Pharmaceutical Research, Vol. 13, No. 11, 1996

Workshop Report

Evaluation of Orally Administered Highly Variable Drugs and Drug Formulations[1]

Vinod P. Shah,[2] Avraham Yacobi, William H. Barr, Leslie Z. Benet, Douwe Breimer, Michael R. Dobrinska, Lazslo Endrenyi, William Fairweather, William Gillespie, Mario A. Gonzalez, John Hooper, Andre Jackson, Larry J. Lesko, Kamal K. Midha, Patrick K. Noonan, Rabi Patnaik, and Roger L. Williams

CONFERENCE OBJECTIVES

An AAPS/FDA co-sponsored workshop was held on March 6–8, 1995 in Arlington, VA to discuss issues and difficulties in the bioequivalence (BE) evaluation of highly variable pharmaceutical drug products and to review criteria and possible solutions in the evaluation of these products. The objectives of the workshop were to:

(1) identify sources of variability affecting assessment of bioequivalence;

(2) identify issues in the evaluation of bioequivalence of highly variable drug products;

(3) propose methods for assessing bioequivalence and acceptance criteria for evaluation of highly variable drugs and drug products; and

(4) develop a workshop report that could be used as a reference for developing guidelines for evaluation of highly variable drug products.

Toward these objectives, the workshop addressed the following questions:

What is a highly variable drug?
What are the sources of variability and can they be overcome?
What are practical and reliable metrics to evaluate bioequivalence?
Do bioequivalence criteria address "switchability", i.e., the substitution of one product for another in a given patient?
Should the acceptance criteria for bioequivalence be changed?

This report summarizes the deliberations with respect to these issues and other topics discussed at this meeting. These conclusions and recommendations represent a general consensus of meeting participants.

STATEMENT OF PROBLEM

Drugs and drug products that exhibit intrasubject variability of greater than 30% Analysis of Variance Coefficient of Variation (ANOVA-CV) in bioavailability parameters are generally referred to as highly variable drugs. Some examples of highly variable drugs include cyclosporine, chlorpromazine, erythromycin, isosorbide dinitrate, nitroglycerine, methylphenidate, sulindac, diltiazem, and verapamil. When products exhibit high intrasubject variability, the number of subjects needed to demonstrate bioequivalence, in studies using current criteria, becomes excessive, often requiring more than 48 subjects. Even for products with 25–30% CV, difficulties occur in demonstrating bioequivalence with a reasonable number of subjects. The current bioequivalence criteria requires that the 90% confidence interval of the difference between the log transformed mean values of the AUC and Cmax of the test product (T) and those of the reference product (R) must fall within the specified bioequivalence limits (BL), which presently are set at 80% to 125% of the reference mean for each parameter. The two critical criteria are the confidence interval (CI) which determines the frequency that an acceptable product would be tolerated, and the bioequivalence limits (BL) which determines the difference between the means of the pharmacokinetics (PK) parameters of two products that are considered clinically acceptable. An important consideration, therefore, is the degree to which a given difference in bioequivalence would result in a meaningful clinical difference.

Generally, a change in bioavailability will be more important for drugs with a narrow therapeutic index (or therapeutic ratio), in which the amount (or concentration) of drug in the body required to produce a therapeutic response is close to the amount that will produce a significant adverse effect. Classes of drugs could be categorized as:

I. Narrow therapeutic index and low intrasubject variability.

II. Narrow therapeutic index and high intrasubject variability.

III. Broad therapeutic index and low intrasubject variability.

IV. Broad therapeutic index and high intrasubject variability.

Nevertheless, the bioequivalence criteria in the U.S. are the same for all drugs regardless of the therapeutic range. It

[1] This document represents a consensus of the personal views of the authors or presenters. It does not necessarily represent the policies or guidelines of the American Association of Pharmaceutical Scientists (AAPS), FDA or any other organization.

[2] To whom all correspondence should be addressed at Office of Pharmaceutical Science, Center for Drug Evaluation and Research, Food and Drug Administration, 7500 Standish Place, Rockville, Maryland 20855.

0724-8741/96/1100-1590$09.50/0 © 1996 Plenum Publishing Corporation

has been suggested that an approved drug with narrow therapeutic range will generally exhibit relatively low intrasubject variability, or it would not have been approved. On the other hand, drugs which exhibit high intra subject variability will succeed only if they possess relatively wider therapeutic ranges. In case of drugs for which these safety generalizations hold, then the bioequivalence criteria for high intrasubject variability drugs may not need to be as rigorous, and the acceptance criteria may be individualized.

