Biopharmaceutical Section

American Statistical Association

Biopharmaceutical Report

Volume 2, No. 1 Summer 1993

Chair: *Bruce E. Rodda, Ph.D.* **Editor:** *Avital Cnaan, Ph.D.*

Essential Efficacy Data Analysis

Gregory G. Enas, Todd M. Sanger, William J. Huster
Lilly Corporate Center, Eli Lilly and Company, Indianapolis, IN 46285

Introduction

We would like to offer some ideas regarding efficacy analyses from the perspective of a more thorough understanding of the efficacy findings. These ideas, incomplete as they may be, are presented here so that we may begin to produce reports and publications of efficacy data that will allow a reasonable skeptic the opportunity to draw conclusions about the data as expeditiously as possible. Since our ideas are incomplete, we invite the discussants and other readers to contribute to this topic as well.

At least two things have become apparent to us regarding the analysis of efficacy data over the past ten years. First, over a wide variety of important clinical conditions, there are many features common to the conduct of clinical trials upon which one can standardize analyses. Second, clinical trialists (including statisticians, physicians, and others) often do not have a thorough understanding of the efficacy findings. For whatever the reason, they do not dig deep into understanding the trial process that generated the data and do not rigorously scrutinize the "robustness" of at least the primary findings.

These issues certainly affect one another, i.e., one usually does not have the time to perform extensive *de novo* analyses while simultaneously scrutinizing the trial process. However, both issues need to be addressed if we want to expedite the approval and acceptance of new treatments by regulators and medical practitioners. We would like to help create a working environment where statisticians have time to spend on important aspects of a clinical drug development program, such as:

1. understanding the medical condition,
2. strategic planning of drug development,
3. planning the design of protocols,
4. specifying what data to collect when and how,
5. negotiating with management and colleagues involved with the trial,
6. understanding trial results,
7. generating new hypotheses,
8. communicating with consultants and regulatory officials, etc.

All of these activities are threatened when there is too much data analysis and reporting to do in too little time. We all recognize that one can perform a superb job of analysis yet still produce a less than satisfactory trial report (and not impact medical practice) if items such as those in the list above are not given close attention.

Contents

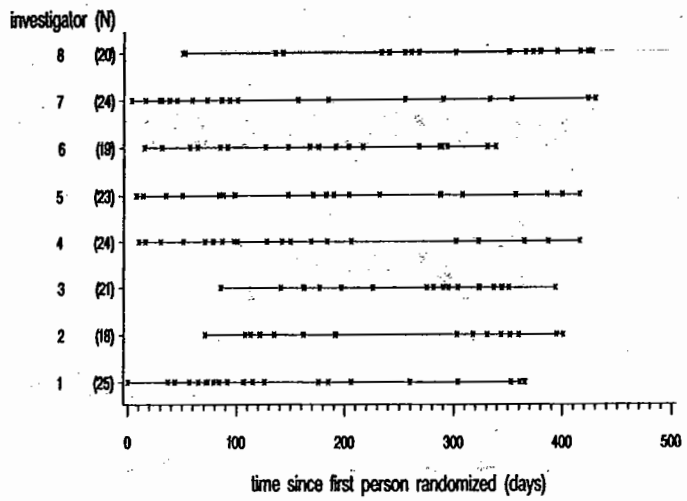


Figure 1: Time course of patient randomization starting when the first patient was randomized by investigator.

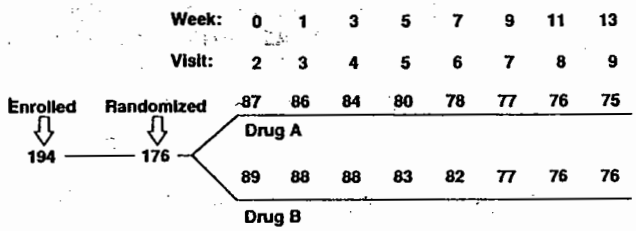


Figure 2: Patient Accounting by treatment and visit/week of the study.

We have heard discussion recently about "standardized" analyses. Much of the activity in this regard has been focused on safety data. The benefits of such safety analyses have been or will soon be realized by investigators, data analysts, reviewers, and medical practitioners alike. Standardization may be a means to help achieve a greater understanding. The ideas we would like to offer regarding efficacy analyses were motivated, not so much from a "standardization" perspective, but from the perspective of a more thorough understanding of the efficacy findings. Fortunately however, many of the data displays and analyses to be presented may be automated rather easily, which will help to free up the statistician to contribute to other important tasks mentioned above.

Philosophy of Analysis

Clinical drug development trials should always state as explicitly as possible what the objective of the trial is. Associated with this is the design of the trial and the analysis by which one measures how well the trial met its stated objective(s). The more nebulous the objective, the less explicit the protocol-defined analysis (e.g, early trials). The closer one moves to ultimate approval and marketing of a new drug, the clearer the objectives (and stated analyses) must be. Our philosophy is that no matter how early or late in drug development a trial is conducted, one should do as much legitimate analysis as is possible. Early on, these analyses will generate hypotheses to be tested in further trials. Later on, these same analyses can be used to measure the robustness of the primary analysis as stated in the protocol. While the objectives of Phase II and Phase III trials may differ, there is no distinction between Phase II and Phase III analyses, per se. It is of critical importance to state, as clearly as possible, the prior information to date, the objective of a new trial, the analyses to be undertaken, and by the time you are close to marketing, the primary analysis to hang your hat on (which is, as you will see, one of the many that should be conducted for any trial).

Obviously, time limits the amount of digging and analysis one can do. Thus, there is a need to come to some understanding of the most helpful analyses for a large number of trials and then automate such analyses. Such analyses could be readily conducted throughout and at the end of a trial. Such analyses should be useful to a reasonable skeptic who either wants to know in what direction(s) the drug development program should go or to answer the question of efficacy in any particular patient population.

Analyses should illuminate trial issues such as treatment balance on important risk factors, relationship to baseline variables to outcome, time trends, impact of missing data as well as the reasons for missing data, subset differences and heterogeneity in outcome due to differences in patient populations, treatment by investigator interactions, enrollment patterns of investigators, and the contribution to the results from high

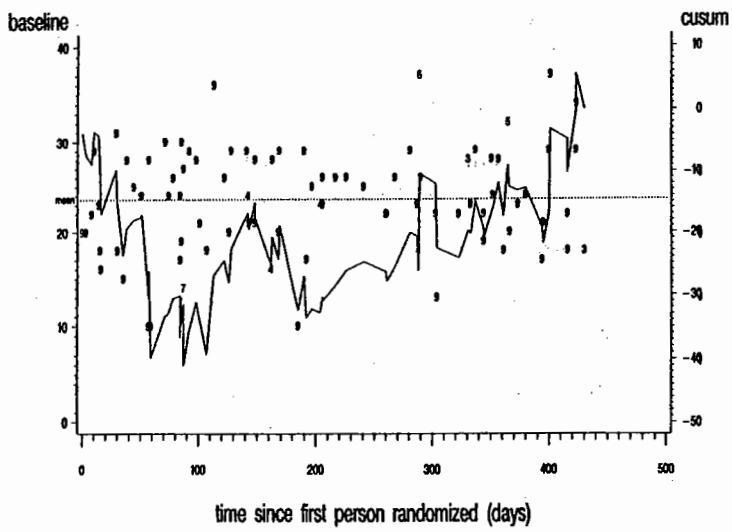


Figure 3: Primary efficacy variable at baseline for treatment A versus time of randomization starting when the first patient was randomized. Cusum is calculated as the cumulative change from the treatment mean. Plot symbol is the last visit attended.

enrolling sites versus sites enrolling fewer patients. If interim analyses are performed, then any impact of such analyses on further conduct of the trial should be explored. Also, do the (many) legitimate analyses support or contradict each other?

We now present an example that illustrates just a few of these points. We trust this will instigate further discussion on other analyses that may be essential to maximize the potential of a trial to influence and promulgate benefits to future patients.

Example

For purposes of illustration, we concentrate only on efficacy data measured on at least an interval scale in a parallel study. Straightforward modification of the methods presented in this paper can be made to accommodate discrete and ordinal data. Crossover studies would also require modification to these summarizations.

Description of the Dataset

The data used for illustration throughout this paper were collected in a 14-week, double-blind, randomized, parallel study of Treatment A versus Treatment B in patients from 8 different investigators. There was a 1-week single-blind placebo lead-in period during which 194 patients were screened for compliance. At visit 2, 87 patients were randomized to Treatment A and 89 patients were randomized to Treatment B. These patients were treated for 13 weeks. Visit 3 occurred 1 week later and Visits 4 through 9 were scheduled at 2-week intervals. The primary outcome measure was the test score, which ranged from 0 to 40. Declining scores over the treatment period indicated an improvement in the patient's condition.

Study Execution

This section describes some basic trial summarizations, analyses, and pictorial representations that help one understand and interpret the efficacy analyses from a clinical trial. These displays are supportive of the analysis of the primary treatment comparisons, which will be discussed later.

For starters, as one way of characterizing the patient sample under investigation, one would like to know how and when patients entered into the trial. Each tic mark in Figure 1 represents when a patient was randomized to a trial therapy for each of the investigators in this multicenter trial. The horizontal axis is the time of patient randomization relative to the time since the first patient was randomized. Any "significant" events that occurred throughout the trial (e.g. interim analyses, protocol amendments, etc.) may be noted on the horizontal axis. This figure indicates that patient entry was fairly uniform throughout the study.

Next, one would like to know about patient "flow" through the trial. Figure 2 gives an example of how a simple graphic might look that describes this process for a parallel study. Sample sizes over time would be described within important strata (such as investigative site or prognostic variables) as well as combined over the strata. The time points in this depiction would be chosen to be representative of the entire trial. From Figure 2, one can see that greater than 85% of the patients completed the trial and that most dropouts occurred during the first 9 weeks of the study.

Time Trends

It may also be important to know if the patient's baseline medical condition varied over time as they entered the study. Time trends and interactions with treatment may be important determinants in the assessment of efficacy. One way to show time trends is displayed in Figure 3 for treatment group A. This is a scatterplot of the pre-randomization, baseline value versus the time of randomization relative to the time the first patient was randomized. Each patient in these scatterplots is identified with an informative symbol. In this case, it is the last visit the patient was observed. Throughout this paper, the symbols used for plotting are informative and a variety of meanings can be

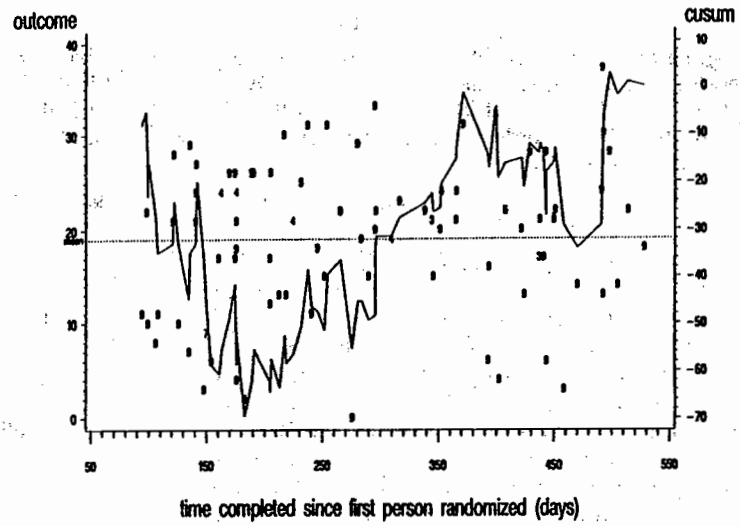


Figure 4A: Primary efficacy variable at endpoint (last patient value) for treatment A versus time of completion relative to when the first patient was randomized. Cusum is calculated as the cumulative change from the treatment mean. Plot symbol is the last visit attended.

