Individual Bioequivalence: A Problem of Switchability

Sharon Anderson
Bristol-Myers Squibb

Introduction

The area of bioequivalence testing is seeing rapid changes. It has been less than a decade since the use of the correct null-hypothesis of non-equivalence was adopted and now, in the past two years, we have seen the adoption of the concepts of individual and population equivalence. The future should be exciting as the appropriate methodology is developed to address these types of equivalence.

This essay reviews the motivation for the new concepts of individual and population bioequivalence and provides a brief overview of where the research stands today. The Discussion Section will summarize where I believe we should be heading in the near future.

For reasons of brevity, this essay will be restricted to only bioequivalence trials for the approval of generic drugs. Also, while many methods have been proposed for testing bioequivalence, only those methods that directly address individual bioequivalence have been referenced. Finally, the concepts of individual and population equivalence extend to all equivalence studies.

Background

A bioequivalence study for the approval of a generic drug is an equivalence study of bioavailability. Bioavailability is the rate and extent of absorption of the active drug or metabolite at the site of drug action. The common parameters are summary statistics based on the measured concentration of the active drug or metabolite in the plasma (or urine) over time. The most common parameters used to characterize the extent and rate of absorption are the area under the plasma concentration-time curve (AUC), the maximum concentration (Cmax) and the time of maximum concentration (Tmax). Presently, standard two-period crossover designs are most common.

In contrast to most clinical equivalence problems, the test for bioequivalence is an interval hypothesis with the alternative hypothesis being that the ratio of the population means, say \( \text{AUC} \), is greater than 1.25. The choice of equivalence interval may differ by country, compound, drug class, and bioavailability parameter. The primary analyses are done on the log scale. Presently in the U.S., the recommendation is to conduct bioequivalence studies with the narrowest 90% confidence interval that falls within the equivalence interval or, equivalently, if two simultaneous one-sided t-tests are each significant at the 0.05 level. The important point to remember is that the present bioequivalence hypothesis test is a test only of the population means. The comparison of only population means is referred to as "average bioequivalence."
As an aside, in the specific type of bioequivalence study discussed in this essay, one is dealing with a surrogate measure problem. The demonstrated bioequivalence of a test and reference formulation allows the test formulation to claim the clinical results for its reference, i.e., assume the same efficacy and safety profiles. However, there is a catch. The relationships between plasma drug-concentration level and therapeutic response or drug intolerance are not always known. Thus, one challenge is setting equivalence criteria for measures based on active drug concentration when the ultimate objective is equivalence of clinical efficacy and safety.

Motivation for Individual and Population Bioequivalence

The motivation for using individual and population bioequivalence came from a recognition that even with the correctly formulated testing procedures, the practical application of the results, namely the daily use of generic drugs, was not necessarily supported by the statistics. Anderson and Hausk (1990). If you were to ask a physician what he or she thought it meant to be an approved generic drug, the answer would almost likely be, "It doesn't matter which one I prescribe for my patient." If you were to ask the patient the same question, their response would be, "It doesn't matter which one I take." Also, in today's health care system, a patient's formulation may be changed without either the prescribing physician's or the patient's knowledge. This is the practical understanding of a statistical demonstration of bioequivalence and a test of only population means, without severe assumptions regarding the correlation in form, does not support this understanding and application.

There are, however, two different clinical situations that need to be distinguished. First, if the patient is drug naive, the physician has no information as to how the patient will respond to either formulation. This situation is referred to as "prescribability." For this situation, one needs to demonstrate that the differences in the distributions or "population" bioequivalence. What about the patient who has been well controlled with minimal tolerance problems? In this situation, both the physician and the patient expect that the new formulation will provide the same efficacy and safety profile for that patient. This latter situation, when "switchability" is required, corresponds to the "individual" bioequivalence problem. The fact that both formulations were shown on average to have equivalent responses in a population (average bioequivalence) is not sufficient for either of these situations.

One way to understand the differences between these two types of bioequivalence is to consider the plots of individual bioavailability responses on formulations labeled 1 and 2. See Figure 1 (from Ekbom and Melander, 1991). Case 1 is the ideal case of individual bioequivalence. The average responses are the same on both formulations, the variances (total variation) are the same and an individual's response is the same across formulations. The data would satisfy individual population and average bioequivalence. In Case 2, the individual responses are not similar. These data would satisfy population but not individual. Case 2, where only the average responses are similar, would only satisfy average bioequivalence.

Population and individual bioequivalence do not imply statements of the problem of bioequivalence. They require that the distributions be equivalent and not just that the population means be shown to be equivalent as is presently done. Table 1 summarizes the three types of bioequivalence in terms of the underlying distributional parameters which are involved in the testing. The underlying model is shown at the top of the table. In the case of individual bioequivalence, the proposed methods differ in terms of whether they work with the individual means or the population means.

So far the argument for considering individual bioequivalence has been conceptual. Anderson and Hausk (1990) showed that one may conclude an overall bioequivalence in the population, only 57% of the population would have their individual ratios fall within the old equivalence interval (0.8, 1.25). The parameter values used
Table 1: Underlying model and types of equivalence

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying Model</td>
<td>( Y_{ij} = \log (X_{ij}) )</td>
<td>Yij = aij + eij, aij = ( \mu_{ij} ), and eij = ( \sigma_{ij} )</td>
</tr>
</tbody>
</table>

Types of Equivalence

- Average: population means (\( \mu \))
- Population: population means (\( \mu \))
- Total: all subjects (\( \mu \))
- Within subject: (\( \mu + \sigma \)) if more than two periods
- Between subject (\( \sigma \)) if more than two periods
- Individual: population means (\( \mu \)) if derived as a statistics
- Individual means (\( \mu \)) if interval estimates
- Within subject (\( \sigma \))
- Between subject (\( \sigma \))
- Correlation (\( \rho \))

in these calculations are not extreme for bioavailability data.

Thus, the reassessment of switchability based on average bioequivalence may not exist for highly variable drugs.

A data set from Chow (1990) based on 24 subjects from a two-period crossover study also illustrates this distinction. The means for the reference and test formulations were 82.56 and 80.27, respectively. The ratio of the means is 0.972, sufficiently close to 1.0. The standard deviations were 20.80 and 21.12 respectively. The 90% confidence interval (extrapolated from log scale) is (0.885, 1.007). Thus, according to the present guidelines these bioequivalences are bioequivalent. Figure 2 plots the individual AUCs, with the newly adopted equivalence criteria of 80% and 125% also drawn in. Applying the equivalence criteria to an individual, only 15 (63%) of the 29 subjects fall within the equivalence boundaries. (It is recognized that one might wish to loosen the criteria on individuals to allow for within-subject variation). Is it satisfactory for this percent of subjects to meet the criteria? This becomes a regulatory issue.

Before reviewing the methods, it is important to mention that the methodology is not clear. There is a tendency for the proposed methods to be referred to by the type of distribution in use, when in fact, a method for population bioequivalence may be inappropriate for individual bioequivalence. The objective is to assure switchability and, in some regulatory justifications, predictability.

The FDA's Generic Advisory Committee in February, 1993, agreed with these concepts. However, how switchability is assessed may vary. For example, if the between subject variances for the test and reference formulations are close to 1.0, i.e., if the subject by formulation interaction is zero, then demonstrating what has been referred to as "average" bioequivalence as opposed to reference method one-sided tests, is sufficient. The challenge for the future is recognizing when a particular method can be applied to assure switchability.

Methods for Testing Individual Bioequivalence

To date, the methods for testing individual bioequivalence were developed from two approaches. One approach, referred to as "tolerance interval methods", works with the individual means (or differences) rather than functions of distribution parameters. Anderson and Hauck (1990) proposed TIE, test of individual equivalence values. (See Table 3). It tests whether the proportion of the population for whom the two formulations are equivalent is greater than the regulatory specified minimum proportion in the population for whom the two formulations are equivalent. In a similar fashion, Estivill and Chinnickil (1992) find the tolerance interval that contains 100% of the individual ratios with probability 1.0. They propose both parametric and nonparametric approaches but the latter is noted by the authors to be too conservative. While both methods are easy to understand, to implement and specify equivalence criteria for, both methods are inconsistent. Schall and Leup (1990) propose looking at the differences or the ratio of the probabilities that the difference between test/reference and reference/reference are bounded from above by some meaningful criteria. They proposed a one-sided test that the difference or ratio be sufficiently large to declare individual bioequivalent. Their proposed methods, however, are flawed. For example, the use of the difference as probabilities, the criteria are too easily satisfied and one is working with bounded probabilities. Similarly, the ratio of probabilities is flawed because it acts like a "relative risk" which would require prohibitively large sample sizes and the criterion is not independent of the probabilities.

The second approach to the problem of individual bioequivalence grew out of the simple concept that the difference between test/reference and reference/reference formulations should be close to the difference between repeated tests of the reference. These methods are referred to as "derived statistics" since they are functions of the underlying distribution parameters and do not work directly with the individual ratios of the bioavailability measures. They have the advantage of directly addressing the FDA's interest in having methods which

Table 2: Percent of population satisfying individual bioequivalence

<table>
<thead>
<tr>
<th>Correlation</th>
<th>% of Population within (0.1, 1.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>92</td>
</tr>
<tr>
<td>0.5</td>
<td>82</td>
</tr>
<tr>
<td>0.7</td>
<td>63</td>
</tr>
<tr>
<td>0.5</td>
<td>51</td>
</tr>
</tbody>
</table>

Takes from Anderson and Hauck (1990).

Nia denotes the population mean on the original scale and at the time this paper was written the accepted equivalence interval was (0.1, 1.2).
would explicitly take into account the within subject variability of the reference formulation. With the present methods it is difficult to demonstrate bioequivalence to a highly variable reference formulation.