The intrasubject variability can be inherently due to the drug substance itself, or it can be due to the formulation factors of the product. New advances in analytical technology and basic science knowledge have improved the understanding and ability to identify sources of variability. Managing variability entails managing its pharmacodynamics, pharmacokinetics, and formulation components. Managing the pharmacokinetics variability requires a good estimate of its extent of intrasubject variability, defining the metabolism pathways involved, and understanding the impact of these variabilities on the drug formulations. The *in vitro* (dissolution rate)-*in vivo* (PK parameter) correlations, if any, can aid in determining some of the formulation variability factors.

The interaction between the nature of the variability and the therapeutic index of the drug product and other concepts, including switchability, individual therapeutic windows and individual bioequivalence were discussed to address the clinical implications of bioequivalence criteria.

The workshop presentations and discussions were focused on the issue of understanding and reducing the sources of variability and thus decreasing the number of subjects required in BE studies without sacrificing the quality of the studies needed for the evaluation and approval processes. The most important consideration was whether the evaluation of bioequivalence ensures that products deemed bioequivalent can be used interchangeably in the target population.

SOURCES OF VARIABILITY AND POTENTIAL SOLUTIONS

Different sources of high variability in AUC and Cmax which affect the assessment of bioequivalence were defined and discussed. These sources include the physiologic and pathophysiologic variables of absorption and post-absorptive disposition (e.g., metabolic clearance) and product formulation factors.

Several variables of GI physiology, such as regional pH, bile and pancreatic secretions, luminal and mucosal enzymes, GI motility, gastric emptying, small intestinal transit time, and colonic residence time, vary widely within individuals which can influence drug absorption and bioavailability. The fasted/fed state of the subjects can greatly influence bioavailability and BE assessment because food affects many of these GI variables. The pH of GI tract is often influenced by food and its buffer capacity which interacts with other variables such as the age and achlorhydric condition of the subject/patient. Complexation of food with certain drug and/or excipient can occur. Presence of food often stimulates bile and lecithin secretion which result in emulsification and solubilization, and may increase drug absorption, especially for drug products with poor aqueous solubility. Food may also greatly influence gastric emptying and to a lesser extent, the transit time of the dosage form through the small intestine. These variations in total transit and regional conditions may significantly contribute to variability in the rate and extent of absorption. However, these factors affect intersubject variability more than intrasubject variability.

External gamma scintigraphy in combination with conventional pharmacokinetics studies provide valuable information about the influence of gastric emptying and small intestinal transit time. Special delivery systems and intubation studies can also provide information on sites of drug absorption. If GI physiological variables are determinants of the rate limiting step for absorption, it is likely that the drug will exhibit high drug absorption variability.

The second major source of variability is in the drug formulation. In order to limit individual variability due to drug product variables, it is essential to maintain within-batch homogeneity. Batch-to-batch homogeneity and consistency of the products across manufacturers (e.g., test and reference) are also essential to minimize variability due to drug formulation. These sources of variability can usually be assessed through *in-vitro*-dissolution profile testing in various media of different pH values.

Metabolic factors may constitute a major source of inter- and intrasubject variability in highly variable drugs. Drug disposition and post-absorptive clearance can be influenced by gender, age, disease state and concurrent drug therapy.

The mechanism and site of first pass metabolism greatly affects the inter- and intrasubject variability of the administered drug. If the metabolism occurs in the intestinal lumen or mucosa it may be more susceptible to formulation effects. First pass metabolism may also result in significantly different metabolite levels. In instances, where a metabolite is a measure for bioequivalence, such as in the case of a prodrug, nonquantifiable parent levels, or where a metabolite is the active species, a highly variable drug presents a different and difficult issue to be resolved. Same criteria may be used when metabolite is the primary analyte (when parent cannot be quantified) in BE evaluation.

It is now well documented that there is a large variability in the activity of cytochrome P450 enzymes, especially 3A4 along the absorption pathway and thus they may play a significant role in determining the extent of absorption and metabolism of drugs. In assessing BE of highly variable drug products where first pass metabolism is considerable, change in rate of absorption may also result in a significant change in the extent of intact drug reaching the systemic circulation. Cytochrome P450 enzyme systems are subject to a large number of external influences, such as age, disease, diurnal rhythm and dietary interactions. Gender may also be a factor. Controlling these external influences of variability in the study design is critical in reducing variability. External and intrinsic influences on P450 activity can alter bioavailability. Therefore, knowledge about the site and type of enzyme system involved in pre- and post-systemic drug metabolism is essential for designing rational bioequivalence studies.