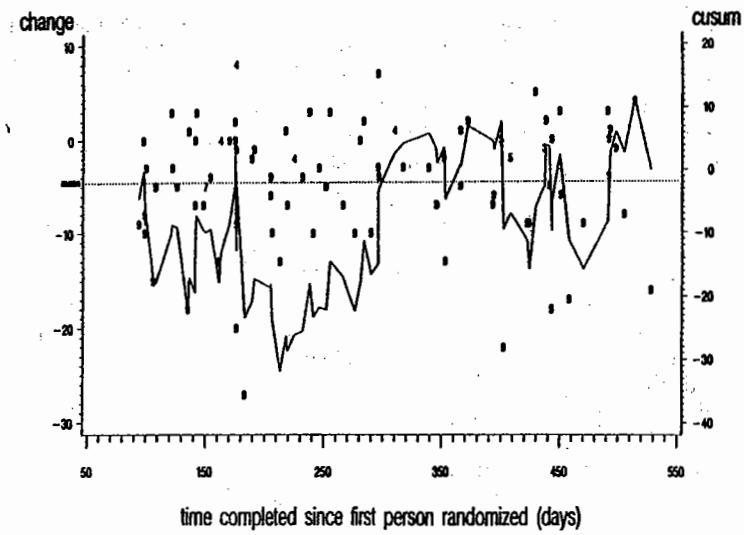


Figure 4B: Change in the primary efficacy variable at endpoint (last patient value) from baseline of treatment A versus time of completion relative to when the first patient was randomized. Cusum is calculated as the cumulative change from the treatment mean. Plot symbol is the last visit attended.

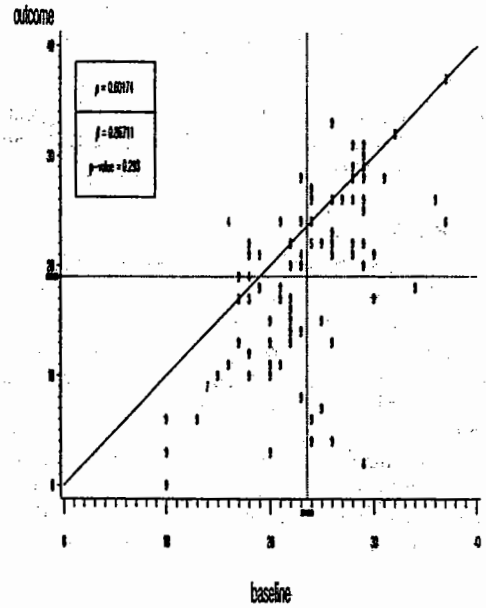


Figure 5: Relationship between the primary efficacy variable at baseline and at endpoint (last patient value) for Treatment A. Plot symbol is the last visit attended and the line of identity is also plotted. The statistic ρ is the sample correlation, ß is the ordinary least squares estimate of the slope, and the p-value tests whether ß is equal to 1.

denoted (e.g. gender, age group, presence of adverse event, etc.).

Overlaid on the points is the cusum calculated as the cumulative difference from the overall mean for treatment A (see O'Brien et al (1989) or Altman and Royston (1988). That is, if $X_{(i)}$ is the baseline value for patient i ordered by time of entry into the study and X_a is the mean for treatment A, then the cusum at time t is just the sum from i=1 to t of $(X_{(i)}-X_a)$. From this figure, it appears that patients who were randomized very early in the study may have had lower baseline scores than those who were randomized thereafter. This figure could also be drawn combining all treatment groups. Other important prognostic variables could be shown as well in this manner. Again, significant events that may have affected the types of patients enrolled can be noted on the horizontal axis.

Similarly, a summary of the efficacy outcome over time may be shown as in Figures 4A, for the raw last-visit-carried-forward (LVCF) value, and Figure 4B, for the change from baseline to LVCF value. The horizontal axis is the time of last patient visit relative to the time the first patient was randomized. The efficacy outcome may be the LVCF value or some other predefined variable of interest (e.g., the last or some intermediate visit, a summary of many visits, etc.). Again, it is helpful to denote each patient in these scatterplots with an informative symbol. One of our favorite symbols is the last visit the patient was observed; this is quite helpful in determining the effect of early dropouts on conclusions. These figures indicate that patients randomized very early in the trial may have performed better than those randomized late. Significant events that occurred during the trial, such as an interim analysis or protocol amendment, could be noted on the plot.

Treatment Comparisons

The primary analysis, especially in a confirmatory trial, should be outlined in the protocol. However, several additional analyses of the primary efficacy variables should be undertaken to investigate the "robustness" of the results elicited by the primary analysis. The figures shown above aid in verifying robustness by providing some support for the assumptions of the models used in subsequent analyses.

An important relationship that should also be investigated is the relationship between the primary efficacy variable at baseline and after treatment. Figure 5 shows how a scatterplot of the baseline versus the endpoint values may be constructed for treatment A. This plot can be supplemented by summary statistics on the axes (that means are used here) as well as by the line of identity to help identify a shift. As in Figures 4A and B, the primary efficacy variable after treatment may be the LVCF value or some other variable of interest and symbols, such as the last visit a patient was observed, may be used to denote each patient in the picture. These plots may be shown for each treatment or pooled over all treatments (with treatment as the symbol). From Figure 5, one can see that most of the patients on

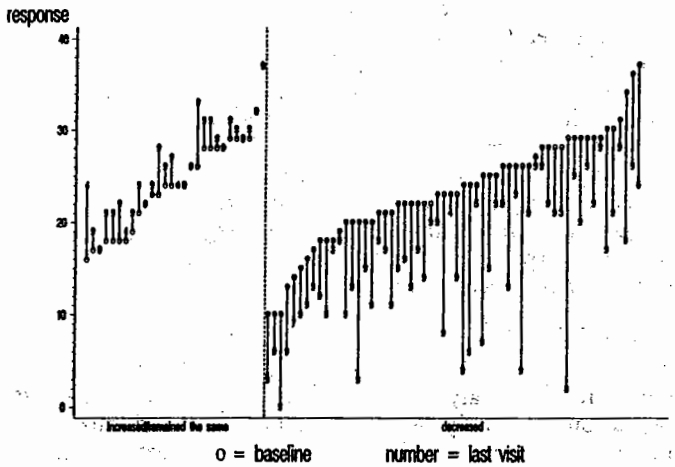


Figure 6A: Paired data plot of the primary efficacy variable at baseline and at endpoint (last patient value) for Treatment A.

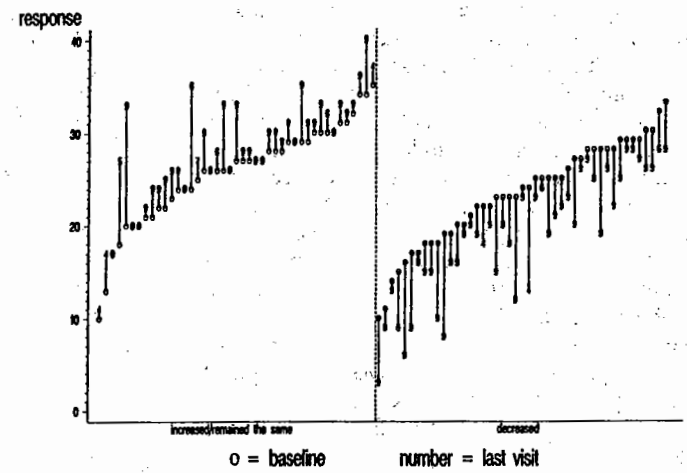


Figure 6B: Paired data plot of the primary efficacy variable at baseline and at endpoint (last patient value) for Treatment B.

Table 1: Treatment comparison p-values from the primary and supplementary analyses of the primary efficacy variable

		Type of Model			
		ANOVA[1]		ANCOVA[1]	
Type of Analysis	Outcome Measure	Original	Ranked	Original	Ranked
Endpoint (LVCF[2])	Raw	<.001	<.001	<.001	<.001
	Change from BL[3]	<.001	<.001	<.001	<.001
Repeated (Linear)	Raw	<.001	.001	<.001	.001
	Change from BL[3]	<.001	<.001	<.001	.018
Visit 9 (Completers)	Raw	.001	.002	<.001	<.001
	Change from BL[3]	<.001	.003	<.001	.004

1 ANOVA and ANCOVA models contained the terms treatment, investigator and treatment by investigator interaction as fixed effects.

2 LVCF = Last visit carried forward

3 BL = baseline

treatment A showed improvement in the primary efficacy variable compared to baseline since the majority of the points fall below the 45 degree reference line. In addition, the regression line fit to this data has a slope that is not significantly less than one, indicating that simple change scores may be analyzed in a straightforward manner.

McNeil (1992) has suggested several alternatives to the typical "bundle of lines" approach for graphing paired data (e.g., baseline to endpoint) which are more informative. As an example, Figures 6A and 6B plot the baseline and endpoint values for each patient on treatment A and B, respectively. Patients whose scores increased or stayed the same during the study are presented first, sorted by baseline value. These patients are followed by the patients whose values decreased, sorted by baseline value. Open circles denote baseline values and are connected to endpoint values portrayed by the number of the last visit attended. From this figure, it is obvious that the majority of patients experienced a reduction in the test score while on treatment A but this does not hold for treatment B. Also, the magnitude of response does not appear to be related to the baseline value in either treatment. In a similar vein, some ideas for visualizing survival and time-to-event data are given in Enas and Rockhold (1987) and Goldman (1992).

To provide support for the robustness of the primary analysis results, one would analyze the efficacy outcome itself as well as the change (and/or percent change) from baseline using several models (parametric versus non-parametric). Table 1 provides the results of several analyses of the primary treatment comparison that are quite helpful in ascertaining the fit of the statistical model under consideration as well as the effect of baseline on outcome. The primary analysis in the protocol was an analysis of the LVCF value using ANOVA with investigator and treatment as fixed effects with a term for the investigator by treatment interaction. As an alternative model, ANCOVA can be used with the response as the dependent variable and the baseline value as the covariate. To evaluate the fit of the model, we like to analyze the data in the original scale as well as some (rank) transformation of the data. In a repeated measures situation over time, one can analyze not only a univariate summary statistic of the outcome (such as the LVCF or endpoint analysis) but can also conduct repeated measures analyses and visit-wise analyses (analyses of those patients observed at each visit). In Table 1, the results reported for the repeated measures analysis are from the test of the linear component of the treatment by visit interaction term. See Crowder and Hand (1990) for a discussion of the assumptions of and restrictions on the repeated measures models. Because the results provided in Table 1 are very consistent no matter what model is used, one would feel comfortable that the results from the primary analysis (ANOVA on the endpoint value for data in the original scale) represent real results.