Schall and Louw (1993) proposed use of the difference in squared error loss between the test/reference and the reference/reference. (See Table 3). The actual test of this quantity being sufficiently small is based on the one-sided bootstrap 95% confidence interval. Steiner (1992), while starting from a different premise than Schall and Louw, proposes a statistic which can be written as the ratio of test/reference squared error loss to that of the reference/reference. Steiner's method, in my opinion, has the advantage of being a ratio which is the most common way people think about bioequivalence. In addition, a series of statistics were proposed by Ekholm and Melander (1991) based on an approximate F distribution. Their most general case can be shown to be a simple scaling of Steiner's statistic. Steiner's method uses a one-sided likelihood confidence interval to demonstrate individual bioequivalence. However, it suffers from being biased and conservative. While the derived statistics do explicitly take into account the within subject variance of the reference formulation and may allow for adjustment of period effects, they are flawed. The major criticisms of the derived statistics are: they are not easily understandable statistics, meaningful equivalence criteria are difficult to specify, and the ability to specify sample sizes is not straightforward, particularly for the method proposed by Schall and Louw. Thus, each of the methods proposed to date suffers from some practical and/or statistical flaw. From a statistical perspective, we are not yet there in assessing individual (or population) equivalence.

**Discussion**

The real question, however, is where should we be? I believe that it is time for the statisticians to step back for awhile in terms of proposing new statistical methods and let our regulatory and pharmacology colleagues decide what they need. Having these proposed statistical methods on the table has helped to focus some of the issues, but we need more direction.

I believe that we need to develop criteria, both practical and statistical, by which we can assess proposed methods. For example, two criteria might be: the ability to set meaningful equivalence criteria and to do sample size (or power) calculations. I stress "practical" as well as statistical criteria, since the sponsors, as well as the reviewers of these studies are not necessarily statistically sophisticated. It is important that, for any given study, what a statistical test is doing be clearly understood. A method that appears to be a "black box" would be less than fully desirable, in my opinion, for either the FDA or the generic drug sponsor. Ideally, we need methods that earn a "sense of the data."

At the same time, we need to see if some of the proposed methods can be improved. Presently, work is being done with respect to the tolerance interval methods as well as Steiner's method. Any new methods that are proposed would need to be statistically validated. Not only do the methods have to be statistically valid, but we need to understand how changes in the parameters of the underlying model shown in Table 2 (or functions of them), affect the methods. The objective is to get a better handle on how to set the equivalence criteria. For example, for Steiner's method, should each of the three components of the numerator below

\[ (\mu_1 - \mu_2)^2 + (\sigma_1^2 + \sigma_2^2)^2 \]

receive equal weighting? Maybe unequal weighting is more desirable, in which case a new test would have to be developed. Work has begun in look at the influence of these parameters using simulations and real data.

Once we have a set of valid methods, we will need to identify the situations under which each method assures

<table>
<thead>
<tr>
<th>Core Statistic</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln((X_{TP}/Y_{TP}))</td>
<td>Distribution-free Bioassay — P Confidence Intervals</td>
</tr>
<tr>
<td>ln((X_{TP}/Y_{TP}))</td>
<td>Parametric and distribution-free tolerance intervals</td>
</tr>
<tr>
<td>2(E[Y_{TP}/E[Y_{TP}]]^2)</td>
<td>Bootstrap confidence intervals</td>
</tr>
<tr>
<td>2(E[Y_{TP}/E[Y_{TP}]]^2)</td>
<td>Likelihood confidence intervals</td>
</tr>
<tr>
<td>2(E[Y_{TP}/E[Y_{TP}]]^2)</td>
<td>Approximate P Confidence Intervals</td>
</tr>
</tbody>
</table>

---

**Table 2: Overview of Methods for Individual Bioequivalence**

- **TOLERANCE INTERVALS**
  - **Andersen & Hauck:** ln(\(X_{TP}/Y_{TP}\))
  - **Eisenhart & Chinchilli:** ln(\(X_{TP}/Y_{TP}\))

- **DERIVED STATISTICS**
  - **Schall & Linn:** 2\(E[Y_{TP}/E[Y_{TP}]]^2\)
  - **Sheiner:** 2\(E[Y_{TP}/E[Y_{TP}]]^2\)

- **Ekholm & Melander:** 2\(E[Y_{TP}/E[Y_{TP}]]^2\)

---

**Figure 2:** Example of average but not individual bioequivalence. Note: AUC values from Chow (1993)
switchability (or preassemblability). The choice of method will depend on bioavailability characteristics of the drug or drug classes. Drugs that are highly variable will most likely require methods similar to those proposed for individual bioequivalence.

These concepts also require more that the standard two-period crossover design. Determining the most appropriate three- or less period crossover designs is an area of current research. The impact of these methods on sample size needs to be reviewed as well.

Progress is being made, but the whole process is likely to 00 many iterations, sometimes like two steps forward and one step back.” I say this, because of the linkage or interdependence between these defining the problem, defining the criteria, validating the methods, and deciding when a given method should be used. The next few years will be exciting as these problems are addressed.

References

Discussion
Carl M. Metzler
Nutwood Associates

For the past 15 years Sharon Armstrong and Walter Hauck, individually and in collaboration, have been major contributors to the development of statistical methods for evaluating bioequivalence. I consider their proposal (1987) of the interval alternative hypothesis to be the best statistical tool on which to base a decision rule for average bioequivalence. The theoretical objections that had no practical relevance; unfortunately it required everyone to completely reverse their thinking about acceptance and rejection regions, power and sample sizes. This may have been most difficult for statisticians. Perhaps for this reason it never received widespread acceptance. In this paper Dr. Anderson reviews some of the current suggestions for changing the definitions of bioequivalence, and some proposed statistical tools to evaluate these new concepts. As she points out, all of the proposed tools still need some fine-tuning. In my opinion the most important part of the paper is the Discussion section. To Sharon’s question “where should we be?” I would add the questions: Where have we been? And where are we now? Twenty years ago Biometrics published my introductory paper on bioequivalence. Although I was listed as the sole author, this publication grew out of, and was indebted to, the discussions of an ASA committee of statisticians considering the then new subject of bioequivalence.

The definition then was that of average bioequivalence, and the accepted rule for declaring bioequivalence was an average bioequivalence of the test drug product that could be shown, with an acceptable risk of a wrong decision, of being not less than 80% nor more than 120% of the average of the standard drug product. With a very few exceptions that is still the rule. Have we not learned anything about pharmacology or biopharmaceutics of drugs in the last 20 years?

In an oral presentation at an FDA hearing in February of 1993 I looked at the top 11 drugs of 1992 (in world-wide dollar volume). For five of these drugs the Physicians Desk Reference gave a starting dose and then recommendations for treating the dose to the needs of the individual. For the other six drugs a single dose was recommended. For three of the drugs something was said about reducing the dose in the presence of reduced kidney function. No other recommendations were given for adjusting dose, and certainly nothing was said about inter- or intra-subject variability in the absorption of the drug products. My point is this: if the top eleven drugs can be given to any patient, whether the little old 85-pound lady down the block or a young 200-pound football player, without regard to differences in volume of distribution, elimination, metabolism, etc., why should we get concerned about relative bioavailability of drug products?

Thus I strongly agree with Sharon that we should step back... and let our regulatory and pharmacology colleagues decide what they need. Why have our regulatory colleagues and their academic counterparts spent so many hours (day, weeks, months?) debating statistical points that have little importance and not addressed the pharmacologic and medical issues? Is it because they perceive the statistical issues to be easier? Is it because as statisticians we have allowed them to think that the statistical issues were really the important ones?

I submit that there are many non-statistical issues that need to be resolved before bioequivalence regulations need be changed: Is there any evidence that the present regulations are not serving well? If so, does this evidence point to the need for regulations that consider populations rather than as covance structures, or does it point to the need for more discrimination among classes of drugs, with different criteria and various levels of risk? Is that plus or minus 20% may not be appropriate for all drugs, a risk level of 5% may not be appropriate for all.

I don't mean to discourage statisticians from developing new methodology, but I do suggest that as statisticians we may be most useful if we can help regulators see which are the most important, and what is the proper role of statistics.

References
Discussion

Murray R. Selwyn, Ph.D.
Statistics Unlimited, Inc.

Introduction

Dr. Anderson has provided us with an interesting and informative article on "individual bioequivalence" and the problem of switchability. I would like to thank Dr. Crain and the Biopharmaceutical Report for the opportunity to comment on this topic.

Anderson motivates the question of assessing individual bioequivalence with the clinical maintenance of a patient who is successfully being treated with one formulation of a drug (the standard), but who is to be switched to a second, alternative formulation (the generic product). She argues that the current approach to bioequivalence testing (called average bioequivalence) is not necessarily satisfactory for a patient in this situation.

In discussing Anderson's paper and the ideas and issues with which it deals, I would like to separate my thoughts into short sections as follows:

(1) General Issues: Is individual bioequivalence a consideration which should be investigated and statistically assessed?
(2) Clinical Issues: If so, what are the clinically relevant questions to be addressed and what would be the impact (if any) on a bioequivalence study design and other related studies?

(3) Statistical Issues: If we perform bioequivalence studies with the assessment of individual bioequivalence as one of their goals, how should we analyze the data from these studies? Further, how would current designs (e.g., the two period crossover) and typical sample sizes be affected?

General Issues

It is tempting for statisticians to pursue slight variations of natural problems to produce interesting statistical research. However, the problems themselves should be important, practical problems in need of pursuit and resolution. In considering "switchability," what are the potential reasons that a difference in the availability of the drug substance (typically measured through blood level AUCs of the active ingredient) is important to the individual patient? Suppose that the patient is already at a high drug level (high AUC) and that switching to the generic will result in an even higher, possibly toxic level. Conversely, suppose that the patient's drug level is toward the low end of the effective range. Will the generics formulation result in an even lower level, perhaps one that will not be efficacious for the patient's disease?