Formulation factors can also contribute to intrasubject variability. Often intrasubject variability is more dramatic with coadministration of the drug with food. Food studies for NDAs are generally undertaken for safety and prescribability concerns. Currently all ANDAs of extended release products require a study to determine the effect of food on the bioavailability of the drug. For immediate release products, ANDAs require food studies only if the innovator's product is known to be adminis-

tered with food or if a special labeling is included. The food studies are generally evaluated using only a point estimate rather than using complete bioequivalence criteria (confidence interval with bioequivalence limits) in the United States. At present, even though the product is labeled "dose with meals," the pivotal study for evaluating bioequivalence criteria is the fasting study. It was suggested that in such a scenario, when the product is labeled to be administered with food, the pivotal bioequivalence study should be a food study.

STUDY DESIGN AND STATISTICAL EVALUATION

The standard bioequivalence study involves a two way crossover study design with test and reference products. For IR products, a fasted single dose study and for ER products, fasted single dose, fed single dose, and fasted multiple dose studies are required. Critical metrics in bioequivalence evaluation include log transformed mean values of AUC and Cmax. The test product should meet the criteria of a log transformed mean value of 80–125% with 90% C.I. compared to the respective mean values of the reference product. Generally, Cmax and Tmax values provide only minimal information on the absorption rate. Use of mean partial AUC 0-t (eg., t = time) and partial AUC equal or greater than 80% of AUC 0-∞ to evaluate both rate and extent of absorption are suggested. Use of mean partial AUC permits more samples to be harvested around Tmax, and thus provide better estimates for rate and extent of absorption. When required, for IR products, the food study is conducted to show that the mean AUC and Cmax of the test product are within 20% of the reference product.

Replicate study designs have shown significant advantages in the estimation of intrasubject variability, particularly to determine treatment-subject interaction, and also have usefulness in individual bioequivalence measurements. Replicate study designs which allow replication of both test and reference product can also provide evidence as to the quality of both test and reference products in terms of their variability. However, under current drug approval criteria, the only practical advantage of replicated crossover studies is the ability to obtain more observations for the same number of subjects. Replicated crossover studies have a number of disadvantages, including greater cost and longer duration of the study which may result in significant dropout of the subjects participating in the study. The major concern with a replicate design is the technical difficulties of analysis and unresolved regulatory issues. The latter is true, for example, when greater than two sequences are used in the replicate design.

Average bioequivalence relies on population means without consideration of variability. Population bioequivalence relies on population distributions, whereas individual bioequivalence takes into consideration the product-subject variability and is linked to the concept of switchability. It was suggested that a new frame work for bioequivalence should be considered in terms of assuring switchability and taking into account the individual therapeutic window.

Several additional designs, such as group sequential and "add-on" studies were discussed. "Add-on" designs are valid, only if they are part of well-designed studies allowing addition of a second panel of subjects to increase statistical power. This study design is now allowed by the Agency for bioequivalence studies for inhalers and topical glucocorticoids where PD observations are recorded and for which intrasubject variability is shown to be relatively high resembling the case of highly variable drug products.

The bioequivalence standard may be modified by changing the bioequivalence limits of log transformed mean value of 80–125% (width of goal posts) while maintaining the current C.I. at 90%, i.e., keeping the consumer risk at 5%. The bioequivalence limits should be determined based on the intrasubject variability, pharmacodynamic/pharmacokinetic data, therapeutic indices or other clinical characteristic of the reference product. The metric of choice for the extent of absorption is AUC, preferably AUC 0-t (equal to or greater than 80% of the estimated AUC 0-∞) which needs to be evaluated and accepted by the Agency. A second metric (designated as reflecting rate of absorption) should also be used and evaluated by C.I. criteria, although the second metric need not have the same bioequivalence limits as AUC. Alternative metrics for rate of absorption should be defined.

CONCLUSIONS AND RECOMMENDATIONS

The following are the recommendations based on deliberations at the workshop. A brief discussion of the rationale follows when applicable.