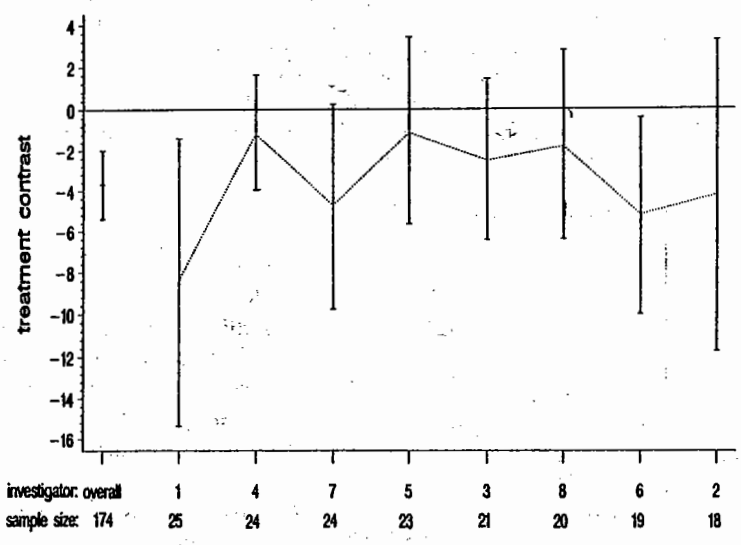


Figure 7: Treatment contrasts (treatment A minus treatment B) with 95% confidence interval by investigator and overall. Investigators are ordered by decreasing sample size.

Visit-wise analyses as well as the LVCF analyses are important to conduct in the presence of dropouts. Disagreement between these analyses should be investigated thoroughly so that any biases due to the missing data can be put into proper perspective. Several authors have proposed methodology to examine the effect of patient withdrawals on the outcome of the study: Gould (1980), Johnson (1992), Brown (1992) and Lagakos et al (1990) discuss analysis strategies for a single outcome while Wei and Johnson (1985),

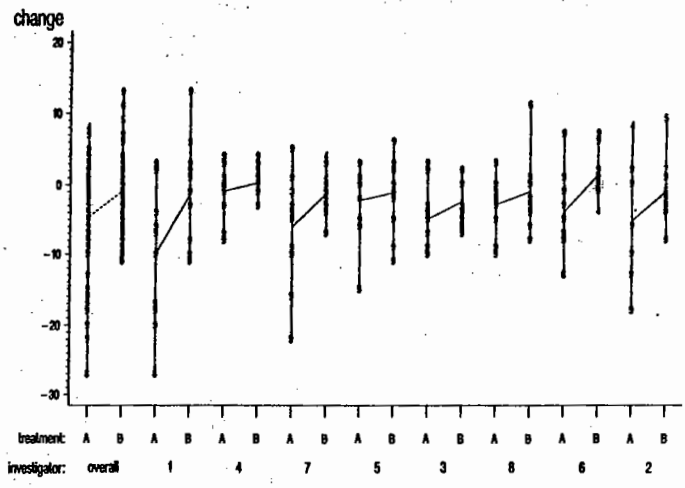


Figure 8: Change in the primary efficacy variable at endpoint from baseline versus treatment and investigator (and overall). Investigators are ordered by decreasing sample size.

Dawson and Lagakos (1991) and Wu and Baily (1989) discuss strategies for repeated measures data. Many other analyses of the data can be conceived depending on the nature of the data and assumptions. In particular, if the previous plots have indicated a potential effect of time of enrollment, this term along with a treatment by time interaction should enter the model. Similarly, the treatment by investigator interaction term should be tested in the model at a predefined level. If it is not significant, the results of the primary treatment comparison should be evaluated combining this interaction with the error term as well as with the interaction split out. If the interaction is significant, evaluation of the results from the SAS Type II versus Type III sums of squares should be compared. ANOVA and ANCOVA with fixed effects are not the only models which should be considered. Evaluation of the results from a mixed effects model for the ANOVA (with investigator as a random effect) or from randomization tests with bootstrapped confidence intervals would provide additional verification of the primary results.

A transformation of the data which is particularly useful to investigate the robustness of the results is to dichotomize the patients into "responders" and "non-responders". This method should use an externally valid definition of what constitutes a responder in the medical condition under study. After patients have been defined as a responder or non-responder, the data are reanalyzed using this binary response. This method is certainly not as efficient as previous methods but it can provide confirmation of results. The set of Cochran-Mantel-Haenszel (CMH) tests (Landis et al (1978) can be used to analyze categorical data across investigators. The CMH test can also be used with the original continuous data without the assumptions of normality and variance homogeneity.

Analyses of the primary efficacy variables should also include a description of the treatment contrasts for each investigator in a multicenter trial. Shown in Figure 7 are the 95% confidence intervals about the treatment contrast for each investigator in order of decreasing sample size. The overall, combined treatment contrast is also shown. The overall treatment contrast can be calculated using an unweighted or a weighted, CMH-like estimator. From Figure 7, one can see that values of the treatment contrast were quite variable across investigators yet one can conclude that the treatment response was consistent (no evidence of qualitative interaction). Alternatively, these figures may be stratified on the values of other important prognostic variables to help identify patient subsets which may benefit from treatment.

Figure 8 shows how the change in the primary efficacy variable from baseline to endpoint varies between the treatments and among the investigators. One can see that the lines connecting the means of treatment A and B have positive slopes indicating that treatment A was superior to treatment B for all investigators. In addition, for most investigators, this figure shows that the variability in response to treatment A was slightly larger than that for treatment B, yet homogeneity of variance seems to be an appropriate assumption.

The results from Figures 7 and 8 may sometimes suggest the presence of treatment by investigator or treatment by prognostic variable interaction. Evaluation of the magnitude of interaction as well as the subsequent course of action in the presence of the interaction is discussed by Gail and Simon (19185) and Chinchilli and Bortey (1991).

Conclusions

There are many more ideas relevant to this discussion which provide ways to understand data and to investigate the consistency of the findings across different subsets of the patients. The methods presented here are essential to reports and publications of efficacy data and will allow a reasonable skeptic the opportunity to draw conclusions about the data as expeditiously as possible. Serendipitously, all of these data displays and analyses may be automated rather easily which will help to free up the statistician to contribute to the other important (or even more important) tasks of trial design, data interpretation, and communication of the findings. This is not to say that automation will take care of all the analytical needs of a study. There may be many additional kinds of analyses that are necessary to make sense of the data that are not automated (e.g. analyses exploring investigator-by-treatment interaction). Yet we hope that by automating some of the basic data descriptions and simple model testing procedures presented above, we will have more time to spend on important aspects of a clinical drug development program mentioned before such as, strategic planning, negotiating with management, and communicating with consultants or regulatory officials.

References

Altman, DG and Royston, JP (1988). The hidden effect of time. *Statistics in Medicine*, 7:629-637.

Brown, M (1992). A test for the difference between two treatments in a continuous measure of outcome when there are dropouts. *Controlled Clinical Trials*, 13:213-225.

Chinchilli, V and Bortey, E (1991). Testing for consistency in a single multicenter trial. *Journal of Biopharmaceutical Statistics*, 1(1):67-80.

Crowder, MJ and Hand, DJ (1990). *Analysis of repeated measures*. New York: Chapman and Hall.

Dawson, JD and Lagakos, SW (1991). Analyzing laboratory marker changes in AIDS clinical trials. *Journal of Acquired Immune Deficiency Syndromes*, 4:667-676.

Enas, G and Rockhold, F (1987). Some useful pictures and methods in survival analysis of oncology data. *Proceedings of the Biopharmaceutical Section of the American Statistical Association*, 266-275.

Gail, M and Simon, R (1985). Testing for qualitative interaction between treatment effects and patient subsets. *Biometrics*, 41:361-372.

Goldman, A (1992). EVENTCHARTS: Visualizing survival and other timed-events data. *The American Statistician*, 46:13-18.

Gould, A (1980). A new approach to the analysis of clinical drug trials with withdrawals. *Biometrics*, 36:721-727.

Johnson, M (1992). Basic Statistics: Patient dropouts and issues with multiple testing. *Journal of Clinical Research and Pharmacoepidemiology*, 6:65-72.

Lagakos, SW, Lim, LL, Robins, J (1990). Adjusting for early treatment termination in comparative clinical trials. *Statistics in Medicine*, 9:1417-1424.

Landis, JR, Heyman, ER, Koch GG (1978). Average partial association in three-way contingency tables: a review and discussion of alternative tests. *International Statistical Review*, 46:237-254.

McNeil, D (1992). On graphing paired data. *The American Statistician*, 46(4):307-310.

O'Brien, E, Perry, I, Sheridan, J., Atkins, N, O'Malley, K (1989). Application of cusums to ambulatory blood pressure data. *Journal of Hypertension*, 7:707-709.

Wei, LJ and Johnson, WE (1985). Combining dependent tests with incomplete repeated measurements. *Biometrika*, 72:359-364.

Wu, MC and Bailey, KR (1989). Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics*, 45:939-955.

Discussion

Gary G. Koch, Ph.D.
Biostatistics Department
School of Public Health
University of North Carolina, Chapel Hill

The authors have presented many helpful suggestions concerning statistical analyses that enable better understanding of the efficacy findings for clinical trials. These suggestions appropriately emphasized informative descriptions for a study as well as strategies to support the robustness of inferential results. The scope for the descriptions included the background status of the patient population, the course of the clinical trial over the time period for its conduct, patterns for patient outcome during the time period of treatment, and patterns of variation over subgroups like investigators for the extent to which one treatment has more favorable outcomes than another treatment. Moreover, well-structured graphical displays effectively enhanced the interpretability of these descriptions. Through both these graphical displays and the inferential results in Table 1, one sees reasonably convincing evidence for better outcome with Drug A in the illustrative clinical trial as well as for logical consistency in the underlying data structure.

A limitation for the illustrative example presented by the authors is its representation of a relatively clear and straightforward situation. For example, all of the p-values in Table 1 are less than 0.01; all investigators entered similar numbers of patients over approximately the same time period; the patterns of study discontinuation were similar for the two treatments; about 85% of the patients in each treatment group completed the study; and patterns of treatment differences were reasonably homogeneous across investigators. The methods presented by the authors effectively support these points, but some uncertainty remains as to how one might proceed for clinical trials for which the data structure and/or the results were not straightforward. In other words, the authors have provided ways of enhancing understanding of efficacy findings for a situation where skepticism of a reviewer may be relatively minimal rather than a situation where it might be low to moderate. Accordingly, an implicit point of the authors relative to their comments on other important areas for the involvement of statisticians is the specification of study designs, data collection procedures, and patient management procedures that are likely to provide relatively clear data structures. In this way, straightforward analyses will be sufficient and thereby will represent a valuable benefit from the collaboration of the statistician in the planning and the conduct of the clinical trial.

Although most clinical trials might have received substantial care and effort from collaborating clinicians and statisticians, the interpretation of their data may be difficult. Some of the potential sources of difficulty are variation in numbers of patients across the respective investigators in a study, variation of treatment differences across investigators, variation of treatment differences across subgroups based on background characteristics of patients, variation of treatment differences across the visits of the treatment period, different patterns of study discontinuation for the treatment groups, and variation in findings across multiple primary outcome measures. The graphical displays suggested by the authors will be helpful in indicating the extent and nature of these difficulties, but may not clarify the robustness of findings to them. Thus, an additional strategy of interest for primary inferences in potentially difficult situations is the use of nonparametric methods with relatively minimal assumptions in combination with procedures to address multiplicity.