A further possibility that one can imagine is that the substitution of the generic drug for the standard results in a sudden change in the AUC or Cmax and that it is the suddenness of the change and its magnitude which causes a problem for the patient. How likely and clinically relevant are these and other possibilities? What studies have been performed in this regard? How much do we know about the seriousness and prevalence of this perceived problem?

Clinical Issues

I would like to review the purpose of bioequivalence studies: to demonstrate the approximate "blood level" equivalence of a new formulation (e.g., generic) to a standard formulation (e.g., currently approved and marketed product). The idea is that formulations which are equivalent in rate and extent of absorption of the same active ingredient will be therapeutically equivalent. That is, once the active ingredient enter the bloodstream, its subsequent actions do not depend on how it got there (whether generic or approved product).

Instead of performing large and possibly difficult clinical equivalence studies in the target patient population, we conduct smaller bioequivalence studies, usually in healthy, normal volunteers. We assume that the findings from the bioequivalence study are representative and sufficient to make inference about the clinical situation.

In considering switchability for the individual patient, what do we know about the dose that patient? Should the clinically relevant characteristics of that patient be considered in order to verify "switchability"? For example, we are dealing with an elderly patient, in one who is renally impaired, or one whose particular clinical situation is substantially different from the "typical" patient? It seems to me that when we are making clinical decisions about individual patients, we should know something about those patients.

I raise these issues not to be critical of Anderson's and other statisticians' methodological proposals, but only to suggest that the questions they consider transmural mere data analysis and statistical design considerations.

It seems to me that before we undertake the development and implementation of additional and perhaps imposing requirements for bioequivalence studies, particularly within the regulatory context, we should carefully consult with our medical colleagues to identify the types and sources of their concerns (with appropriate literature references of examples) and not rely on mere speculation about the clinical nature of the problem.

Statistical Issues

I generally agree with Anderson and Hauke's (1990) formulation of how to assess acceptable individual bioequivalence, i.e., a high probability that the product's availability ratios will be between suitable limits. I was somewhat surprised, however, that Anderson and Hauke chose to collapse the quantities otherwise available from (AUCs, AUCp, Cmax) into a simple binary outcome indicating inclusion within such an interval (see step 5 of TIER). Instead, given the long history of the linear mixed model for assessing bioequivalence, I would take advantage of this formalism. Specifically, for the two period crossover, define as the model equation for AUC (Grizzle 1965; Selwyn et al. 1981):

\[ Y_{ij} = \mu + S_{i} + F_{j} + E_{ij} \]  (1)

for \( i = 1, \ldots, N; j = 1, 2 \); \( \mu \); is an overall mean, \( S_{i} \) represents the \( i \)-th subject effect, \( F_{j} \) is the \( j \)-th period effect, and \( E_{ij} \) and \( F_{j} \) are the effects for the test and reference formulations, respectively, and \( \mu \) is a term for unexplained variation. We further assume that the \( \mu \) is independent of the \( (S_{i}), (F_{j}), (E_{ij}) \) and impose the usual constraints on the fixed effects in the model (\( P_{1} + P_{2} + F_{1} + F_{2} = 0 \)).

Because the period effect is a nuisance parameter inherent in the crossover design, we further define:

\[ W_{TF} = \mu + S_{i} + F_{j} + E_{ij} \]  (2)

for \( i = 1, \ldots, N \); as unobservable AUCs for test and reference formulations, respectively, for future subjects, \( N+1, \ldots, \).
We can base assessment of individual bioequivalence on quantities derivable from the predictive density of \( P(W_{t1}, W_{t2}) \), given the experimental data, and obtained from

\[
g(W_{t1}, W_{t2}) = \int P(W_{t1}, W_{t2}) g(W_{t1}, W_{t2}) \, dw_{t1} dw_{t2},
\]

where \( g = \{F_1, F_2, G \} \) is the vector of unknown parameters and \( +_1 = (0, 0, 0, 0, 0, 0, 0) \) is the observed data vector.

Given a suitable range (A, B) for relative availability, we would be interested in quantifying such as \( P(W_{t1}, W_{t2}) \) calculated from (3) \( \hat{g}(W_{t1}, W_{t2}) \). Similarly, we could compute the probability that at least 100(1-\% confident patients would have relative availabilities between A and B (e.g., see Geiser 1982, equation 3.17) using (3) as the basis.

Returning to the question of specific patients or specific subsets of the patient population, we could quantify informally some bioavailability for these individuals. For example, if a certain patient \( \{W_{1}, \ldots, W_{7}\} \) was known to have or was suspected of having an apparently low availability (ULC, or one which was near the border of the effective range), we could compute probabilities such as

\[
P(W_{t1}, W_{t2}) \mid A < W_{t1} < W_{t2} < B.
\]

To do this, however, would generally require more knowledge (such as value for A and B) and possibly differently designed bioequivalence studies than those we are now conducting.

I have chosen to base individual bioequivalence assessment on the Bayesian prediction distribution (3) and Bayesian tolerance intervals. One could alternatively consider the sampling distribution of \( W_{t1} \) or \( W_{t2} \) or statistical tolerance intervals for such ratios (see Einbacht and Chinchilli 1994 for a frequencist analysis based on log ratios). Other modeling considerations would include use of log of AUC as AUCs and generalization of approach (1).

I would like to point out one practical consideration arising from Bayesian bioequivalence assessments. It seems, out of necessity, that bioequivalence studies having the capability to make assessments at the individual level will need to be considerably larger than current studies (see Anderson and Haux 1990 for TIBS sample sizes, Einbacht and Chinchilli 1994a regarding those for frequentist tolerance intervals).

**Concluding Remarks**

The paper by Anderson introduces and discusses a potentially important problem, that of assessing individual bioequivalence rather than average bioequivalence. Carefully much of the early development in statistics has focused on problems involving mean values, e.g., the comparison of two means via Student's t-tests and more than two by ANOVA. The realization that mean values only are not sufficient to describe populations seems to a more sophisticated attitude and possibly a better appreciation of the problem at hand.

In bioequivalence, we should ask ourselves to first identify the relevant clinical issues. While assessment of individual bioequivalence may be statistically challenging, is it clinically relevant? If so, let our medical colleagues identify, by providing actual data, the types of questions with which we as statisticians need to cope. I believe that we have the tools to mathematically formulate and solve such technical problems. Bayesian and other inferential methods are certainly available, once the specific issues have been defined.

**Additional References**


**Discussion**

*Jen-pei Liu*

**Biosistats, Berlex Laboratories, Inc.**

The views expressed in this discussion are those of the authors and are not necessarily of Berlex Laboratories, Inc.

I would like to congratulate Dr. Anderson for a comprehensive review of the concept of individual bioequivalence and a much-needed statistical critique of the current methods for evaluating individual bioequivalence.

Bioequivalence studies are surrogate trials for assessing equivalence without actually conducting clinical trials to establish similar effectiveness and safety of the test formulation, because the fundamental bioequivalence assumption implies that bioequivalent formulations are therapeutically equivalent and can be used interchangeably (Miettinen, 1971). Interchangeability for this purpose and equivalent efficacy and safety will be achieved (or not achieved) to the extent that the formulation is prescribed to the patient by the physician. The author used the term "bioequivalence" for this purpose.

On the other hand, many diseases, such as hypertension, are treated and can be controlled by chronic administration of medications. Interchangeability for these purposes, for example, drug concentration levels and safety level will be achieved (or not achieved) to the extent that the formulation is prescribed to the patient by the physician. The author used the term "bioequivalence" for this purpose.

My additional comments are as follows:

*Population Bioequivalence* as pointed out by the author, the assessment of bioequivalence, in fact, is an examination of the closeness between distributions. If the distributions of the pharmacokinetic (PK) response for the test and reference formulations are obtained from a population of subjects, it is referred to as "population bioequivalence" (Chow and Liu, 1992).

If the PK responses follow approximately a normal distribution, it then requires to establish bioequivalence of both average and variances. Current applications of Schattman's two-sided exact test (Schattman, 1983) for evaluating bioequivalence on average completely ignores the importance of the difference in intrasubject variances (Liu, 1990).

For drug products with a narrow therapeutic window, bioequivalence in intrasubject variances poses a serious...
danger regarding drug interchangeability between the reference and test formulations based only on the evidence of bioequivalence for average bioavailability.

The parametric and nonparametric two-one-sided Pitman-Morgan tests procedures were proposed by Liu and Chow (1992) for evaluating bioequivalence of intrasubject variabilities between two formulations. One important feature of the two one-sided bioequivalence test procedures is that the power for average bioequivalence. Consequently, the individual nominal significance levels for average and variability can be pre-determined based on the relative importance of the average to the intrasubject variability with respect to a certain drug class after an overall nominal significance level for the population bioequivalence is chosen. However, power properties and sample size determination for the two one-sided Pitman-Morgan tests procedures still are not clear and require further research.