Formulation Factors

Variability Due to Formulation Should Be Monitored and Minimized by More Rigorous Efforts to Develop in Vitro *(Dissolution)*/in Vivo *Correlations*

The critical manufacturing variables for each orally administered product should be defined relative to appropriate dissolution procedures which have a demonstrated relationship to *in vivo* (bioavailability) data. *In vitro* data should be obtained throughout the pH range expected in the GI tract/pH 1–7.5. *In vivo-in vitro* correlation should be established, wherever possible to aid in specification of the drug product.

Chirality

Chirality in BE evaluation should be studied and established, because there may be differences in absorption and disposition between individual isomers of a drug product.

Food Studies

The Current Guidelines for Food Studies Do Not Address Switchability Issues and Should Be Modified

The food study should use rational acceptance criteria, rather than the current ±20% which does not have a statistical basis for food study when the dosage form is administered with food. The acceptance criteria should meet confidence interval criteria when the drug is indicated for administration with food. The bioequivalence limits could be different than the standard 80% to 125% based upon intrasubject variability and the therapeutic range of the drug product. In such cases, the fasted study should be evaluated in terms of point estimate i.e., ±20% of the mean value of the reference.

Bioequivalence Requirements Should Be Harmonized for the NDA and ANDA Processes

There are many inconsistencies in the processes and criteria used to establish BE for a generic product development (ANDA process) as compared to a new drug formulation development process e.g., formulation and site changes. Sometimes the differences simply reflect lack of standardization, for example, the different types of food composition used in ANDA and NDA food studies.

Study Design

Statistically-valid Multi-stage Designs Should Be Permitted

Multi-stage studies would include group sequential designs and interim "blinded-looks" to determine sample size. This recommendation excludes "add-on" designs where additional subjects are simply added to a completed and analyzed study. Studies must include proper statistical methodology that is defined in the protocol. It is emphasized that new procedures should not reward studies with excessive variability (i.e., poorly controlled studies). Gender, weight and age should be taken into account when soliciting subjects.

Manufacturers Should be Encouraged to Submit Data from all Studies Including Failed or Incomplete Studies. This Includes all ANDA Studies and Post-approval NDA Studies.

FDA lacks sufficient data to assess the importance of formulation factors contributing to intra- and inter-subject variability and hence bioequivalence. Industry should be encouraged to provide additional information in the form of a summary on:

(1) failed bioequivalence studies.
(2) *in vitro-in vivo* correlations, and;
(3) results from replicate dose studies;

These studies will provide the FDA necessary information on study variables such as degree and sources of variability. It is unfortunate and inconsistent that sponsors of drug products approved by the ANDA process are not required to provide the Agency with information on all studies that have been done, in contrast to the NDA process. These data should be requested, only if the studies were conducted by the firm.

Statistical Criteria

The Choice of Metrics to Quantify the Extent of Absorption Is AUC

One clear consensus was that the area under the plasma concentration time curve (AUC) is the primary metric to assess the extent of absorption. It is recognized that metrics evaluating unchanged drug measurements define the extent and rate of availability, rather than absorption. Recommendations within this document refer to the absorption, however, in concert with the stated Federal Register definitions of bioavailability and bioequivalence. There is controversy on how the value of AUC should be assessed when data are incomplete and do not define the entire systemic concentration-time profile. A particular example would be the use of AUC 0-t to assess extent of absorption of drugs with very long half lives.

Explore the Use of Truncated AUC to Quantify Extent of Absorption

For some drugs with a long or complex elimination phase, the appropriate use of truncated concentration-time curves can decrease variability, cost, and time and allow more detailed characterization of the absorption process. It is not necessary to always require collection of biological samples for 3 half lives or until 90% of the area is obtained. Specific guidelines should be developed on when and how truncated systemic concentration time curves can be used as a measure of extent of absorption.

A Second Bioequivalence Metric Should Be Used. It Should Also Be Evaluated by CI Criteria, But the BE Limits for the Two Metrics Could Be Different

Generally the second metric should be a measure of the rate of absorption. However, it was not concluded that the metrics be limited to two or that the second metric should exclusively assess rate. The second metric (e.g., Cmax) could be more variable than AUC and therefore might warrant a wider BE limits.