One useful class of nonparametric methods for primary outcomes is that for extended Mantel-Haenszel tests. These methods have a scope that includes dichotomous outcomes, ordered categorical outcomes and continuous measurements. Also, their validity does not formally require any assumption concerning treatment x investigator interaction (although the presence of such interactions reduces their power, particularly when it involves conflicting direction). Moreover, the robustness of findings to potential heterogeneity of treatment differences across investigators can be evaluated by applying the extended Mantel-Haenszel test with alternative weightings of investigators beyond the usual ones (which are roughly proportional to sample size); e.g., investigators can be weighted equally or proportional to square roots of sample size. Another attractive feature of extended Mantel-Haenszel tests is that they can be applied with a nonparametric covariance adjustment for baseline or other prognostic variables without involving assumptions for a particular type of relationship between outcome and covariables or for parallelism. They are applicable to multivariate outcomes either jointly or through patient-wise averages for rankings of them in settings where missing values are either maintained as missing, assigned most recent prior values, or assigned conservative values. For more complete

discussion of these methods and examples for their application, see Koch et al (1982) and Koch et al (1990). Strategies for addressing multiplicity include the average ranking principle as in O'Brien (1984) and Lachin (1992), closed statistical tests (for multiple contrasts relative to groups or outcomes) in logical hierarchies as in Bauer (1991), Koch (1991), Koch et al (1993), and Phillips et al (1992), and statistical assessment of a pre-specified nested sequence of subgroups corresponding to prognostic status by use of significance levels in accordance with an Interim analysis spending function like that of O'Brien and Fleming (1979). The theme for addressing multiplicity here is the conduct of as many logically relevant assessments as possible at significance levels near 0.05 and 0.025 while maintaining the experiment wise significance level at 0.05.

The role of the previously described nonparametric methods in combination with appropriate multiplicity procedures is the inferential assessment of treatment differences in the presence of difficulties in data structure. Moreover, their use for this purpose is *specifiable* in the study protocol prior to study initiation and does not require alteration because of difficulties in data structure; but it may require reinforcement to confirm robustness. Such reinforcement may consist of alternative forms for the application of the extended Mantel-Haenszel method or may involve evaluations with related statistical models. Such models include the logistic regression model for dichotomous outcomes, the proportional odds model for ordinal outcomes, and the multiple linear regression model for continuous outcomes; see Koch et al (1990), Phillips et al (1992), and Koch et al (1993) for illustrations and further discussion. For univariate outcomes, computing procedures to fit these models are conveniently available; and for repeated measures outcomes, they are emerging with those recently developed for sample survey data being particularly convenient to use; see Schmid et al (1991) and Shah et al (1991). The specifications for these models can include treatment, investigator, and explanatory variables for aspects of patient background and medical history; and for repeated measures, they can additionally include explanatory variables for the measurement condition (e.g., visit) and have sample survey regression implementation (by use of patient as the primary sampling unit (i.e., source of intra-patient correlation) and weights of one for the respective repeated measures). With the previously outlined structure, these models enable assessment of treatment in a way which accounts for the other variables in the model and thereby shed light on the robustness of findings for treatment effects. Another important use for models is evaluation of homogeneity of treatment differences across patient subgroups (or measurement conditions) through identification of whether the model needs to include corresponding treatment x subgroup interactions (or treatment x measurement condition interactions). Those interactions which are identified by such a process can then be described through displays like those in Figure 7 and 8 of the authors' paper. Since the use of models for such supplementary assessments is after the establishment of confirmatory inferences with extended Mantel-Haenszel methods, it can be undertaken in a flexible and exploratory way.

In conclusion, the authors have provided many useful strategies and tools for essential efficacy data analysis. This discussion has suggested some additional methods that may be helpful for situations with areas of difficulty in their data structure.

References

Bauer, P. (1991). Multiple testing in clinical trials. *Statistics in Medicine* 10: 871-890.

Koch, G.G. (1991). Comment. *Statistics in Medicine*, Vol. 10, 13-16.

Koch, G.G., Amara, I.A., Davis, G.W., and Gillings, D.B. (1982). A review of some statistical methods for covariance analysis of categorical data. *Biometrics* 38, 563-95.

Koch, G.G., Amara, I.A., Forster, J., McSorley, D., and Peace, K.E. (1993). Statistical issues in the design and analysis of ulcer healing and recurrence studies. *Drug Information Journal*, Vol. 27, in press.

Koch, G.G., Carr, G.J., Amara, I.A., Stokes, M.E., and Uryniak, T.J. (1990). Categorical data analysis. Chapter 13 in *Statistical Methodology in the Pharmaceutical Sciences*, D.A. Berry, editor, Marcel Dekker, Inc.: New York, 389-473.

Lachin, J., (1992). Some large-sample distribution-free estimators and tests for multivariate, partially incomplete data from two populations. *Statistics in Medicine* 11: 1151-1170.

O'Brien, P.C., (1984). Procedures for comparing samples with multiple endpoints. *Biometrics* 40: 1079-1087.

O'Brien, P.C., and Fleming, T.R. (1979). A multiple testing procedure for clinical trials. *Biometrics* 35: 549-556.

Phillips, J.A., Cairns, V., and Koch, G.G. (1992). The analysis of a multiple-dose, combination-drug clinical trial using response surface methodology. *Journal of Biopharmaceutical Statistics*, 2(1), 49-67.

Schmid, J.E., Koch, G.G., and LaVange, L.M. (1991). An overview of statistical issues and methods of meta-analysis. *Journal of Biopharmaceutical Statistics*, 1 (1), 103-120.

Shah, B.V., Barnwell, B.G., Hunt, P.N., and LaVange, L.M. (1991). *SUDAAN User's Manual*, Release 5.50 Research Triangle Institute, Research Triangle Park, N.C.

Discussion

Robert L. Davis

Astra/Merck, Wayne, PA 19087-5677

Greg Enas always demonstrates an insightful grasp of the practical issues in clinical trials. In this paper, he and his co-authors have laid out a thorough battery of diagnostics for efficacy analyses. Even though pharmaceutical companies' approved submissions are available through the Freedom of Information Act documents, it was interesting for me to see how another company analyzes clinical trials. The standard analyses at drug companies are determined by such factors as which methods were emphasized at the graduate school the highest ranking statistician attended or by the most recent discussion at regulatory agencies. Thus, statistical methodology can vary between companies but often converges to what FDA wants.

Although I usually prefer a good table to graphics, I found the authors' displays to be better than tables. Their Figure 1 shows time since first patient randomized for each investigator. In their example, the authors note that patient entry was generally evenly distributed over the study. But, suppose all of the patients from Investigator #1 had entered the study on days 1-100 and all of Investigator #8's patients on days 401-500. What do we learn from this information? Perhaps Investigator #8 was added late only because patient accrual was too slow. Or maybe Investigator #8 initially could not find

patients who met the entry criteria so then bent the rules to meet his quota. In the latter case, results from Investigator #8 should be examined carefully. Since the authors do not really describe how they use this graphic, it would be interesting to see some examples of how it has been useful to them.

In Figure 2 the basic information on patient dropouts is displayed. A variation on the graphic would be to slant the lines for each treatment group, with the steeper slope representing the faster dropout rate. I am curious about how a patient who misses a visit would be counted in this table. Suppose a patient's observations fell as follows:

Week	0	1	3	5	7	9	11	13
Visit	2	3	4	5	6	7	8	9
Observed	X	X		X	X	X	X	X

Would he still be counted in the Week 3, Visit 4 total? Would his Week 5 observation become his Visit 4 observation, because it is his fourth visit, or stay as Visit 5 because its timing is consistent with a range of days at Visit 5? If there are a number of such missing values, even a simple table like Figure 2 can be very tricky. Since bookkeeping issues may lead to reanalyses to satisfy regulatory authorities, it is important to carefully lay out the ground rules by which a patient appears in displays such as Figure 2.

The graph of the patient's baseline versus the time he entered the study (Figure 3) could be a useful adjunct to Figure 1. Curiously, in this example patients randomized early into the trial were less severe and better responders than those randomized later. My experience has been that the more severe patients, who become better responders because they have more room to respond, are entered earlier. Then, as investigators get more desperate for accrual, they enter the milder patients.

Figures 4A and 4B display the last-visit carried forward (LVCF) data versus time, with the last visit observed plotted as the symbol. There is a lot of information captured here, maybe too much. In general, graphs this complicated, and using scary sounding things like CUSUMS, will be useful to a statistician but of less value to a clinician, the end user. If such graphs are included in a statistical report they should be included in an appendix, rather than in the main body of the text.

The graphical comparison of baseline to last value seems obvious, but statisticians probably rarely provide it. The paper implies that the authors' approach to analysis of change from baseline data is to see if the regression slope is significantly different from 1.0, and if not, analyze change directly; presumably, if the slope is not 1.0, an analysis of covariance would be used. Over the years my method of choice has rotated from 1) analysis of covariance on treatment value versus baseline to 2) analysis of covariance on change versus baseline from baseline to 3) analysis of variance on ranks of changes, then back to 4) analysis of covariance on changes versus baseline. Thus Table 1 is especially useful for statistical wimp like me who cannot quite decide which method to use to analyze change data. It provides p-values for almost any reasonable method one could propose. It is interesting that the authors, as do I, use an analysis of the LVCF as primary. This "endpoint analysis" originated in analyses of psychotropic drugs, then became the method of choice when regulatory agencies began requiring intent to treat analyses. Using the last value as an estimate for all subsequent time points can be silly when a patient drops out early. The main virtue of the LVCF is its simplicity, but better estimates are available. The mixed model analyses of Laird and Ware (1982), Cnaan (1991), and Getson and Cnaan (1993) provide more reasonable estimates of missing data but suffer from being computational cumbersome. Table 1 does report the repeated measures p-values to balance the LVCF analyses.

I really like Figures 6A and 6B, the McNeil graphs. One gets a real sense of the treatment effect, the consistency of response and the effect of dropouts all at once. Pictures of this type could be used in the main body of the report of the clinical study. The presence of even a few early dropouts can greatly change the between-treatment p-value, depending on which method one uses. Trials of psychotropic or anti-inflammatory compounds especially are influenced by dropouts. The authors provide a good review of this topic but are vague about how "biases due to the missing data can be put into proper perspective".

The authors' suggestion to perform an additional analysis on responders is reasonable but probably will not be useful in the approval process. Regulatory authorities tend to be wary of categorization, since one can find literature "justification" for many different definitions of responders.

Figures 7 and 8 are useful for identifying "problem investigators" or inconsistencies of response in levels of concomitant variables. Again, these could be included in the report if they addressed an important issue. The authors suggest that these analyses should be automated and I agree totally. Moreover, I believe that an "expert system" could use p-values and summary statistics to write much of the statistical report and free up statisticians for the tough analyses and the strategic thinking.

Enas, Sanger & Huster have covered the major issues arising in an efficacy analysis and I congratulate them on sparking discussion on this topic. I also would be interested in seeing comments from the FDA about how the analyses presented are viewed by regulators.

References

Cnaan, A. (1991) "Application of longitudinal methods with differential follow-up due to chronic disease." *Proceedings of the Biopharmaceutical Section of the American Statsitical Association*, 124-129.

Getson, A. and Cnaan, A. (1993) "A comparison of some longitudinal methods to analyze treatment response in the presence of missing data." Presented at Biometric Society ENAR Meeting, Philadelphia, Pennsylvania. PA, March 1993.

Laird, N.M. and Ware, J.H. (1982) "Random effects model for longitudinal data." *Biometrics*, 38, 963-974.