**Individual Bioequivalence**

The statistical concept of individual bioequivalence is referred to as the comparison of the closeness between the two distributions of the PK responses from the same subject obtained under a repeated administration of test and reference formulations. As indicated by the author, the current methods for evaluating individual bioequivalence, e.g., the test of individual equivalence ratios (TIER) by Anderson and Hausk (1990), tolerance interval approach by Einhas and Chinclini (1991), or use of ratios or differences of squared error loss by Sheiner (1992) and Schall and Lus (1993), are all not satisfactory. However, I would like to point out that Liu and Wang (1992), in their previous work, have shown that, under the normality assumption of the log-scale for the PK responses, the hypothesis of individual bioequivalence by TIER can be reformulated in terms of the average as well as the variance. In essence, the individual bioequivalence using TIER is in fact evaluated through a bioequivalence measure that combines average and intrasubject variability together. On the other hand, the measures proposed by Sheiner (1992) and Schall and Luss (1993) are all not satisfactory. However, in a squared form. The danger of combining averages and variabilities is the possible masking effects. It is quite possible that two tests are claiming equivalence based on a single measure which combines average and variability. The same two formulations, however, might not be bioequivalent with respect to other bioequivalence measures based on average or variability bioequivalence alone. As an example, suppose that the average and intrasubject standard deviation of the AUC(0-24) for the reference formulation are 100 μg x h/ml and 20 μg x h/ml, respectively, while for the test formulation they are 10 μg x h/ml and 2 μg x h/ml, respectively. Hence, the standardized average is 5 for both formulations even though the reference formulation is 10 times more bioavailable than that of the test formulation.

Wellek (1993) suggested another measure for individual bioequivalence. This parameter is the difference between the probability distributions for the same subject, the AUC of the test formulation is less than that of the reference formulation and 1/2. The optimal statistical testing procedure (uniformly most powerful invariant test) exists for this measure. However, Wellek's method suffers from the same drawbacks as the TIER. It is only valid in the absence of period and sequence effects and it is only applicable in the standard 2×2 crossover design. In addition to the drawbacks and flaws pointed out by the author, the exact distribution of any bioequivalence measure based on the functions of squared error loss, either under the null or alternative hypothesis, appears to be unknown. Consequently, the power functions are not clear and no simple methods for sample size determination are available for these methods. Furthermore, the standard 2×2 crossover design, the most commonly employed for bioequivalence studies, is not able to provide necessary information (e.g., an identification problem) required for estimation of some measures based on squared error loss such as the one proposed by Sheiner (1992).

**Comments**

Another disadvantage of the current methods for evaluating individual bioequivalence under the standard 2×2 design is that for each individual subject, they fail to provide a description of the characteristics for the distributions of the PK responses under the two formulations. On the other hand, a repeated 2×2 crossover design (Liu, 1993) provides the multiple PK responses from the repeated administrations of the same formulation within each period. As a result, for each subject, not only the individualized test and reference averages can be obtained but also the individualized intrasubject variabilities for the test and reference formulations can be estimated. A repeated 2×2 crossover design both generates the data for comparing equivalence of the time to reach the steady state and also provides the information for assessing equivalence of pharmacokinetic profiles between the two formulations after the steady state is reached. This design is therefore an ideal design for testing the hypothesis of switchability.

The current trend for evaluating bioequivalence, either population or individual, tends to analyze the data on the logtransformed scale as shown in Table 1 of S. Anderson's paper. Herein, the information of both population and individual bioequivalence is ignored, the average and variability on the original scale of PK responses. Caution is required for the interpretation of the data using the TIER method. Furthermore, this problem (Chow and Liu, 1994). In particular, the lognormal distribution is a very flexible distribution because it can be used for both the average and variability of its corresponding log-transformation. The between-subject, within-subject variabilities, or subject-by-formulation interaction on log-scale do not have the same meaning on the original scale. Therefore, criteria for evaluating bioequivalence on the log-scale should consider their transformation and presentation of results on the original scale.

It is the third and current innovative drugs on the market will lose their patents by the year 2000. Generic drugs will become a crucial component of cost reduction in health care. Consequently, exchangeability between the test and reference formulations should be considered in terms of prescribability and switchability for assessment of bioequivalence. Considering all of this, it is also given the information of prescribability and switchability for generic copies of innovative drugs. Development of meaningful, interpretable, and valid statistical methods for evaluation of prescribability and switchability remains a tremendous challenge for us.

References

Discussion

Daniel J. Holder

Merck Research Laboratories

Dr. Andersen's characteristic clarity and insight always made her articles a pleasure to read. In this article she has given an excellent exposition of current ideas on individual and population bioequivalence. In my comment I would like to focus on her discussion of individual bioequivalence by supplying some details and pointing out some important issues regarding criteria for individual bioequivalence.

I was quite pleased to see that Dr. Andersen dispelled her discussion of the motivation and criteria for individual bioequivalence from her discussion of proposed testing methods. Lack of distinction between criterion and testing method has led to much confusion in the area of bioequivalence. The 75/80 rule is an example of a decision rule which was employed for testing bioequivalence without reference to a specific set of hypotheses or criterion. It is important to clarify what we mean by terms such as individual bioequivalence and state the hypotheses before we propose tests.

Although I do not disagree with the description of individual bioequivalence given in Table 1, I would like to formulate the statistical notion of individual bioequivalence in terms I find easier to understand. Each subject, j, has a mean difference of the log of the bioavailability measure, γ, which, using the notation of Table 1, we call Δ = γj - γ0. The parameter Δ has a distribution, F, with mean Δ and variance σΔ^2 = σ2 + σ2γ - 2σ2γΔ over the population of possible subjects. Note that σγ^2 is essentially the variance component for a subject-by-formulation interaction.

Average bioequivalence has to do with how close θ is to zero. Individual bioequivalence is concerned not only with whether θ is close to zero, but how tightly the δs are clustered around zero. Anderson and Hauck (1993) argue that an important measure of individual bioequivalence is the natural logarithm of the ratio of the distributions for which T and R are bioequivalent, that is the proportion of δs contained in the same interval around 0, say [δ0-δ1, δ0+δ1]. This leads to the definition that T and R are individual bioequivalent if and only if

p(δ | [δ0-δ1, δ0+δ1]) = Φ^p, (1)

where p > p0 and p^* are determined by regulatory and pharmacological considerations. The constant p should be large, maybe >0.75 or greater, since we want to insist that T and R are bioequivalent for a large proportion of the subject population. The definition above is intuitive since it is simply a translation of the notion of within-subject variance for a large proportion of the population into statistical language. An important property of the definition is that it involves only the distribution of the δs and does not involve the within-subject error variance. This is advantageous since the within-subject error variance can be statistically dependent. Another consequence of this definition is that drugs which are highly variable with respect to within-subject variance are measured against precisely the same criteria as drugs which are less variable.

By framing the concept of individual bioequivalence as above, it is clear that the methods that Dr. Anderson describes as "based on derived statistics" in Table 3, test something different than criterion (1), since all of these methods measure the magnitude of the difference between formulations relative to the within-subject variance of the reference formulation. This should be regarded as a reflection of the philosophies of these individual authors towards individual bioequivalence and not as evidence that use of derived statistics necessarily implies rejection of (1) as a useful definition. Holder and Hauck (1993) have shown for bioequivalence intervals that are symmetric about 0 (i.e., -b,b), that by making assumptions about the shape of F, one can choose c <, so that the criterion Δ - Δθ ≤ c(b) implies (1). That individual bioequivalence in terms of (1) can be established by showing that an upper confidence limit to Δ^* - Δθ is less than or equal to c(b).

Note that determination of individual bioequivalence requires information not only on means but subject variance as well. Therefore, in Dr. Anderson points out, the standard two period crossover design is inadequate for determination of individual bioequivalence. The reliance of statisticians on this design has been an impediment to the investigation of individual bioequivalence.

Lastly, I would like to emphasize that individual bioequivalence is a necessary and important concept. Although consensus on a statistical method for assessing individual bioequivalence or even defining a suitable criterion has yet to be reached, it is a natural and important concept. As such, consensus on a statistical method for assessing individual bioequivalence or even defining a suitable criterion has yet to be reached, and one of the most important concerns is the selection of the criteria for bioequivalence.
Biopharmaceutical Report, Winter 1993

Discussion

Vern Chinchilli
Penn State Center for Biostatistics & Epidemiology
Hershey Medical Center

Sharon Anderson should be congratulated for a timely article, which hopefully will spur further research and discussion in the area of biomarkers. As a biostatistician, I will begin my commentary by making some general remarks about individual bioequivalence and follow it with some specific remarks about Sharon Anderson’s article.

My impression is that the typical prescription-drug user does not understand the distinction between population and individual bioequivalence. When asked about their understanding of bioequivalence, such a person is likely to say what bioequivalence means, but it is placed within the context of generic drug substitution. Then the standard reply to this question is something analogous to Sharon Anderson’s definition of switchability, i.e., it does not matter whether the user takes the pioneer-developed drug or a generic version because the delivered dose will be approximately the same. The consumer is quite surprised when he/she learns that the FDA approval process for generic drugs is based on average bioequivalence instead of individual bioequivalence.

Therefore, the issue for debate is whether the FDA should require a company to establish individual bioequivalence of pharmaceutical formulations, because this is what the general public expects. However, many pharmacists feel that the requirements for establishing individual bioequivalence are unnecessarily rigid and that the current FDA process based on average bioequivalence provides adequate efficacy and safety for generic drug substitution. My personal opinion is that this may be true for many drug classes. However, it may not be appropriate for certain drugs with moderate-to-high intrasubject variability especially if the disease is life-threatening. I am ignorant of the field of pharmacoeconomics and there may exist a consensus of studies that may have investigated this issue. If not, a series of epidemiological studies in diseased patients across a broad spectrum of disease categories and drug types may help answer the question of whether the FDA should base approval on the establishment of individual bioequivalence.

Some specific remarks in response to Sharon Anderson’s article are listed below. The distinction between prescriptiveness and test therapy is important, but I would like to add that if a formulation is approved for use because it satisfies average bioequivalence criteria, then the consequences could be less dire for prescriptiveness than switchability. When a physician prescribes a particular drug for a patient, there is usually a treatment phase at the onset, in which the physician monitors the patient’s responses to the drug and adjusts the dose accordingly. On the other hand, the physician may not be aware of the patient switching to a different formulation as a later stage, so that efficacy and/or toxicity within that patient may be altered and go undetected by the physician.

