and \( \sigma_{WT} \) is meaningful if they are low which is typical for NTI drugs. The within-subject variation is a net effect of errors from many different sources. When \( CV_{WR} \) is low (say, less than 20%) then error sources include, first, the bioanalytical error. The bioanalytical guidelines (29, 30) require only that the precision determined at each concentration level should not exceed 15% except at the lower limit of quantitation (LLOQ) at which it should not be higher than 20%. Thus, the bioanalytical error component can be quite substantial, particularly for \( C_{max} \). Other potential and often neglected error sources are the differences between the actual and nominal contents of active ingredients and the within-batch variations.

Furthermore, warfarin is a BCS Class 1 drug (31). Tablets of warfarin have uncomplicated technology and its bioavailability is nearly complete. If warfarin were not an NTI drug then its bioavailability could be decided based on in-vitro tests. Therefore, it is questionable that screening for product-related specific differences between \( \sigma_{WR} \) and \( \sigma_{WT} \) is a reasonable approach for warfarin. Thus, separate estimates of \( \sigma_{WR} \) and \( \sigma_{WT} \) may not be needed for low-variability BCS Class 1 drugs.

Consistency between Regulatory Requirements
There are many regulatory requirements which should be met by generic products to get approval. For instance, the United States Pharmacopeia (USP) specifies for content variability that 10 tablets from any batch must contain 85% to 115% of the labelled strength, with a standard deviation of less than 6%. This requirement is more liberal than the proposed bioequivalence criterion. Thus, a situation could arise that allowed the tablet-to-tablet difference of the reference product to be potentially larger than the allowed difference between the test and reference formulations. This is counterintuitive and forces to rethink the consequences of changing only one of the regulatory rules. Stability requirements could be another, similar issue. Thus, bioequivalence expectations should be set in harmony with other quality requirements in order to avoid inconsistencies in the regulatory framework.

For warfarin generics, producers often apply stricter content variability standards than the originator (32). It is not clear that such stricter criteria are general and applicable in other cases.

Alternative Approaches
It is still debatable whether a specific regulation is needed for the bioequivalence of NTI drugs. There are only sporadic clinical reports about negative consequences of switching. Hard evidence in the form of randomized clinical trials is not available and not even expected to be available in the near future. But using data of a small bioequivalence trial with carbamazepine (CBZ), which is an NTI drug, it was demonstrated that the concern is not imaginary. Two have been offered which could support the cautionary approach to generic switching with this drug (33, 34). First, it has been reported that the relative risk of neurological adverse effects increases by as much as 50% due to switching between two CBZ formulations which were otherwise bioequivalent. In that case, \( C_{max} \) was not a sensitive metric which would be able to capture clinically important differences. As a remedy, the use of partial AUC was suggested for assessing the bioequivalence of CBZ tablets (33).

A second explanation is based on the observation that the assumption that bioequivalence results obtained in a healthy population in single-dose studies can be extrapolated to the target patient population at the steady state, is not always true (34). CBZ is a strong enzyme inducer (35) and the \( C_{max} \) and \( C_{min} \) concentrations depend on the clearance. A small difference between \( C_{max} \) values in a healthy population can be significantly higher in the target group. Two formulations which are bioequivalent in a single-dose study with healthy
volunteers may not be bioequivalent at steady state in patients.

Neither of these problems, the inadequate metric and/or inadequate study population, is addressed directly by the proposed new regulation. Instead, it focuses on the presumed problem of variation difference between the test and reference formulations. It remains to be demonstrated that this is a real clinical problem or just an interesting statistical concept.

**CONCLUDING REMARKS**

The procedure in the new draft guidance of FDA on the BE of warfarin sodium products is remarkably innovative (1). Most importantly, it introduces the application of the scaled average BE procedure to the BE of formulations of a drug having a narrow therapeutic index. Another example is the new draft guidance on the comparison of tacrolimus formulations (2).

The present communication recognizes the main feature of the new draft guidances, the introduction of the procedure of scaled average BE for certain NTI drugs. At the same time, suggestions are offered in order to reduce the regulatory burden.

The first recommendation is simple and straightforward. It is not necessary to calculate the 90% confidence limits with unscaled average BE (Step 3 in the draft guidance) in order to ascertain if the limits obtained with SABE would or would not exceed the 0.80 to 1.25 range. If the estimated within-subject variation of the reference product does not exceed 21.42% then, automatically, the unscaled average BE limits are not penetrated.

The second suggestion observes that FDA introduces a new, additional regulatory criterion comparing, by an F-test, the within-subject variances of the two formulations and placing a cap on their ratio. This information is interesting but not more important and essential for comparing formulations of NTI drugs than of other drugs. It is further noted that the additional regulatory criterion modifies the statistical features of the primary criterion that of comparing the means. Therefore it is suggested that Step 4 described in the warfarin draft guidance (1) is not needed. Also, it is not necessary to obtain replicate measurements of the test formulation within individuals.

The suggestions and remarks in the present communication arise from and immediately apply to the recently issued FDA draft guidances on the BE of warfarin sodium and tacrolimus products (1, 2). However, the agency indicated that it intends to apply the approach of scaled average BE, more generally, to products of drugs having narrow therapeutic indices (3, 25, 36). The Advisory Committee on Pharmaceutical Science and Clinical Pharmacology supported this intention (37). Thus, our suggestions and remarks may have wider relevance.

It is often assumed that NTI drugs have small within-subject variations (3, 38) and that the assessment of BE for their products should involve narrower than the usual BE limits. The assumption of small intrasubject variation is not always valid. For example, for dalbigatran, in a replicate-design study, the geometric CVs of the two formulations for AUC were 62.2% and 44.3%, and for C<sub>max</sub> 64.7% and 45.5% (39).

As a consequence of this assumption, regulatory guidances call for narrower bioequivalence limits by, for instance, EMA, Health Canada and, now, FDA (2, 3, 8, 40). However, questions can be raised about this regulatory expectation. For instance, alternative metrics such as partial AUC, could provide a more effective approach for dealing with the issue of BE for NTI products (33).

**ACKNOWLEDGMENT**

We appreciate the perceptive comments of Dr. Yu Chung Tsang of Apotex on an earlier version of this manuscript.

**REFERENCES**