Discussion

Richard A. Stein

Food and Drug Administration, Washington, D.C.

The introductory discussion of Drs. Enas, Sanger, and Huster brings up issues that have nagged me over the years. Although my work environment is surely different from those of the authors, the similarities are striking. Points that come to mind first are:

- "How can a reasonable skeptic draw conclusions about the data as expeditiously as possible?"
- "... too much data analysis and reporting to do in too little time."

- "We would like to help create a working environment where statisticians have time to spend on important aspects ..."
- "There is a need to come to some understanding of the most helpful analyses for a large number of trials and then automate such analyses."

When I read past the introductory part of this paper and into the author's example, the similarities stopped. My outlook is different, and I will try to look into this.

Most clinical trials have protocols. Written into most protocols is a proposed statistical analysis of each primary efficacy variable. These analyses are essential, necessary, but usually not sufficient to convince reasonable skeptics. We have to deal with two sets of skeptics, i.e., ourselves and others. This seems like a silly distinction but the distinction seems important. I will have start at the edges to work into explaining it.

I believe that the authors are considering a situation where the protocol analysis has been completed, and results are favorable on the surface. Now is the time for some healthy skepticism. Here, the skeptic should be ourselves. We ask what might have gone wrong? What data diagnostics are efficient for deciding whether the protocol analysis results are robust? The paper of Enas et al. proposes several graphical diagnostic data displays that could easily come from a parallel group arthritis or dental pain trial. Depending on what these graphs show, the authors cite many literature references designed to examine what might be happening. If you, yourself are the skeptic, many different choices are available. Given the same diagnostic information, a colleague of yours might choose to analyze and report the situation correctly, but quite differently. So far, there is nothing wrong. Nevertheless, a problem arises when the skeptic is no longer yourself. Some skeptics don't have much free time.

The authors examine a variety of features of a clinical trial in ways that both the statistician and the non-statistician can understand visually. My own job gives me a special perspective that I imagine is not the same as the authors. Suppose the authors' plots reveal features of a study that make an outside skeptic uncomfortable. Are the authors' plots going to provide solid ground for establishing that the protocol analysis should be abandoned; or even more simply are the authors' plots best for showing that a trial has failed?

While "standardizing" common features of the analysis and reporting of a clinical trial are not the authors' focus, it is one of mine. In the clinical trial protocol, the primary statistical analyses should be laid out. In order to have time to think about what seems to be a constant barrage of interesting problems, I feel that expediting the mundane without sacrifice of quality is first.

For instance, this means identifying what is the most reasonable way to analyze a dental pain study. Currently, most dental pain protocols are not only abundant but they are very similar. It seems unnecessary each time to rediscover the statistical analysis of the primary efficacy variables, i.e., the pain relief and the pain intensity scales. However, there exists an abundant variety of statistical analyses applicable to the very same clinical data. To speed and to bring more uniformity into the drug decision making process, I see the need for statisticians from FDA, from the pharmaceutical industry, and independent knowledgeable statisticians to identify how we feel these trials are best analyzed and formatted for reporting.

The benefit of "standardized" analyses need not be limited to expediting the task of reporting and reviewing clinical trial data. The medical community is interested in bringing together results from many drugs that have been reported from many trials of a given class. When these trials are analyzed by highly diverse methods, one feels little confidence in combining individually published summary statistics. If a class of trials is well analyzed in uniform fashion and if the summary statistics are planned with present and future use in mind, science should benefit.

Discussion

Lloyd D. Fisher

University of Washington and
The Fred Hutchinson Cancer Institute

This paper by Enas, Sanger and Huster is entitled "Essential Efficacy Analysis." The paper has two central themes. First, it is argued that we should have more standardized ways of examining data to free biostatisticians to spend a larger proportion of our time on tasks requiring more ingenuity and deeper thought. This is a goal we can all support and such support may be analogous to being for "baseball, motherhood, and apple pie." Second, examples of data analysis are given to help progress toward the goal of analyzing data about drug efficacy in standardized ways. The examples include (among others) examination of patient characteristics and treatment effect over time of enrollment; also examined is the presentation of clinical site variability.

"Expert Systems" and Software

There has been considerable effort over the last decade in developing "expert systems." Such systems, almost always, embody in software the expert judgment of selected experts in a field and reduce (as far as possible) the judgment process to an algorithmic form. The resulting systems are only as helpful as the expert judgment. It strikes me that we are proceeding in this direction as we follow the path of this paper. The authors note that automation of their methods is necessary. This standardization of data analysis begins (at least implicitly) to implement expert judgment.

One method of serving the pharmaceutical community (including the public, pharmaceutical companies, FDA and interested academics) would be the development of software guided by an informed steering committee with representatives from, say, the FDA, PMA and interested academics. It is important to note that such a group must not be composed exclusively of biostatisticians — others including knowledgeable clinicians and pharmacologists should help select and refine useful methods of data analysis and presentation. Such software appropriately validated and available at a relatively modest cost for both mainframes, workstations and PCs, would go far toward actually implementing any standardization. Clearly for some aspects of drug development the FDA would prefer standardization (as shown by suggested tables in the draft NDA guidelines). If such an effort were to be performed it would be necessary to keep in mind some limitations: 1) The implemented methods of analysis and data presentation are only suggestions. It would be a big mistake to imagine that such complex science could be embodied in an expert system. Moreover the standards and methods change with time and revisions to such software might sometimes lag behind the best current approaches. 2) As just suggested newer and more preferable methods of data presentation can be expected to emerge over time. Thus a continuing group would be needed for the continuing development and modification of such software.

3) Analyses of drug efficacy endpoints can vary by indication. It might be that there were relatively distinct versions of "standard" analyses.

Robustness of Data Analysis and Control of Type I Error

Table 1 of the paper presents a number of analyses (in fact 24) with respect to the analysis of the example data. In the happy circumstance where all are significant, as occurs here, it is comforting. However the approach presented here would be an inappropriate precedent. I personally believe that in a regulatory setting with extremely large investments and potential returns it is important to have the level of proof required specified as well as possible; further adherence to the criteria should be fairly rigorously followed. Thus table 1 could introduce more problems than it cures. Would Eli Lilly consider a trial unconvincing if, e.g., all ranked analyses were not statistically significant? In anti-hypertensive studies the between patient variability is usually much larger than the variance of change in the individual patient and adjustment for baseline blood pressure (as a covariate or using the change) is needed by study design. Further, analyses of completers can be quite biased (in either direction) and such analyses can often be misleading for efficacy. ANOVA and ANCOVA of ranked data has some theoretical limitations and I find it less convincing than the usual models (provided the model assumptions are tenable). (For ranked analyses randomization p-values are preferable.) For all these reasons many of the presented methods may not be desirable — not to mention the multiple comparison problem. Would the FDA want to consider a study positive if 16/24 analyses satisfied $p< 0.05$, but the prespecified analysis does not? While exceptions can be argued on a study by study basis, I do not like the idea of routinely doing a large number of analyses to demonstrate robustness. [We fight over the multiple comparison problem when we cannot avoid it — such as for adverse events. Why deliberately introduce the problem when it could be avoided?] It is better to assure robustness by examining the assumptions of the one or few prespecified analyses directly — by examination of influential points, examination of residuals, by using randomization p-values, etc. Robustness may be better investigated by sensitivity analyses with respect to assumptions, e.g. if there are many patients with no follow-up data, how aberrant would their experience need to be to "remove" the statistical significance? Or if the censoring of time to an event data may not be statistically independent of the outcome, how sensitive is the conclusion to this assumption? [See Fisher and Kanarek (1974)].

Essential Efficacy Analysis by Time of Enrollment

In many trials the enrollment characteristics vary over time for many reasons. 1) The enrollment rate at each clinic may increase variably over time as study organization improves, knowledge of the study brings in more referrals and pressure is brought to bear on the investigators. 2) Often clinics begin enrollment at different times — this may be due to delays in IRB approval, approval by regulatory agencies in different countries at different times, and late recruitment of clinical sites because of slower than expected patient recruitment. Of course there are many other reasons for variability over time. Thus changes in entry characteristics over time are not unusual in large studies. Of how much concern is this? I suggest it usually not "essential" for data analysis and if such graphical presentations are desired they would usually be in an appendix to a report, NDA, etc. Of more interest is the relationship of patient characteristics to treatment response — especially if there are qualitative interactions. Of more interest are treatment center differences, this frequently occurs as a main effect and not infrequently as an interaction with treatment response. Sometimes this may be "explained" by differing patient covariates as an alternative possible explanation. But in may cases we never know why the interactions occurred. In summary, variability over time of entry characteristics is less important in many cases. An important exception could be when concomitant therapy changed dramatically over the course of a long entry plus observational period.

More Detailed Comments

While it may be more enjoyable to discuss relatively general issues as above, the real progress must come with specific detailed approaches to the data analysis and presentation. I am a strong supporter of more graphical presentations and applaud the authors particularly for addressing graphical presentations of data. Thus this discussion concludes with a few more detailed comments.

Figure 2 would be more telegraphic if the numerical information were also graphical. For example the number available at each time point (visit) could be represented on arrows whose width is proportional to the number available at the specific visit. The actual numbers would also be on the figure. [See CASS Investigators (1983) for an example.] If many patients miss visits (apparently not the case here) the density of the shading could indicate the proportion of patients with future visits to come who were observed at the particular visit. Figures could also incorporate drop-out information with a figure below the numbers evaluated figure; this second figure could give by treatment arm, bar graphs subdivided with the reason for drop-out — for example the causes might be lack-of-efficacy, adverse events, both, patient refusal, administrative reason, lost to follow-up, and other. Figure 5: it would be best to use the same scale on both axes. Figures 6A and 6B: I liked the figures but would combine the increased and decreased to make it easier to see the relationship to baseline values. Figure 7: With a relatively small number of clinics (as in the figure) one might consider Box plots with very dark lines for the 95% confidence intervals for the mean effect.

Plotting numbers for the number of the last visit adds interesting information to the plot. An alternative approach would be to either: 1) have circles whose area is proportional to the number of visits; or 2) a symbol with shading proportionate the number of the last visit. I do not know that this would be better, but merely suggest that many approaches need to be examined in a number of settings before suggesting "standard presentations."

The authors are to be complimented for furthering discussion of standardized methods of presentation and analysis in these important areas.

References

CASS Principal Investigators and their Associates (1983). Coronary Artery Surgery Study (CASS): A randomized trial of coronary artery bypass surgery. Survival data. *Circulation*, 68:939-950.

Fisher, L. and Cyanuric, P. (1974). Presenting censored survival data when censoring and survival times may not be independent. *Proceedings of the SIAM Conference on Biometry and Reliability*, pp. 303-326.

Rejoinder

Gregory G. Enas, Todd M. Sanger, and William J. Huster
Eli Lilly and Co.

The intent of our paper was to put forward some initial, yet admittedly incomplete, ideas regarding the analysis of efficacy data from a clinical trial data. In that vein, our "rejoinder" to the discussion must reflect that we appreciate the comments and suggestions the discussants have offered. Each discussant is an experienced hand with clinical trials and their opinions are tried and true. Our hope was to get the ball rolling with some basic ideas; the discussants have added substantively to the topic. We will comment briefly here to each discussant and trust that the reader will improve their own practice of clinical trial data analysis with some of the ideas that they have seen here. In general, we note that there can be no rule as to what figures or tables should appear in the text of the report versus an appendix or elsewhere. This depends on the audience and the intent of the study. In turn, the intent of the study dictates whether the tools we presented are used in a confirmatory or exploratory manner.