Figure 1 and Table 1 presented by Sharon Anderson succinctly represent the distinctions among average bioequivalence, population bioequivalence, and individual bioequivalence. Although some biostatisticians may quibble over the listing in Table 1 because of slight differences in interpretation, my opinion is that the figure and table illustrate the basic concepts well enough to be understood by pharamaceutical scientists.

Analogous to Sharon Anderson’s example in which average bioequivalence was established but individual bioequivalence was not, Estabert and Chinchilli (1994a) observed the same result in all three of their examples (they use tolerance intervals for investigating individual bioequivalence). It is very likely that this is true for most formulations that have been approved using average bioequivalence criteria. This is not surprising because individual bioequivalence is a more stringent requirement and the sample sizes in current studies are designed for establishing average bioequivalence. Estabert and Chinchilli (1994a) investigate the sample size issue for individual bioequivalence (again based on tolerance intervals) and the standard 24-subject bioequivalence trial may need to be doubled in size in order to have any chance of establishing individual bioequivalence.

Sharon Anderson states that if the subject-by-formulation interaction is null ($\delta_{11} = \delta_{12} = \delta_{21} = \delta_{22} = 0$ in her notation), then the requirement of average bioequivalence may be sufficient. I do not necessarily agree with this conjecture. In this situation individual bioequivalence still could depend upon within-subject variability. For example, in a two-period crossover design in which it is determined that the subject-by-formulation interaction is null, the test and reference formulations could be so highly variable within each subject that the test and reference observations within each subject could be so different from each other that one questions whether the definition of switchability could be satisfied.

Sharon Anderson concludes her article by discussing the various methodologies that have been proposed for determining individual bioequivalence. I agree that much more statistical work is needed because none of the proposed methods is ideal. However, I would not dismiss any of them so quickly. A number of statisticians and pharmaceutical scientists are working on these problems and refinements to some of these procedures may soon appear in the literature and become well-accepted. I do wish she would have elaborated more on some of her criticisms, such as her claim that (1) the tolerance interval approach is inconsistent, and (2) the derived statistics are flawed. Hopefully, she or other researchers will start to investigate these claims about the current methodologies for individual bioequivalence in more detail.

Additional Reference


Rejoinder

Sharon Anderson

When Walter and I defined individual and population bioequivalence in our 1990 paper, we never expected that these concepts would stimulate the amount of research that they have in the relatively short time period of three years. I would like to thank Tulli for the opportunity to discuss these concepts and the Discussions for their comments. In an effort
to be brief, I will highlight what are, I believe, the most salient comments.

While it is exciting to develop new statistical methods, the point made by Carl never been understood. Namely, in order to successfully tackle the concepts of individual and population equivalence we need the input of our regulatory, pharmaceutical, and clinical colleagues. As statisticians we can continue to publish papers but we should not continue to be the driving force. We need to step back.

Murray Selwyn asks for input from our clinical colleagues, specifically to provide "actual data." Some data does exist in the literature regarding treatment failures upon substitution, for example, for psychoactive and anticorrelated drugs. Systematic data, however, does not exist for a variety of reasons. First, the lack of efficacy may not be as obvious as it is for psychoactive or anticoagulant drugs. Second, there is no national reporting system for treatment failures, only for safety. couple this with the fact that both physician and patient may be ignorant of any substitutions. Finally, the use of the standard two-period crossover design has resulted in little available data to explore within-subject variation. The FDA is presently reviewing the data they have, based on 3 and 4 period designs, but this is minimal. This is a situation where we may never have what would be considered "sufficient" data. To overcome this issue, stimulations are being done to better understand where the use of average bioequivalence will be insufficient to assure individual or population bioequivalence, for various classes of drugs.

One issue that need to be addressed was mentioned by Jen-pei Liu. Namely, the interpretation of methods and equivalence criteria based on the log scale. As he point out, "the between-subject, within-subject and subject-by-formulation interaction on the log scale do not have the same meaning as on the original scale. Whether or not these are finally adopted must be interpretable to a variety of audiences.

Daniel Holder raises an interesting argument for development of the within-subject variability: the within-subject variability is study dependent. The approval to market a generic drug may be granted on the basis of the within-subject variability. The within-subject reference variation will, therefore, be determined by the subjects the generic firm enrolls in this one study. The concern is whether the within-subject variability on the reference formulation that is seen in the study, is representative of the variability that would be expected if the reference formulation is generally reproducible. Since ignoring the within-subject variability is contrary to what I understand is derived from the FDS, the more discussion of this approach would be interesting.

Finally, Vem Chinaveli asked for more details concerning the impact of the tolerance interval approach as well as the flaws of the derived statistics. The tolerance interval method (as stated in a 1993 unpublished manuscript) because it did not take into account the within-subject variability suffered the same problem as TIER. The "flaws" in the derived statistics are (1) interpretability, (2) ability to be meaningful criteria; (3) in the case of Shlizer's method, as he states in his paper, it is biased; and (4) in the case of Shall and Liu's method, the RSD ratio is more appropriate as the authors have concluded in their work. As I concluded in the main text, this is an exciting and important area of research because of the changes in our health care system. It is also hoped that the work done for bioequivalence testing will be extended into positive control clinical trials. The future should be interesting!
To the Editor, Biopharmaceutical Report

In response to a book review by Dr. C.M. Mezler on our book entitled "Design and Analysis of Bioavailability and Bioequivalence Studies" which appeared in Biopharmaceutical Report, Vol.10(3), 1992, we wish to make the following comments:

For assessment of bioavailability, estimation is at least as important as hypothesis testing. This is the reason why more than half of Chapter 3 is devoted to the point and interval estimation procedures of formulation, period, and other effects under the additive model. Moreover, the whole of Chapter 6 is devoted to the point and estimation procedures under the multiplicative model (actually lognormal linear model).

On pages 77 and 129, we present our viewpoints on the use of the classical confidence interval for assessment of average bioequivalence. The coverage probability is far the unknown true bioequivalence effect which has nothing to do with the probability of the random interval being within the fixed bioequivalence limits. Our viewpoints on this subject might be different from that of Dr. Mezler, it is correct.

We agree with Dr. Mezler that it is crucial to examine the residuals for the adequacy of the selected model. However, Dr. Mezler's viewpoint is somewhat misleading. In his review, he indicated that "...the authors repeat the common mistake of justifying the log-transformation by the skewness of the observed data (e.g. AUC)... However, in our book, we did not justify the possible use of logarithmic transformation mainly based on the skewness of the observed data. The importance of examining the model residuals has, in fact, been repeatedly emphasized. In particular, there are eight figures of intra-subject or inter-subject residuals in Chapter 8 alone. Our comprehensive and updated summation of statistical design and analysis for bioavailability and bioequivalence studies from which readers will benefit.

Shein-Chung Chow and Jen-pei Liu

Although we have used the awards program to encourage people to present papers in our sessions, we felt that a greater commitment to developing students in the biopharmaceutical area is critical to the future of the section. To provide tangible support for this philosophy, the Executive Committee authorized the awarding of 5 (five) $1,000 student travel grants to attend the annual meeting in Toronto. These grants will be awarded to students contributing papers to the annual meeting. To be eligible, a student must be attending a full time program and be the primary author on the manuscript. We feel that this program, which will be administered on an annual basis, will encourage much more student participation in section activities and will result in a broader interest in biopharmaceutical topics by statistics students.

Dr. Mark was the 1993 program chair, and under his able leadership we sponsored 5 (five) sessions at the Biostatistics Society ENAR meetings in Philadelphia, co-sponsored the midwest Biopharmaceutical Statistics Workshop, and co-sponsored the Applied Statistics Conference in Atlantic City. In addition, the section sponsored 3 (three) invited sessions and 10 contributed paper sessions at the annual meeting in Toronto. Nick Teoh coordinated 8 (eight) round table discussions at these meetings. These contributions to ESAS were very successful in providing an excellent opportunity for discussion of various points of interest to section members.

The section continued to publish all manuscripts presented at the annual meeting and the proceedings of the Biopharmaceutical Section and the most popular and most effective procedures published by the ASA. Chris Czerny continues to direct the effort.

One of the most exciting accomplishments of the section during the year was the request to sponsor the 1995 Winter meetings. Earlier in the year, Liam Gehr suggested that we consider sponsoring the winter meetings. We then approached Sue Hunter about the possibility of the section sponsoring these meetings. In the past, we have co-sponsored sessions at the winter meetings and have had mixed success. Sue was very receptive to our proposal, and we received official approval from Miles Pfaller at our meeting in San Francisco. Ken Kozya is the 1994 Program Chair and will coordinate this as well as the other meetings we sponsor. He will aid in this effort by Joe Heye, the program chair-elect for 1993. The next winter meetings will be held in January of next year in the research triangle area. We will be working very closely with the North Carolina chapter of the ASA to make these a success.

Considerations are underway for publishing the proceedings of the entire meeting and making them available for distribution. The Biopharmaceutical Section has well over 1,000 members, yet its activities are directed by a small group of individuals. Growth and vigor in the section require that different people from different disciplines participate in the complete variety of activities in which the section is involved. I encourage each member of the section to work closely with Dick Northcote, Liz King, Bob Davis, Ken Kozya, or any member of the Executive Committee with suggestions, ideas, or any thoughts you might have which would make the section more effective and responsive to its membership and to the membership of the ASA. We believe that these are the goals of the meetings of the Biopharmaceutical Section. The sections belongs to its members and is there to serve them. To do this properly, it needs participation by all.