The importance of assessing rate of absorption is clear from a scientific and therapeutic point of view. It is also important from a regulatory and legal point of view since the Hatch-Waxman Statute clearly states that generally both extent and rate of absorption must be shown to be the same to establish bioequivalence between a test and reference formulation. The rate of absorption may be very important for some drugs such as nifedipine relative to hypotensive efficacy or theophylline relative to potential toxicity, but may be unimportant for others such as the bisphosphonates. The basic problem is that the traditional indicators of rate, Cmax and Tmax, are not reliable metrics of the absorption rate. Tmax is greatly dependent upon the sampling frequency and Cmax is a measure of both extent and rate. Although Cmax may be confounded as a measure of rate, it is often a very critical metric for the assessment of the potential adverse effects for some drugs. Therefore, from a clinical point of view, Cmax may be very important regardless of its determinants. Because of these pharmacodynamic differences, it was concluded that the decision to use Cmax and/or some other metric to assess rate should be handled on an individual drug basis.

Alternative Metrics for Rate Should Be Defined

Because of the limitations of Cmax and Tmax as indicators of rate of absorption, it is important to validate better methods to assess differences in rate of absorption. This is particularly important for drugs that exhibit rate-dependent pharmacodynamics and for controlled-release products that are expressly developed to modify the rate of absorption. This may have to be done on a drug-by-drug basis or even on a product-by-product basis.

Bioequivalence Criteria

Development of Methods of Evaluating Individual Bioequivalence Are Critical to the Issue of "Switchability" of Products in an Individual Patient

Current regulatory methods of assessing bioequivalence that use only mean data may obscure differences in products that may be important in individual patients who are substituting one product for another ("switchability"). Previous regulatory requirements used to include the "75-75" rule which evaluated the ratios of individual AUC values. While this approach was flawed for lack of statistical rigour, it made sense clinically. Differences in the individual distribution of values and detection of formulation dependent subsets of individuals could be detected. It is important that these clinical objectives be realized by methods that are also sound statistically.

The Issue of "Switchability" Is Not Adequately Addressed by Current Criteria which Evaluates Only Mean Population Data. Individual Bioavailability Assessment Should Be Added

The criteria for bioequivalence should include criteria for individual bioequivalence. This includes a measure of intra subject variability in both reference and test groups. The methodology should allow detection of significant subsets of the population which may have different bioavailability profiles than the average. The old "75-75" rule which was included in previous acceptance criteria served some of these needs but was discarded because it was not statistically sound. Newer proposals such as probability based and approximate moment-based criteria have statistical properties to allow better definition of individual bioequivalence.

For Some Highly Variable Drugs and Drug Products, the Bioequivalency Standard Should Be Modified by Changing the Bioequivalence Limits (BL) While Maintaining the Current Confidence Interval (CI) at 90%

The recommendation to vary the bioequivalence limits (currently set at 80–125), rather than the CI (currently set at 90%, was done to maintain the current consumer risk (5%). For example, for a drug demonstrated to exhibit high intrasubject variability in rates of absorption, the bioequivalence criteria for the 90% confidence interval for Cmax might be increased from a BL of 80 to 125% to a BL of 70 to 143% while maintaining the consumer risk at 5%.

The Bioequivalence Limits Should be Determined Based in Part Upon the Intrasubject Variability for the Reference Product (δ^2_{WR})

Recommendations were not made on how the δ^2_{WR} should be determined. Ideally this value should be derived from the studies using replicate design. The value of this statistic would be different than the value of δ^2 from residual error which reflects all sources of variability.

Bioequivalence Limits Should Also Be Defined Based Upon Therapeutic Indices and Other Clinical Considerations

δ^2_{WR} is one criterion, but not the only criterion, that should be considered in widening the BL. The importance of the therapeutic window relative to the variability of the drug and drug product is equally important. The BL should not be increased at the cost of clinical efficacy or safety. While the goal is to reduce the numbers of subjects required in a study, this should not preempt considerations of variability due to subsets of individuals which may not be identified in studies with small numbers of subjects. The issue of sources of variability is complicated by the fact that current standards compare only the means, not the individual values, of the bioavailability parameters.

It was concluded that for highly variable drugs, the bioequivalence standard may be modified by changing the bioequivalence limits, i.e., widening the width of goal posts while maintaining the current level of C.I. at 90%. The bioequivalence limits can be varied and changed, rather than altering the level of C.I. for highly variable drugs. Any modification of acceptance criteria should be done on a case-by-case basis using pharmacodynamic or pharmacokinetics data and intrasubject variability for individual drugs or drug products. The 90% C.I. represent the current consumer risk (of 5%) and should be maintained for assuring switchability. The bioequivalence limits should be determined based on the intrasubject variability for the reference product. The therapeutic indices and other clinical considerations should be evaluated for determining the bioequivalence limits for highly variable drugs and drug products.