Although our paper encouraged the use of Mantel-Haenszel (MH) type estimators and tests, we are grateful that one of the leading experts in this field has shared his comments with us. Professor Koch has given a very complete and useful synopsis of the extended MH and related methodologies. The methods developed for sample survey data also look very attractive. We heartily concur with his recommendations and feel that the MH, randomization and possibly other tests would be a useful adjunct to those presented in Table 1, particularly from the standpoint of verifying the robustness of the results from the pre-specified primary analysis.

In his commentary, Dr. Davis has provided a good example of the utility of Figure 1 to identify problematic investigators. Figure 1 could also be used to determine the effect of the occurrence of an external event (e.g. interim analysis) on the patient accrual patterns in the study. As Dr. Davis suggests, Figure 2 can be very tricky, not because of any inherent trickiness, but because of the nature of clinical trials. Unfortunately, patients miss visits. The figure we show can be used in a number of different ways, and we agree with Dr. Davis that the ground rules should be carefully laid out. The purpose of this figure is to show how many patients were still in the trial at each visit. Thus, patients would appear even for those visits they missed.

The influence of dropouts on treatment comparisons is a lengthy subject. Suffice it to say that in the presence of dropouts, many sensitivity analyses that address a number of legitimate ways to handle missing data may be warranted. Analysis of responders is useful when applying agreed upon definitions that were specified a priori. Dr. Davis' comments concerning the usefulness of the plots and analysis strategies for regulatory submissions were enlightening, particularly since he is a seasoned veteran of regulatory interaction.

We appreciate Dr. Stein's perspective on this topic and we are glad that standardization is a topic of interest to him. He poses some interesting questions for which answers can only be derived with use of the tools we have suggested. He offers many solid reasons for standardization, both from a speed perspective as well as a quality perspective. Faster decisions can be made when the analytical wheel does not need to be reinvented. Quality decisions can be made when a large number of trials with the same drug or spanning many drugs from a given class are reported and analyzed in the same manner. Meta-analyses are enhanced in this fashion.

We did not mean to imply that standardization is a system which provides "only one way to do it". Standardization provides one the opportunity to automate a variety of ways to look at the data. In a confirmatory mode, these analyses will assess the robustness of the primary analysis. In an exploratory mode, they will allow one to explore!

Professor Fisher offers some very helpful suggestions. Development of "expert systems" must be accompanied by expert judgment which will help the system meet the expectations of the user. The limitations Dr. Fisher notes will be helpful in setting future expectations. Dr. Fisher makes other additional comments which could improve the graphical figures we presented. We agree that the relationship between response and patient characteristics or treatment center differences may often be more important than variability over time. We present ways to show all of these graphically. Time trends do become more important in studies where treatment allocation depends on the outcomes observed to date (such as "play-the-winner" designs) or in long-term studies of chronic disease.

With reference to Table 1, we agree wholeheartedly that the level of proof required should be pre-specified as much as possible. This level depends, in part, on the intent of the study -exploratory or confirmatory. To investigate the robustness of the pre-specified primary analysis in a confirmatory study, Table 1 offers one way to do exactly what Dr. Fisher recommends - examine the assumptions of the analysis. In fact, the development of an expert system as Dr. Fisher advocates will facilitate the production of tables like Table 1. This table does not replace examination of influential points or residuals, but puts them in perspective. One should not do many different analyses just to see how many are less than .05 nor should one necessarily be concerned that "all ranked analyses were not statistically significant". On the other hand, further exploration would be warranted if the range of p-values from these analyses ranged from .001 to .500, for example. Different analyses make different assumptions; congruence among the results of different analyses implies that the results are invariant to differing assumptions. The issue here is similar to the likelihood principle: will the cohort of reasonable skeptics generally agree on the same conclusion as that offered by the pre-specified primary analysis? Analyses in addition to the primary analysis are indeed sensitivity analyses with respect to assumptions; consequently, there is no multiple comparison problem here. Rather, the problem is a multiple "comparator" problem—there are many people looking at the pre-specified primary analysis and assessing its strengths and weaknesses.

We thank all the discussants for their comments and hope that statisticians and their medical colleagues will continue to make progress along these lines in order to expedite the delivery of quality, cost-effective medicines to waiting patients.

Section News

Minutes of the August 10, 1992, ASA Biopharmaceutical Section Business Meeting

Camilla Brooks opened the Business Meeting by welcoming the attendees. She then asked if there were any corrections to the 1992 meeting minutes as published in the *Biopharmaceutical Report*, Summer 1992. There were no corrections and the minutes were approved.

The new elected officers were introduced. The new officers are: Bob Starbuck, Chair-elect; Bob Davis, Secretary-Treasurer; and Program Chair-elect, Kenneth Koury. The new executive committee members, Jerome Wilson and Nguyen Dat, were also introduced.

The awards for the outstanding papers presented at Biopharmaceutical Section sponsored sessions at the August 1991 meetings were presented by Lilliam Kingsbury. She reported that three persons received the top score so rather than a first and second place awards, as planned, the total award money was divided three ways. The awardees were: Gayle S. Bieler for the presentation "Ratio Estimates, the Delta Method and Quantal Response Tests for Increase Carcinogenicity", Arthur J. Roth for the presentation "A New Statistical Method for Analyzing the CHO/HGPRT Mutation Assay" and Howard T. Thaler for the presentation "Pharmacodynamics of Analgesia Produced by Morphine and One of Its Metabolytes." Because of the positive response to these awards, the competition was being extended to the 1992 meeting; evaluation forms were dispersed at the Biopharmaceutical Section sponsored sessions at the 1992 meeting and the awardees will be announced at the August 1993 Biopharmaceutical Section business meeting.

Gladys Reynolds reported for Program Chair Dan McGee. The following sessions were scheduled for the August 1992 conference:

3 invited sessions, 1 special topics session, 4 contributed sessions and 1 poster session. Of particular note, was the presence of the Section at the spring WNAR Conference; Gladys Reynolds chaired a Biopharmaceutical Section sponsored session. The hope was expressed by Dr. Reynolds that the Biopharmaceutical Section would not only continue but would expand participation in the WNAR Conference.

Nick Teoh, Chair of Work Groups, reported that the ten luncheon roundtables scheduled for the August 1992 conference were sold out. Discussion then ensued on how to increase the number of tables. The ASA suggestion of scheduling roundtable sessions during the technical sessions was presented. Discussion on the suggestion, as well as the suggestion to increase the size of the roundtable discussion group, followed. The suggested change for scheduling round table sessions and technical sessions concurrently was not considered to be acceptable because the roundtable sessions would be competing with technical presentations; attendees may very well want to attend both. Increasing the size of the roundtable discussion group was rejected because it was felt that the advantages of small group interaction could not be maintained with a larger group. A suggestion was made form the floor for "Brunch" and/or "Breakfast" Roundtables - similar to the noontime luncheon roundtables. There was considerable support for this type of additional session and the chair was asked to pursue this concept.

Lilliam Kingsbury, Chair, reported on the Quantitative Literacy Committee. She stated that Katherine Rowe is the ASA contact and suggested that if a need to talk to Ms. Rowe arises the contact be made directly rather than through the Section. A newsletter is prepared 3 times a year and assistance is needed to prepare an article encouraging participation in the program.

Camilla Brooks stated that a second issue of the *Biopharmaceutical Report* has just been distributed and, as noted by many, it looked great! She stated that Tuli Cnaan has done an outstanding job as editor.

Gary Neidert, Chair of the Membership Committee, reported on activities for increasing membership in the Section. He reported the committee had designed and printed copies of a pamphlet about the Biopharmaceutical Section that would be used as a handout, along with copies of the *Biopharmaceutical Report*, for the recruitment booth during the annual meeting. It was also reported that a draft copies of the Section's *Manual of Operations* were distributed. Comments were to be submitted to Dr. Neidert by the end of January 1993.

JeAnne Burg, treasurer, presented the treasurer's report (see attached). At the conclusion of the report, considerable discussion followed whether the current section dues was too high given the treasury balance. It was pointed out that this balance did not reflect all of the costs associated with the current issue (Summer 1992) of the *Biopharmaceutical Report* or

Letter from the Editor

In the last issue, we discussed safety, and this time our concern is efficacy. Greg Enas, Todd Sanger and William Huster wrote a stimulating paper with wonderful graphics that provoked interesting discussion. If a picture is worth a thousand words, then this is our longest issue yet!

We owe apologies to Larry Gould. We switched his figures around a bit. Some of you thought it was an intentional IQ test, but we simply goofed. SORRY! We hope everybody figured it out by now.

Well, it is summer. Relax, kick off your shoes and enjoy the report.

Avital Cnaan
Editor

other committed expenses and the true balance was therefore less than that amount. A commitment was made to include a complete treasurer report with the publication of the business meeting minutes. (Note: The treasurer's report and a discussion of Section expenditures are included in this issue of the *Biopharmaceutical Report*. John R. Schultz, Chairperson of the Biopharmaceutical Section Finance Committee, discusses initiatives and future plans being discussed and considered by the Section Executive Board for implementation in 1993).

The meeting was adjourned.

Submitted by JeAnne Burg
Section Secretary/Treasurer

Letter to Section Members

Camilla Brooks
CB Quantitatives

In January the officers for 1993 took over the reins of the Section. However, I felt it was not too late to write to the membership and express my appreciation for the chance to work with you during the last year. Included in this issue of the report are the minutes of the Business meeting held in August at the annual meetings and a report from the finance committee, so I have tried not to be overly repetitious.

I believe one of our biggest accomplishments as a Section in the past year was publication of our first *Biopharmaceutical Report*. We may thank Avital Cnaan (we all know her as "Tuli") for the fine job that she did and is continuing to do with the report. It has taken an enormous effort on her part to produce this very informative and professional-looking document. The purpose of this report, of course, is to serve the members by discussing issues of importance to the industry and conveying news of interest to the Section, and she would appreciate your input in expanding its benefit to the readers. More book and software reviews, consulting questions, news items, and suggestions for section involvement from you, I am sure, will be welcomed.

Another avenue for your input will soon be available. Akbar Zaidi agreed last year to work on a new round of our recurring Section membership survey, which was last conducted in 1989. You should be sure to use this opportunity to suggest ways in which the Section can better serve you and the profession as a whole, including naming activities in which you would like to be actively involved. John Schultz in his Finance Report has discussed some of the initiatives suggested by the membership and considered by the Executive Committee. We hope you will use the above avenues—Biopharmaceutical Report and the Section's membership survey—to comment on the initiatives discussed as well as suggest others for consideration by the Section.

The past year also saw the revision of the Section charter. Gladys Reynolds headed the committee for this and the membership approved it through balloting last spring. In addition to the charter, work began on revision of the Section handbook. Gary Neidert was kind enough to volunteer for this task. Gary also used the knowledge he gained while working on the handbook to design a brochure on benefits of belonging to the Biopharmaceutical Section for the Council of Sections recruitment table at the annual meetings. Both of them, along with members of their committees, are to be commended for their hard work.