The last year has been a very challenging one for me personally and for the section. We have accomplished a lot and for that I thank all the members of the section. We could not have accomplished what we did without the commitment and hard work provided by these people. These include our past chair Cameron Freed, John Schork, who helped in the finance considerations; Nick Teoh who

Biopharmaceutical Report, Winter 1993

Editorial Board

Avital Cnaan
University of Pennsylvania
Editor

Thomas Bradstreet
Merck Research Laboratories
Associate Editor

Eric Sampson & Alison Sarm-Dunyak
American Statistical Association
Layout and Design
Minutes of ASA Biopharmaceutical Section Executive Committee Meeting, August 9, 1993

Attendance:
Bruce Rodda
Bob Starbuck
Bob Davis
Mark Scott
Ken Korus
Dick Binneman
Nick Tesh
Liang Yuh
Mike Boyd
Nguyen Dat
Tuli Crum
Chris Newsat
John Lambert
Akhil Zaidi
Lee Decker ASA Officer
Mike O'Fallon (Chairman of Committee on Meetings)

Bruce Rodda opened the meeting by announcing the winners of the recent Section elections: Lillian Kingsbury for Section Chairperson and Joe Heyne for Program Chair for 1995.

1995 Winter Meeting
Mike O'Fallon and Ken Korus announced that the ASA Committee on Meetings had authorized the Section as the sponsor of the 1995 Winter Meeting. The theme of the meeting will be "Interfacing Statistical Science with Other Scientific Disciplines." Ken noted that we need to work with Lee Decker to select a site for the meeting. Raleigh/Durham seems to be the first choice but concerns were expressed about the uncertain weather in that area in early January. The Biometrics Section was invited to co-sponsor some sessions at the meeting but they appeared somewhat reluctant to participate actively due to their own planning requirements for the 1994 IBC Meeting.

Mike O'Fallon reminded the committee that the Winter Conference will be a major effort requiring the support of many people. Nick Tesh, Joe Heyne, Bob Starbuck and Mike Boyd were appointed to help Ken with the planning.

Assignment: The North Carolina Chapter will appoint academic and FDA Committee members to the Planning Committee.

Lee Decker noted that the Winter Conference can be quite flexible, and the possibility of having two sections over the same weekend is another option and should be set up through Bob Mason of Continuing Education. The outline of the invited session needs to be in place by October 1993 followed by a call for contributed papers in 1994. A session set up by Gladys Reynolds and a Health Economics session by Joe Heyne were suggested. Lee suggested that we consider having companies sponsor the coffee breaks. Four hundred registrants is the break-even point, although non-member invited speakers do not pay registration fees. We should get announcements of the conference in each issue of Annuar News.

Assignment: Ken Korus will apprise the Executive Committee of progress.

1993 Business Meeting
Bruce announced that the Business Meeting would be held Wednesday, August 11 at 6:00 pm. Judy Goldberg of the Biometrics Section had asked that we try not to conflict with that section's meeting.

Assignment: Bob Starbuck will schedule the 1994 meetings to not overlap with Biometrics.

1994 Winter Meetings
The Sections will sponsor sessions organized by Bob Small and Joe Heyne at the 1994 Winter Meeting in Atlanta.

Secretary and Treasurer Report
The March Executive Committee minutes were approved. Bob Davis announced that the Section had a $51,170.49 cash on hand as of June 30. The proposed Sectional 1994 budget was approved. The Section's projected income for 1994 is $24,100. With projected expenses of $23,300.

Assignment: Bob Davis will submit the budget to the ASA office council.

Several suggestions for expenditures were made. Tuli Crum proposed that the Section's projected income ($1,500.00 in 1994) be used to fund a special invited speaker. Lee Decker noted that unless this session were property structured, it would count as one of the Sections invited sessions allocated by ASA.

She suggested that if the Section always pays an honorarium for the Speaker, the proposal might be acceptable. Realistically, the 1995 Annual Meeting is the earliest it could be implemented.

Assignment: Tuli Crum will draft a proposal for an invited speaker. Ken Korus will present it at the January meeting of the Committee on Meetings if the Executive Committee decides to proceed.

The idea of sponsoring students to present papers at the 1994 Annual Meeting was revisited.

Assignment: John Lambert will place announcements and application forms in the Annuar News.

Students should write on their abstract from that they wish to be considered for the Biopharmaceutical Section competition. The student must be the presenter and first author. The abstracts will be due at the ASA by February and the final paper by April 1. Up to five winners will be selected, each receiving a $1,000.00 travel award and a certificate.

Assignment: Ken Korus will forward all abstracts and final papers to Mark Scott, Nick Tesh, Bob Starbuck and John Lambert or their designees for review.

The committee agreed that there will be no change in Section dues for an individual but corporate membership will be reduced from $500.00 to $300.00. Each corporate member will be listed in the masthead of the Biopharmaceutical Report.
Assignment: Bruce Rodda will write an relevant corporate report announcing the solicited membership sale.

1993 Meeting
Mark Scott noted that the room size for the Section-sponsored session continues to be inadequate.
Assignment: Mark Scott will gather attendance information from section chairs and forward to Lee Decker.

1994 Midwest Statistics Conference
Dick Breitman reported that there were 140 registrants and 25 speakers at the 1993 conference. Linning Yuh thanked David Elton for being the emcee speaker at the 1994 meeting, co-chaired by Earl Northrop and Tony Segenti.

Work Groups
Nick Tich reported that the only active work group is the Population Pharmacokinetics group chaired by Linning Yuh, who will become the new coordinator for the work groups. Nick urged the Section to encourage the younger statisticians to get involved with the work groups.

Roundtables
Nick Tich announced that the Section was sponsoring eight roundtables luncheon discussions, all fully subscribed. He suggested that there be more tables, perhaps at other times of the day.

ENAR Meeting (1994)
As noted by Ken Krouse, there is little progress on the Section’s session for the 1994 meeting in Cleveland. He had received no proposals for sessions.

Post Meeting Notes:
With the assistance of Linning Yuh and Nick Tich, we have obtained commitments for 16 invited paper sessions at the 1994 ENAR meeting. They are listed elsewhere in this issue.
We will also be sponsoring a one-day course on “Bootstrapping-based Multiple Testing” given by Peter Westfall and Stanley Young.

Annual Meeting 1994
Ken Krouse reported that five proposals have been received for consideration as invited paper sessions at the 1994 ASA meeting in Toronto.

3. Invited Paper Session on AIDS Clinical Trials, proposed by Neil Palmus (also submitted to Biomatrics, ENAR, WSNAR).
4. Pharmacoeconomic Issues in Clinical Trial Drug Development, proposed by Christopher Burton (also submitted to Biometrics and Economics).
5. Multiple Endpoints in Clinical Trials, proposed by Ji Zhang.

Yen may negotiate with other sections to sponsor some of these, such as pharmacoeconomic session.
Assignment: Bruce Rodda will convene a subcommittee of the Executive Committee to decide the rank-order of these sessions and to make other plans for 1994. (Note: Meeting has been scheduled for November 2 at Bristol-Myers Squibb).

Biopharmaceutical Report, Winter 1993

Training
Mike Boyd updated his proposal for a “Training Workshop for 1993 Biopharmaceutical Statisticians” to be run in Toronto at the 1994 Annual Meeting. The tentative schedule follows:
1:00 pm Xyle of Statistics to Drug Development - Bob Starbuck and Bruce Rodda
2:15 pm Break
2:30 pm FDA Guidelines for “The Format and Content of the Full Justification Clinical and Statistical Report of a Controlled Clinical Study” - Gordon Pledger
3:45 pm Break
4:00 pm Technical Writing - Howard Smith (medical writer at VRA)
5:00 pm Adjourn

The idea behind the workshop is for experienced Section members to interact with the participants. The workshop is not meant to compete with existing PMA courses. Following the Executive Committee meeting and after consultation with ASA, Mike Boyd and Bob Starbuck established that the workshop would be limited to 40 attendees, with a fee of $200.00 for members, $20.00 for students. ASA would take care of the finances, room and meals.
Assignment: Mike Boyd will talk to Bob Mason and Marty freshman at Continuing Education about the proposed plans for the workshop.

In other Continuing Education news, Nguyen Dinh noted that Lillian Kington is still encouraging Joe Heyse to organize a Health Economics course in Toronto.
Assignment: Bruce Rodda and Joe Heyse will decide whether to make a proposal for a Winter course. We will aim for 1995 for this course.

Applied Statistics Conference
Bruce Rodda reported for Earl Paetl that the Atlantic City conference will feature the following sessions:

A. Session 1: "Innovative Applications of Randomized Design", Hufnagel, Beden, David Borenet, Wake Forest
B. Session 2: "Population PK/PK Models: an FDA Perspective", Seyla Mathias, FDA, Tom Ludden, FDA
C. Session 3: "Problems in Multiple Arm Trials", Nancy Glenn, NHBL; Michael Petuch, NASA
D. Two-day Short Course: "Population PK/PK Models, Analysis and Applications", Nick Hollander, BRC; Tom Ludden, FDA; Sten Machado, FDA, Steve Olson, Parkers

ASA Fellows
Bruce announced that John Schmaltz of Uofl was elected Fellow for 1993.
Assignment: All Section members should send any proposed Fellow nominations to Bruce.

FDA Advisory Committees
There was no progress or recommendation for the committees.
Assignment: Bruce will review any progress with Vern Chickil.

Biopharmaceutical Report
Toni Cram reported that the second newsletter of 1993 would be published immediately and the third newsletter is in progress.
Assignments: Bruce Rodda and Bob Starbuck will appoint a new Editor and Associate Editor of the newsletter by August, 1994.