The awards for "Best Presentation" that were instituted in 1991 were presented at the 1992 Annual Meetings by Lilliam Kingbury. More details of this are given in the minutes of the Business meeting. The purpose of the awards are to foster good presentations at the Section-sponsored paper sessions. We are pleased that presenters of contributed papers continue to agree to participate in the awards programs and that the purpose of the awards is being realized; it is evident that the presenters have been making an effort to give good presentations. We would like to thank Lilliam for chairing this committee last year and for all the work that it involved. Criteria for awards to students is evolving; we are grappling with the issue as to whether the main aim of the awards should be a means of rewarding excellence or of encouraging an interest in the profession. Travel awards will be given this year to the annual meetings based on "best abstract"; however your input on future disbursement of awards is needed.

As in past years, there was continuing concern for education of its Section members through the Continuing Education Committee chaired by Nguyen Dat, training to statisticians, and general quantitative literacy. We as members, are encouraged to participate as individuals in the Quantitative Literacy programs sponsored by our local chapters. Some ideas for Section-sponsored activities in this arena are discussed in John Schultz's finance report.

The Section was active as usual in the annual meetings, conferences, and workgroups. We participated in the ENAR spring meetings, the Annual Meetings, and for the first time, the WNAR meetings. We would like to thank Dan McGee, the program chair for 1992, and Nick Teoh for his work with the roundtable discussions at the August meetings. Nick, also head of the Workgroup, reported at our Executive meeting that The Population Model Workgroup has completed its report; there may be a special contributed or invited session on the results in the future. Peter Imrey representing the Task Force on Design and Analysis in Dental and Oral Research (an informal Section Working Group) reported on Task Force activities, including developing guidelines to the American Dental Association for evaluation of therapeutic products. Karl Peace continued to do a fine job with the Conference on Applied Statistics, and Patrick O'Meara and the new representative to the Midwest Pharmaceutical Workshop, Frank Rockhold, continued in their work.

In addition to those people mentioned above, I would like to thank all the officers and members of the Executive Committee, as well as the Section members for keeping our Section active in the ASA and the profession as a whole. In particular, I would like to thank Gladys Reynolds, whom I called on an inordinate number of times, for information, advice, and assistance during my year, and John Schultz for his helpful advice on financial and other matters. Additionally, I would like to thank Bruce Rodda who was always willing to give me input and support; Chris Gennings for her work with the publications committee and with whom I had numerous conversations; JeAnne Burg for kindly taking over the last year of Sharon Anderson's term as Secretary-Treasury (Sharon was elected Chair of the Council of Sections) and for volunteering to pitch in, in time of need; and Janet Begun for agreeing to serve the last year of Nancy Flournoy's term (Nancy is Chair-elect of the Council of Sections).

We again would like to thank our Executive Director, Barbara Bailar, Carolee Bush, Penny Young, Alison Stern-Dunyak, and other members of her staff for all their help. They were not only very responsive to our questions, but were always very patient and pleasant while doing so.

Report of Biopharmaceutical Section Finance Committee

John R. Schultz
Chair, Section Finance Committee

One of the provisions of the new ASA constitution is to allow the Sections greater freedom to sponsor activities and programs that promote their interests. It is the Section's responsibility to raise funds to support these activities and programs. The Biopharmaceutical Section has a history of promoting activities reflecting the interests of the membership. These include organization of workshops, symposia, study groups and special sessions related to biopharmaceutical statistical methodology and applications.

A number of new initiatives have been suggested by members and considered by the Section's Executive Committee. As ideas were generated, it became clear that a source of funding would be required to implement them. The traditional source of funding is through membership fees. In order to begin

Section News

implementation of some of the initiative, fees for individual members were set at $11.00 ($10.00 for Section and $1.00 for ASA). This level was set to fund three items: a newsletter, "best presentation" awards, and "best manuscript" awards. The first two of these have been implemented. All members now receive the *Biopharmaceutical Report;* awards for "best presentation" were made for Section-sponsored sessions at the 1991 and 1992 meetings. After further study, the logistics of presenting a "best manuscript" award seemed too difficult to manage at this time.

There are still a number of proposed initiatives with a great deal of merit which would require additional funding. These include sponsorship of speakers and guests at Section-sponsored meetings by supporting travel and other expenses; providing scholarships for statistics students interested in biopharmaceutical applications; sponsorship of activities related to recruitment of students into the statistics profession in general and biopharmaceutical applications in particular; providing seed money for relevant workshops, short courses and seminar; and supporting sabbaticals and professional exchange programs to enhance interest and understanding of biopharmaceutical applications.

The above initiatives, along with several others, have been reviewed by the Executive Committee. A decision was made not to implement new activities at this time as this may necessitate a long-term commitment. Any long-term commitment cannot be made without considering future cash flow as adequate cash flow and cash reserves are mandatory for responsible support. The Executive Committee did not feel the current cash flow was adequate for making further long-term commitments; further sources of revenue were needed. To this end, the Executive Committee decided to maintain individual membership dues at the same level for 1993 and to continue recruiting corporate members.

Hints on Preparing a Nomination for ASA Fellow

Robert R. Starbuck
Wyeth-Ayerst Research

- A nomination package must be assembled and submitted to the Chair of the Committee on Fellows. Because the package will take some effort to prepare, the person preparing the package *must* be motivated and committed to completing the task.
- It is recommended that the person submitting the package be an ASA Fellow, or that an ASA Fellow be added as a co-author of the package if the primary preparer is not a Fellow.
- The deadline for submission of the nomination package is 1 March. To meet this deadline, begin the effort on the package at least 6 months in advance.
- Determine whether the person you wish to nominate for Fellow (nominee) is not already a Fellow. If not, verify that the nominee is a member of the ASA and has been for at least 3 years prior to the submission deadline of 1 March.
- Check with the nominee to determine his or her willingness to be nominated. Though few would decline to support efforts on their behalf, some persons may choose not be nominated.
- Ask the nominee for a current curriculum vitae and names of persons who know the nominee well and could provide supporting statements.
- Authors need to be recruited to supply supporting statements for the activities listed on the nomination form. They should focus on selected activities, and the assignments of activities to the authors should provide adequate coverage of the list of activities on the form. A supporting statement from the ASA Section(s) that the nominee is a member of is recommended, and statements from ASA Fellows are strongly advised. Keep the number of authors from the nominee's place of employment to a minimum.
- Give the persons assigned to prepare supporting statements a deadline. Keep in touch with them periodically; regular contact is usually necessary to insure that the statements are received on time.
- Using your knowledge of the nominee and the information in the supporting statements, prepare a high quality presentation that will convince the members of the Committee on Fellows that the nominee deserves to be an ASA Fellow. Be sure to include the key points from the supporting statements. Make the presentation visually appealing, e.g., use a letter quality printer and a suitable font. Have your presentation reviewed and commented on by at least one other person.

Future Biopharmers of America

Michael N. Boyd
Pharmaceutical Research Associates, Inc.

During the recent Biometrics meeting in Philadelphia it was my pleasure to attend the session on "Training the Next Generation of Biostatisticians." One of the speakers made the statement that "The United States has the best graduate

education in the world...but the secondary education stinks!" The statement made an impression on me, and as I looked around the room, which was quite full, I couldn't help thinking that if each of the statisticians in attendance would be willing to make some sort of statistical presentation in their local schools, it would have a positive effect. An easy way to get started doing this is to participate in the ASA Council of Chapters Adopt-a-School program.

The goal of the program is to have professional statisticians visit local schools three or four times a year and work with the teachers to introduce statistical concepts to the students. There are many, many ways to do this — the choice is up to you. For many years now my wife and I have been giving two presentations: one involving flipping pennies and counting M&Ms (Grades K to 3), and another involving an actual randomized study of how pulse rate is affected by exercise (Grades 3 and up). There is a nice notebook of materials available from the Council of Chapters that is designed to facilitate such efforts. John Boyer, Chair of the Council, tells me that the new kits will be ready in August or September of '93. If you are interested in obtaining one of them, send your name and chapter affiliation to:

John E. Boyer, Jr.
Department of Statistics
Kansas State University
101 Dickens Hall
Manhattan, KS 66506

or call him at (913) 532-6883. John and Steve Ruberg (Marion-Merrell Dow) were instrumental in getting this program started. Interested members should also remember that Kathryn Rowe, Center for Statistical Education-ASA, is a valuable resource. She can be reached at ASA, 1429 Duke St., Alexandria, VA 22314-3402; (703) 684-1221; e-mail: kathryn@asa.mhs.compuserve. com.

While the focus of the Adopt-a-School program is on primary and secondary education, we have also been challenged by our ASA president, Stuart Hunter, to extend our efforts to non-statisticians at all levels. I know that some of you have been involved with short-courses within your companies.

In short, participating in the Adopt-a-School program is fun, satisfying, and easy to do. With some consistent participation, informal as it may be, we can make a tremendous impact in the area of statistical education.

Book Review

The Use of Restricted Significance Tests in Clinical Trials. David S. Salsburg, Springer-Verlag, 1992

Reviewed by Robert D. Small
Burroughs Wellcome Co.

This interesting book covers a tremendous amount of ground and should be recommended reading for any statistician working on clinical trials. In fact, it would be useful to any applied statistician, though the author, a well known practitioner of clinical trials, couches all of his examples and comments in terms of clinical trials.

The book is a conglomerate of the author's views on significance tests, a short history of the development of hypothesis testing, some philosophy about testing, scientific method applied to clinical trials, and a good number of examples of innovative ways to analyze data from trials. The style is easy and flowing so anyone with a graduate degree in statistics could easily read the book in a short time. There are large parts of the book that could profitably be read by non-statistician clinical trialists.

The book comes in two parts. The first, which lasts 48 pages and 5 chapters, is entitled "Philosophical and Scientific Problems when Applying Statistical Methods to Clinical Data." It contains a short history of Neyman, Fisher and Pearson's efforts on significance testing, some of the author's philosophy about scientific method, clinical research and the meaning of probability. The second part, which the author admits can be read without the first part, is entitled "Techniques for Applying Restricted Tests to Data from Randomized Controlled Clinical Trials." This is a wide ranging discussion of methodology that would be useful to any applied statistician. Some of it is connected to the ideas of restricted tests. I do think, however, that all of the discussion could have gone on without mentioning restricted tests. There are some parts, like the introductory description of the bootstrap and the jackknife, that have nothing to do with restricted tests. These are means of constructing test statistics and the construction of a restricted test has to do with the choice of an appropriately restricted set of alternate hypotheses.

There are several chapters in the second part that show the vast experience of the author in clinical trials. The first chapter gives a novel way to analyze data from a two-period cross over design. It relies on the details of the design and a good bit of knowledge of the medicine underlying the observations. The standard analysis is avoided. This is good since it gives a null result despite the fact that a cursory view of the data seems to indicate a strong effect. An analysis and set of alternate hypotheses is set up based on the medical question. The result is an overwhelmingly positive test statistic and an estimate of the effect of the drug that is much more useful than the p-value.

The author then proceeds in subsequent chapters to discuss combining data across measures, analyzing counts data and several other useful things. At times, the discussion is little more than would be found in a good applied statistics text. For example, the discussions of permutation, jackknife, and bootstrap methods give a relatively simplified description of these techniques. It would have been interesting to me to see a problem that the author believed had no solution other than using the bootstrap. In fact, none of these techniques are applied to any data in the book.

The final long chapter is a fitting climax to the book. The author takes a study in clinical depression and analyzes the data in great detail using some of the techniques and philosophy that he has developed in the course of the book. This would be a good training module for a new statistician arriving at his first job at the clinical trial research group. Any applied statistician would enjoy reading the chapter. The data is presented in great detail; many of the techniques previously described are used in an appropriate way, hypotheses are

justified from a medical point of view, and results are interpreted. There is a discussion of how to handle dropouts.

In the analysis of these data, Salsburg constructs an unconventional F test. First, factor analyses on previous data sets are quoted as justifying the formation of four factors from among 17 questions. These are described as being stable in other patient populations and so are not dependent on the data at hand. Standard ANOVAs and F-tests were carried out on each of the factors. The F-tests were significant in two of the four cases.

Salsburg then argues that it is important to have a single overall test statistic that gives a single answer to the question of whether there is an important medical difference between the two treatments. To get this statistic, he adds the mean squares for treatment over the four independent factors and calls this a numerator. He then adds the four error mean squares to get a denominator with four times as many degrees of freedom as any of the individual tests had. Of course, the newly constructed F-test is highly significant and we can conclude that there is an important medical difference.

This general approach is consistent with the author's views throughout the book. He tries to use knowledge from sources other than the data in the experiment at hand to develop hypotheses, construct tests or suggest analyses. He argues forcefully that we know enough to do more than a t-test and report the standard p-value. I think that most applied statisticians with extensive experience in any field, including clinical trials, would agree with him. I do have some problems with the approach in this problem and some cautions in general.

I am not convinced that the overall F-test described above tells us any more than we knew without it. There is no doubt that given the assumptions (factors independent, normal theory, etc.) that the constructed statistic has an F distribution with the appropriate degrees of freedom. However, I cannot see what hypothesis is being tested. Further, if the factors are independent, isn't it important that the treatment affected two of them differently and another two identically? If we believe that the factors mean something, then why would we need to mix them up with a contrived test statistic?

Salsburg often stresses that our results need to be useful to a user of the results of the clinical trial. I think in a few cases in his drive to search for all of the information available, he has gotten to a point that would make it difficult for some of the users to use his results. For example, I do not believe that very many reviewers—either journal referees or regulatory—would accept his test statistic on the depression data. Though I think that his analysis of the data from the cross-over trial is brilliant, most reviewers would ask to see the usual analysis, and if there were differences in results they would worry about things like whether the more sensitive method that Salsburg uses was also sensitive to or made more assumptions. That would be a reasonable concern.

It is interesting to note that the simplest example of a restricted significance test is a one tailed t-test. Very few of the people who admire p-values (regulatory reviewers, clinical journal editors, etc.) can even tolerate the existence of one tailed p-values. In part, this is because the introduction of a restricted test has made the situation more complex than they can handle and so the result is not useful to them.

Not withstanding these few alternate viewpoints, I applaud the author for a stimulating and well written book and highly recommend it to all analysts working in clinical trials.

1993 Joint Meetings

Come to the Biopharmaceutical Section Sponsored Sessions at ASA!

August 9-12, 1993
San Francisco, California

7. Monday, August 9, 8:30 a.m.–10:20 a.m.
Some Issues and Methodologies in Factorial Clinical Trials - Invited Papers

35. Monday, August 9, 10:30 a.m.–12:20 p.m.
Meta Analysis and Multiple Comparisons —Contributed Papers

72. Monday, August 9, 2:00 p.m.–3:50 p.m.
Issues on Bioequivalence Studies—A Review of FDA Guidance on Statistical Procedures —Invited Panel

102. Tuesday, August 10, 8:30 a.m.–10:20 a.m.
Multicenter Clinical Trials - Contributed Papers

131. Tuesday, August 10, 10:30 a.m.–12:20 p.m.
Use of Neural Networks in Analyzing Clinical Trial Data - Special Contributed Papers

159. Tuesday, August; 10, 12:30 p.m.– 2:00 p.m.
Biopharmaceutical Section Roundtable Luncheon

169. Tuesday, August 10, 2:00 p.m.–3:50 p.m.
All Patients Have Been Followed After Drop-Out: What to Do with Their Data? —Invited Papers

197. Tuesday, August 10, 4:00 p.m.–5:50 p.m.
Estimation and Testing in Dose Response Settings—Contributed Papers

228. Wednesday, August 11, 8:30 a.m.–10:20 a.m.
Survival Analyses and Normal Theory—Contributed Papers

252. Wednesday, August 11, 10:30 a.m.–12:20 p.m.
Combination and Multiple Drug Trials—Contributed Papers

274. Wednesday, August 11, 12:00 p.m.–2:00 p.m.
Poster Sessions —Contributed Papers

290. Wednesday, August 11, 2:00 p.m. –3:50 p.m.
Longitudinal Data Analysis —Contributed Papers

320. Thursday, August 12, 8:30 a.m.–10:20 a.m.
Estimation and Testing in Bioequivalence Studies — Contributed Papers

334. Thursday, August 12, 10:30 a.m. - 12:20 p.m.
Sample Size and Power - Contributed Papers

ASA Biopharmaceutical Section Executive Committee—1993

Immediate Past Chair

Camilla A. Brooks, Ph.D.

President, CB Quantitatives, Inc.
PO Box 659
Upper Marlboro, MD 20773-0659
301-868-1097
301-868-4474 Fax

Chair

Bruce E. Rodda, 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 - Elect

Robert R. Starbuck, Ph.D.

Assistant Vice-President
Clinical Biostatistics and Data Management
Wyeth-Ayerst Research
145 King of Prussia Road
Radnor, PA 19087
(215) 341-2070
(215) 341-2092 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

Section Financing

John Schultz, Ph.D.

The Upjohn Company
200 Portage Road
Unit 9165-32-2
Kalamazoo, MI 49001
(616) 385-7427
(616) 329-5548 Fax

Program Chair

Mark Scott, Ph.D.

Director, Biometrics & Medical Systems
Clinical and Medical Affairs
ZENECA Pharmaceuticals
Office Wing 1
Routes 202 & 141
Wilmington, DE 19897
(302) 886-8495
(302) 886-1235 Fax

Program Chair Elect

Kenneth J. Koury, Ph.D.

Building 60, Room 203
Lederle Laboratories
North Middletown Road
Pearl River, NY 10965
(914) 732-3827
(914) 735-5249 Fax

Work Group Coordinator

Nick Teoh, Ph.D.

Biostatistics
Shering-Plough Research Institute
2000 Galloping Hill Road
Kenilworth, NJ 07033
(908) 298-5986
(908) 298-4688 Fax

Chair of the Midwest Biopharmaceutics Statistics Workshop

Patrick D. O'Meara, Ph.D.

Director, Statistical Services Division
624 Peach Street
P.O. Box 80837
Lincoln, Nebraska 68501
(402)-476-2811
(402)-476-7598 Fax

Representative, Applied Statistics Conference

Karl E. Peace, Ph.D., FASA

President, Biopharmaceutical
Research Consultants, Inc.
4600 Stein Road, Suite B
Ann Arbor, MI 48105
(313) 663-4440
(313) 663-7797 Fax

Publications Officer

Chris Gennings, Ph.D.

Biostatistics Department
Medical College of Virginia
Box 32
MCV Station
Richmond, VA 23298-0032
(804) 786-9824
(804) 371-8482 Fax

Biopharmaceutical Report *Editor*

Avital Cnaan, Ph.D.

Division of Biostatistics, Department of Pediatrics
University of Pennsylvania
The Children's Hospital of Philadelphia
34th Street and Civic Center Boulevard
Philadelphia, PA 19104-4399
(215) 590-4486
(215) 590-4487 Fax

Chair of Committee on Advisors to FDA

Vern Chinchilli, Ph.D

Associate Professor
Department of Biostatistics
Medical College of Virginia
1 East Marshall Street
PO Box 32, MCV Station
Richmond, VA 23298-0032
(804) 786-9824
(804) 371-8482 Fax

Council of Sections Representative

Edward S. Nevius, Ph.D.

Division of Biometrics
Food and Drug Administration HFN-713
5600 Fishers Lane
Rockville, MD 20857
(301) 443-4594

Council of Sections Representative

Janet M. Beegun, Ph.D.

Vice President
Pharmaceutical Product Development
1400 Perimeter Park Drive, Suite 100
Morrisville, NC 27560
(919) 380-2000
(919) 380-2022 Fax

Liaison Activities

Helen Bhattacharya, Ph.D.

Director of Biostatistics
Sanofi Pharmaceuticals, Inc.
40 E. 52nd Street, 13th Floor
New York, NY 10022

Executive Committee

John Lambert

Director, Biostatistical Operations
Sandoz Pharmaceuticals, Inc.
59 E. Route 10
East Hanover, NJ 07936
(201) 503-6914
(201) 503-6363 Fax

Executive Committee

Gary Neidert, Ph.D.

Director of Clinical Data Management
9165-298-139
The Upjohn Company
301 Henrietta Street
Kalamazoo, MI 49001
(616) 329-8591
(616) 329-5579 Fax

Executive Committee

Akbar Zaidi

CDC Mathematical Statistician
Division of STD/HIVP
Center for Prevention Services
1600 Clifton Road, N.E.
Atlanta, GA 30333
(404) 639-1283
(404) 639-2555 Fax

Executive Committee

Lilliam Kingsbury, Ph.D.

Director of Biostatistics
Bio-Pharm Clinical Services
512 Township Line Road
Blue Bell, PA 19422
(215) 283-0770
(215) 283-0733 Fax

Executive Committee

Jerome Wilson, Ph.D.

Director of Biostatistics and
Data Management
Warner-Lambert
170 Tabor Rd., Room 3015
Morris Plains, NJ 07950
(201) 540-2422
(201) 540-4300 Fax

Executive Committee

Nguyen V. Dat, Ph.D.

Senior Technical Advisor
Biometrics Research Institute
1300 North 17th Street
Arlington, VA 22209
(703) 276-0400

Election Results

Congratulations to Our Newly-Elected Section Officers!

The following members were chosen to serve the Biopharmaceutical Section in the recent ASA elections. All terms begin January 1, 1994.

- *Chair-Elect:* **Lilliam Kingsbury,** Bio-Pharm Clinical Services, Inc.
- *Program Chair-Elect:* **Joseph F. Heyse,** Merck Research Laboratories
- *Section Representatives:*
  Denise J. Roe, University of Arizona
  Harji I. Patel, Berlex Laboratories, Inc.

Back Issues Available

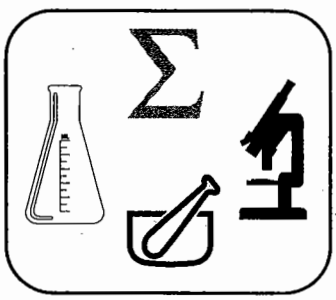
A limited number of additional copies of the first three issues of the *Biopharmaceutical Report* are still available from the editor. Please call, write, or fax your address and the issue numbers that you did not receive, and we will mail you the *Biopharmaceutical Report* while supplies last.

Let's Hear from You!

If you have any comments or contributions, please contact:

Dr. Avital Cnaan
Division of Biostatistics
The Children's Hospital of Philadelphia
Room 9012
34th Street and Civic Center Blvd.
Philadelphia, Pa 19104-4399
(215) 590-4486
(215) 590-4487 Fax

The Biopharmaceutical Report *is a publication of the Biopharmaceutical Section of the American Statistical Association.*



Biopharmaceutical Report
c/o American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3402
USA

FIRST-CLASS MAIL
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 9959

FIRST CLASS POSTAGE

2419 12/93
DR. ROBERT L. DAVIS
215 MANOR ROAD
HARLEYSVILLE PA 19438