Publications
Chris Jennings reported that as of mid-July over 260 issues of the 1993 Biopharmaceutical Proceedings had been purchased. The number of copies run to date is 770, up 4% from last year. The price per volume (set at $25.00 for pre-publication, $30.00 for members and $45.00 for non-members) has not changed since 1990. The average cost per volume is $38.70. To date we have a surplus from the production of the proceedings of $5,975.00, an increase of 21% from last year and 65% from 1991. The 1993 Proceedings includes 49 papers with an average length of 6.5 pages per paper. The organization for the 1993 Proceedings is underway. We are mailing letters to the ASA office of all invited speakers in Biopharm-sponsored sessions from ENAR in Philadelphia. The Applied Statistics Conference in Atlantic City, and the Midwest Biopharmaceutical Statistics Workshop in Muncie. Lists already exist for invited speakers in Biopharm-sponsored sessions at the joint Meetings. ASA seeks information for paper submission to the authors on these lists on August 1. The papers are due back to ASA by October 15, 1993.

Bruce noted that the Editor of STATS Magazine requested from the Biopharmaceutical Section a contributing editor to solicit one featured article each year.

Assignment: Bruce has named Steve Ruberg as the contributing editor.

Council of Sections
Anna Nevius, filling in for family member Ed, reported that the Council reviewed the issue of certification of statisticians. If certification proceeds, the Section may need to send representatives to an ASA committee on job analysis.

Quantitative Literacy
Assignment: Mike Boyd will submit to the Biopharmaceutical Report a small write-up on the Section's quantitative literacy activities.

Awards
Bruce announced that the Section would present three awards at the Business Meeting for the best papers presented in Biopharmaceutical sections sponsored contributed paper sessions. The recipients are Ron Helms, Joe Heyse, and Carolin and Michael Frey (co-authors).

Membership Activities
John Lambert plans to send a highlighted Operations Manual to each Section member. Alsh Zaiki reported that the membership survey questionnaire has been completed and will be mailed to the Section through the ASA office. The "Hot Pink" Section flyer will be included in the mailing. John Lambert has received an offer from Steve Albrecht of Clin Trials to provide the data entry for the survey at no charge to the Section.

Call for Statisticians
Committee for Recommending Statisticians to FDA Advisory Committees (CRSFC)

The U.S. Food and Drug Administration (FDA) maintains a number of advisory committees which provide expertise and guidance on scientific issues. The committees within the FDA's Center for Drug Evaluation and Research are as follows:

- Anesthetic & Life Support
- Anti-Infective
- Antiviral
- Anesthesia
- Cardiovascular and Renal
- Dermatology
- Drug Abuse
- Endocrinologic and Metabolic
- Fertility and Maternal Health
- Gastrointestinal
- Generic
- Medical Imaging
- Neuropharmacology
- Oncology
- Peripheral and Central Nervous System
- Psychopharmacological
- Pulmonary-Allergy

In addition, there may be vacancies for statisticians on advisory committees for the Center for Biological Evaluation and Research and the Center for Devices and Radiological Health. At least one statistician serves on each of these committees. The Biopharmaceutical Section of the ASA formed a Committee for Recommending Statisticians to FDA Advisory Committees (CRSFC) whose function is to solicit interest from qualified statisticians and make recommendations to the FDA as to appointments.

At this time, the CRSFC needs to update its list of qualified statisticians. The qualifications for a statistician to serve on an FDA advisory committee are:

1. Sound training in statistics;
2. Considerable experience in clinical trials in the appropriate drug specialty;
3. Well-proven abilities and skills to make persuasive arguments on important issues in a committee forum.

Individuals who are interested in serving should be aware that all advisory committee members are subject to government conflict of interest regulations. Therefore, candidates with extensive involvements with regulated industry on matters which will come before their committees frequently will be recused from participation.

Any qualified statistician who would like to be considered for an appointment to an FDA advisory committee should submit a C.V., prior to March 31, 1994, with a cover letter indicating areas of interest to the chair of the CRSFC at the following address: Vernon M. Clinch, III, Ph.D., Chair, CRSFC, Center for Biometrics & Epidemiology, Penn State University, Hershey, PA 17033.
Book Review

Cross-over Trials in Clinical Research. Stephen Senn
John Wiley & Sons Ltd., Chichester, England (1993)

Reviewed by James A. Bognole
Associate Director
Biometrics Research
Merck Research Labs, Rahway, NJ

This book provides a user-friendly, work-sheet approach to the analysis of crossover trials. It is intended for (1) the statistician using cross-over trials who either has to analyze them himself or needs to interpret their analyses by others, and (2) the applied statistician with no particular experience with cross-over. The author presents many helpful tips on crossovers and clinical trials. In addition to a self-admired biased viewpoint on carryover effect,

Chapter 1 provides useful descriptions of crossovers, discusses their advantages and disadvantages, and describes the mathematical level (heuristic arguments with little algebra) and general attitudes taken in the book. Among the attitudes are a bias on pre-testing of factors and model reduction based on data from the current trial, and the recommendation to design cross-over trials to avoid carry-over and augment it in the subsequent analysis. Nearly all considerations described in the book can be achieved with a hand calculator, but SAS® examples are also included. Chapter 2 describes the statistical level assumed: concepts of random variables, estimators and SE's, confidence intervals, significance levels, t-tests, F-tests, ANOVA, and computer package knowledge of linear regression. A review of linear combinations of random variables, associated variance, and their use in estimation completes the chapter.

The AR/BA design with normally distributed data is extensively discussed in Chapter 3. The computational approach taken to assess treatment differences follows by sequence group, this is expanded with a discussion of why carry-over cannot be properly assessed. Incorporation of adjustment for a covariate is skillfully included in the author's straightforward spreadsheet type approach to the analysis. This basic approach ought to be easily understood by the non-statistician. Chapter 4 continues on the analysis of AR/BA design, but centers on data transformations, non-parametric approaches, and binary outcomes. This chapter contains an excellent intuitive description of the sign test, the signed-rank test, and associated confidence intervals which are adjusted for period effect. The author presents a very nice progression through non-parametric rank-based methods and their application.

In Chapter 5, designs with 2 and 3 treatments are assessed in terms of treatment blocked by sequence group using methods described in Chapter 3 for the AR/BA design. Balancing for period effect, but not for carry-over effect is recommended, because the author feels that the latter can only be dealt with by an adequate washout period. Non-parametric analyses for these designs are discussed in Chapter 6, included are Cochran-Mantel-Haenszel statistics, Hodges-Lehman estimators, and Mantel-Haenszel statistics for stratified binary outcomes. As with previous chapters, these two chapters are also skillfully presented and should be easily understandable by a mathematically sophisticated non-statistician. However, this reviewer wishes the author had cautioned the practitioner employing the recommended pairwise treatment comparison approaches of the multiplicity and correlation issues.

"Special designs" are discussed in Chapter 7; they include factorial, incomplete block, n=2 trials, and bioequivalence studies. The discussion of single-patient trials is very informative, worthwhile, and insightful regarding the assumed probability of type I error. The author highlights well the problem with relying on n=2 trials, pointing out the proper interpretation based on randomization tests. Chapter 8 describes some basic graphical and tabular presentations to which non-statisticians should pay close attention; however, there is nothing new here for the statistician. This reviewer was surprised at the omission of the usual response-over-period plots of treatments-by-group means, they were not even referenced.

The following design issues were addressed in Chapter 9: parallel versus cross-over, use of washout to eliminate carry-over, a cursory look at choices of sequences, sample-size considerations, and a discussion of missing data. This chapter contains a very useful and clear description of the difference between intra- and inter- subject variance, and their relationship to the correlation coefficient. Use of a covariate is also skillfully tied to this discussion. A handy SAS® program for computing power is included.

Having removed it to the throughout the book, the author has done a solid job on analyzing for carry-over effect in Chapter 10.

(1) "It implies then the investigator can design a trial which eliminates it." But this reviewer feels that such designs may be impractically long to carry out, especially for large studies.

(2) "It is implausible given the unearned pharmacodynamic (modeled by sum of exponentials) and pharmacokinetics (Hill equation) theory." Persuasive mathematical arguments are presented, but psychological carry-over is not addressed, and the author has not been given the opportunity to comment on the needed carryover.

(3) "The carryover which incorporates it self-contracting." Using a Williams square for the 4th crossover involving treatments P, A, P, B, and AB, the author states that who failed to notice the interaction is accounted for, the interaction of strain effect of A with carryover of B is not. He feels that these should be equally important, therefore, the design is self-contracting. This reviewer feels that the importance issue is subjective, not necessarily self-contracting.

(4) The estimators based on it are inefficient (compared to unadjusted-for-carryover estimates). But the author does not say that the latter are biased. Simple carryover is present, an event the author contends never occurs.

(5) The designs associated with it are not necessarily better than others." This depends on the assumptions about the true underlying conditions.

These arguments are clearly and persuasively presented with counter arguments to many issues, but not to those raised above by this reviewer. Following these advances is a further 10 perspectives to analyzing for carryover effects, "if the study is not persuaded by these arguments." The analysis is presented in a nicely intuitive and understandable fashion, involving the derivation of weights for linear combinations of calls to estimate pairwise treatment differences adjusted for carryover.
The book contains many clearly presented, useful tips about clinical trials, among which are the following:

- clinical trials don’t really represent random samples from populations, rather, results are interpreted/assumed to represent that which is expected;
- concentrate on estimation, and avoid over-emphasis on “p<0.05”;
- use of baselines and covariate analysis can improve precision;
- guidance on interpreting interaction, assessing p-values for hypothesis generation in borderline cases;
- constructing variances to justify assumptions and signal their effect by treatment; and
- the t-statistic is generally robust except in extreme cases.

In general, regarding style, there is a gross lack of commas throughout the book, some of which makes reading sentences difficult. The prevalence of typos and grammatical errors indicated a lack of professional proof-reading. The choice of fonts for SAS® example program statements and output match too closely those used for regular text. This author’s pet peeve was the incorrect citation of the reviewer’s bible for design as “Cox (1957),” instead of “Cochran & Cox (1957),” maybe “Cox (1958)” was intended.

Other peculiar problems worth noting are the following:

The statement on page 80 about excluding, as uninformative, patients who are missing 1 of 2 periods in a 2x2 crossover, should have included a caution about informative censoring. The author rightly does not approve of power calculations as an aid to interpret non-significant results, and suggests looking at likelihood under the alternative, but states that this is beyond the scope of the book. This reviewer would have liked to see some reference to the use of confidence intervals, which are not beyond the scope.

With the above reservations, this book is recommended as a guide to a straight-forward approach to the analysis of crossover trials, or as a reference in support of ignoring carryover effects in the choice of treatment sequences in crossover design and their subsequent analysis, if anyone would want to take that approach.

Book Review

Multiple Comparisons, Selection, and Applications in Biometry
Edited by Fred M. Hoppe
Marcel Dekker, 1993
Reviewed by Dror M. Bum
Rhone-Poulenc Rorer

This book contains a collection of papers presented at a Symposium on Biostatistics held at McMaster University, Hamilton, Canada in honor of Charles Dunnett.

As its title indicates, the book is divided into three sections: Multiple Comparisons Selection, and Biometry & Design.

The first chapter is a most interesting conversation with Professor Dunnett about his military, industrial, and academic

1994 International Research Conference on Lifetime Data Models in Reliability and Survival Analysis
Boston 15-17 June 1994

This research conference will be held in Boston, MA on June 15-17, 1994. The conference is co-sponsored by the Statistics Department of Harvard University, the Boston Chapter of the American Statistical Association, Kluwer Academic Publishers and Where, Inc. The conference has been organized to bring together specialists and practitioners from all fields who have an interest in lifetime data analysis, including biomathematics, the natural sciences, engineering, business, and the social sciences. The following list indicates the wide range of conference topics that are of interest:

- Accelerated failure rate models,
- Bayesian lifetime models,
- Estimating and translating, classes of lifetime distributions,
- Competing risk models, counting processes, degradation processes, inequalities for reliability bounds, maintenance policies and replacement models, meta-analysis of life data, models for multiple states, models for nonconformance, multivariate failure models, network reliability models, nonparametric estimation of survival functions, nonproportional hazards models, parametric estimation and prediction inference, parametric regression models, proportional hazards models, quality-of-life models, random effects models, rank tests for comparing lifetime distributions, and surrogate marker processes.

Registration fee is $150
Student registration fee is $30
Thursday night dinner ticket(s) $35 x (# of tickets)

The registration fee covers one mixer ticket, meeting schedule, abstract booklet and a copy of the Conference Proceedings, which will be published after the conference. The Thursday night dinner ticket is not included.

The student registration fee is for full-time students only. This reduced fee includes the meeting schedule and abstract booklet, but not the mixer ticket, Proceedings, or dinner ticket.

Contact
Dr. Mei-Ling Ting Lee for registration forms at:
Chemistry Laboratory, Harvard Medical School
180 Longwood Avenue, Boston, MA 02115
email: stree@geus.mc.harvard.edu
ENAR Meeting Schedule

Monday, April 11, 1994

3:30 - 5:15 pm
19. Sequential Designs in Clinical Trials
Sponsor: ASA, Biopharmaceutical Section
Organizer: C.F. Jeff Wu, University of Michigan
Chair: Dan Friedman, Schering-Plough Research Institute

3:30 Two-stage design of bioassays in pharmaceutical studies: An overview. C.F. Jeff Wu, University of Michigan
4:00 Practical experience using a two-stage study design. I. Molina, D. Roth, Schering-Plough Research Institute
4:30 Some recent experiences in two-arm and multi-arm adaptive clinical trials. Roy N. Tamura, Eli Lilly and Company Incorporated

5:00 Floor discussion

Tuesday, April 12, 1994

9:30 - 10:15 am
25. Multiple Endpoints in Clinical Trials

Biopharmaceutical Report, Summer 1993

Sponsor: ASA, Biopharmaceutical Section
Organizer/Chair: Ji Zhang, Merck Research Laboratories

8:30 Testing the hypotheses that matters. Thomas G. Carey, Ji Zhang, Merck Research Laboratories
8:55 Improving some methods for multiple endpoints. Dei-in Tang, Nanci S. Kline Institute for Psychiatric Research
9:20 Using multiple endpoints to answer clinically relevant questions. David S. Salsburg, Phoenix Central Research Institute
9:45 Discussion. Abdul J. Sankoh, FDA and Bruce W. Fasnacht, Cornell University

10:05 Floor discussion

Tuesday, April 12, 1994

3:30 - 5:15 pm
38. FDA-industry Session: Analysis of Longitudinal Data in CD Trials
Sponsor: ASA, Biopharmaceutical Section
Organizers: Sivani D. Dzaleu, FDA, Nick W. Troch, Schering-Plough
Chair: Samuel M. Heit, Schering-Plough

3:30 Overview of methodology for longitudinal data analysis within the regulatory context. Manuel Talamanc, FDA
3:50 Application of longitudinal data analysis methodology to account for dropouts. Corin Daring Sanders, Ohio Vet. Ho, Lindal Env. Schering-Plough Research Institute
4:10 Longitudinal data analysis for heartburn trials. Mohamed F. Bouic, Abdul J. Sankoh, Ferrin, D. Horak, Sivani D. Dzaleu, FDA
4:36 Application of generalized estimating equations (GEE) to the analysis of repeated ordinal data. David Shaw, Bristol-Myers Squibb (Brussels), Mike Kerward, Rothamstead Experimental Station, Luc Biemans, Jean Bender, Bristol-Myers Squibb (Brussels)
5:00 Discussion. Bruce E. Rodda, Bristol-Myers Squibb, H.M. James, EPA

Wednesday, April 13, 1994

9:30 - 10:15 am
44. Methods for Handling Dropout in Comparative Efficacy Trials
Sponsor: ASA, Biopharmaceutical Section
Organizer/Chair: Chuck Davis, University of Iowa

Wednesday, April 13, 1994

10:30 - 12:15 pm
46. Analyzing Survived Data: Recent Advances in Methodology
Sponsor: ASA, Biopharmaceutical Section
Organizer/Chair: Robert Strawderman, University of Michigan

10:30 Unbiased estimation in the Cox proportional hazards model with missing covariate data. Marian Pugh, Stuart Lipetz, David Hafting, Harvard School of Public Health, Dana Farber Cancer Institute
10:55 Evaluating surrogate markers of clinical outcome when measured with error. Urania G. Dafnis, Susan A. Thall, Harvard School of Public Health
11:20 Semiparametric analysis of general hazard-based survival models. D.Y. Lin, S. Yang, University of Washington
11:45 Discussion. Robert Wolfe, University of Michigan

12:00 Floor discussion
ASA Biopharmaceutical Section Executive Committee—1994

Immediate Past Chair
Bruce E. Bodea, Ph.D.
Vice President, Biostatistics & Data Management
Bristol-Myers Squibb
Pharmaceutical Research Inst.
PO Box 4000
Princeton, NJ 08543-4000
(609) 252-5776
(609) 252-6060 fax

Chair
Robert R. Starchuck, Ph.D.
Assistant Vice-President
Clinical Biostatistics and Data Management
Wyeth-Ayerst Research
145 King of Prussia Road, B-2
Radnor, PA 19087-4517
(215) 341-2070
(215) 341-2092 fax

Chair Elect
Lillian Kingbury, Ph.D.
Director of Biostatistics
Bio-Pharm Clinical Services
512 Township Line Road
Blue Bell, PA 19422
(215) 283-0770
(215) 283-0733 fax

Secretary - Treasurer
Robert L. Davis, Ph.D.
Director, Biostatistics
The Astra/Merck Group of
Merck & Co., Inc.
725 Chesterbrook Blvd.
Wayne, PA 19087-5677
(215) 695-1070
(215) 889-1288 fax

Program Chair
Kenneth J. Kazuy, Ph.D.
Dept. Head, Clinical Statistics
American Cyanamid Co.
Lederle Laboratories
Building 60, Room 203
North Middletown Road
Pearl River, NY 10965
(914) 732-3827
(914) 732-5518 fax

Program Chair Elect
Joseph F. Hoeye, Ph.D.
Director, Health Economics Statistics
Merck Research Laboratories, IL-32
Sweezytown Pkwy
West Point, PA 19486
(610) 397-2334
(610) 397-2931 fax

Executive Committee Member
Michael A. Boyd, Ph.D.
Director, Biostatistical Services
Pharmaceutical Research Associates
240 Old Ivy Road
Charlottesville, VA 22903-4826
(804) 977-2772
(804) 293-7084 fax

Executive Committee Member
Charles S. Davis, Ph.D.
Associate Professor
Dept. of Preventive Medicine
University of Iowa
2837 Steindler Bldg.
Iowa City, IA 52242-1176
(319) 335-9625
(319) 335-9200 fax

Executive Committee Member
Nguyen V. Dat, Ph.D.
Vice President, Project Operations
Biometrics Research Institute
1300 North 17th Street
Arlington, VA 22209
(703) 276-0400
(703) 243-9746 fax

Executive Committee Member
Sally A. Greenberg, Ph.D.
Senior Statistical Research Scientist
Syms Research Development
3601 Hillview Avenue, A6-100
Palo Alto, CA 94304-1397
(415) 813-4888
(415) 813-4894 fax

Executive Committee Member
C.S. Wayne Wing, Ph.D.
Statistician Manager
Shering Laboratories
2000 Gallowing Hill Road
Kenilworth, NJ 07033
(201) 298-5457
(201) 298-5720 fax

Executive Committee Member
Liangying Yu
Senior Manager
Biometrics Department
Walter-Lambert/Phire-Davis
2800 Plymouth Road
Ann Arbor, MI 48105-2430
(313) 966-7894
(313) 966-1796 fax