FDA and Industry: Working Together to Expedite the Development of New Pharmaceutical Products
Christy Chuang-Stein

The Biopharmaceutical Section held its 1997 workshop on October 27 and 28 in Bethesda, Maryland. The workshop, with a theme of FDA/industry partnership, attracted approximately 200 registrants. One objective of the workshop was to create a congenial environment to foster exchange between the FDA and the industry statisticians. During the two-day workshop, exchange was earnest with communication at the grass-roots level identified as a priority for future workshops of this nature. In this regard, the workshop achieved its goal of serving as the beginning of an open dialogue between the FDA and the industry statisticians which, hopefully, will continue and flourish in the many years to come.

The opening session of the workshop, Working Together Toward a Common Goal, included a critical examination of the future challenges of the statistics profession in new drug development by Bruce Rodda and a historical review of the charge of the biometric units at FDA’s Center for Drugs Evaluation and Research (CDER) by Chuck Anello. Rodda discussed the supporting role of the statistics profession in the ‘60s and ‘70s and its leadership role in the ‘80s and ‘90s in the pharmaceutical industry. He challenged the workshop participants to learn from the evolution of our profession and seize the opportunities to lead in the strategic development of entire research projects. Anello reviewed some key events that led to the increasing number of statisticians in the regulatory environment. The events include the thalidomide disaster; Kefauver-Harris Amendment (1962), requiring substantial evidence of efficacy before approval; Drug Efficacy Study Implementation (1970); Government Performance and Results Act (1993); Prescription Drug User Fee Act (1992); proposed Mutual Recognition Agreement; and the International Conference on Harmonization (ICH). Bob O’Neill (CDER), the third speaker of the first session, encouraged statisticians to expand support to areas in need of statistical input. One example is the decision whether to revise a product’s safety profile based on the product’s periodic safety update reports. The latter practice will soon become a requirement per ICH E2C. O’Neill also mentioned the efforts of several statistical working groups in CDER to address various statistical issues. O’Neill urged workshop participants to regard challenges as opportunities and to use wisely the increasing resources to meet the present-day challenges and the challenges of the next millennium.

The second session focused on Discovery to Early Clinical Trials. Jeff Meeker from Bristol-Myers Squibb discussed statisticians’ contributions to lead finding and drug screening. Opportunities in this area include, but are not limited to, high throughput screening, high throughput chemistry, compound libraries, genomics and bioinformatics. Daphne Lin, a Statistical Team Leader in the anti-infective area at CDER, discussed statistical designs and analyses of animal carcinogenicity studies. Richard Simon from the National Cancer Institute proposed some accelerated titration designs for...
phase II oncology trials. Simon's designs include an accelerated phase of dose escalation based on one patient per cohort during the early phase of the trial and intra-patient dose escalations for subsequent courses. Simon proposed a new analytic method that incorporates multi-course information. Richard and his colleagues conducted simulation studies to evaluate the performance of these designs.

The third session featured three presentations by CDER statisticians on Planning the Confirmatory Activities. Not surprisingly, ICH E8 and E9 were frequently quoted during this session. The first speaker Hoi Leung, whose topic was Points to Consider for Confirmatory Study Designs, gave three interesting examples on study designs. The first example emphasized the point that a study would be judged with respect to the stated objectives. Based on the results from prior studies, a company decided that a new therapy would be evaluated first in mono-therapy patients, followed by previous responders among mono-therapy patients, and lastly the entire population. Unfortunately, the results showed that the new therapy was significantly better than the placebo in adjunctive patients, but not in mono-therapy patients and the entire population. The study was considered a failed study by the agency. The second example concerned a proposed composite endpoint. In cancer patients with intractable pain not satisfactorily controlled by morphine, a sponsor proposed to test a new non-opioid medication which was to be used as an adjunctive to morphine. The proposed measure was a composite endpoint combining reduction in pain and morphine use. However, since pain reduction and morphine use were confounded, the sponsor was advised to find another endpoint. The third example illustrated the power of suggestion. In an oral analgesic trial on acute pain, the primary endpoint was time to remedication. According to the instructions to the patients, patients could request, as needed, every 6 hours either the same medication or something else for pain relief. As it turned out, clusters of remedication occurred every 6 hours in all groups. A similar trend of clustering around the suggested time period occurred when 4 hours was mentioned in another situation. It appeared that the suggestion of a time interval for re-dosing might have mistakenly conveyed to patients the message that they should ask for re-dosing at the suggested time intervals. Instead of a suggested time period, Leung stated that patients should be instructed to request re-medication when they feel the need without suggesting a time interval. The second speaker, Nancy Smith, in her presentation Points to Consider for Electronic Data Submission, struck a familiar tune on the need for a comprehensive outline of procedures and analysis plans for confirmatory trials. Smith indicated that sponsors may choose to alter the original analysis plan based on accumulating masked data. Following Smith, Aloka Chakravarty gave a presentation on Points to Consider for Electronic Data Submission, and indicated that CDER is moving away from accepting a sponsor's hardware for electronic data submission. The trend at CDER is to load the information on the agency's network server to allow simultaneous access to the data by multiple medical and statistical reviewers. Chakravarty emphasized the importance of prospective design and seeking FDA's input at the phase II stage. She also stressed the desirability of using CANDA to prepare the NDA documents to help ensure consistency between the agency's and the sponsor's own files. The take-home message here is to start a dialogue with the FDA on the electronic submission plan before the confirmatory trials. This can be most efficiently done by requesting a small group meeting to agree on the architect and structure of the electronic files.

The last session on the first day discussed The Assembly of an NDA from industry's perspective. Bob Chew from Pfizer gave an overview of Pfizer's process for assembling the clinical data in an NDA. Pfizer's presentation was followed by Michael Lu from Abbott who contrasted diagnostic assay development to that of a new drug. The last presentation in this session was given by Susan Kenny from Quintiles. Kenny gave an update on industry's effort in the area of computer assisted new drug applications. Kenny discussed some current trends in the electronic submission including Web-based technology to create a CANDA environment. Kenny also noted the trend to develop CANDAS for in-house sponsor use during the entire phase of drug development. Thus, electronic submission to the FDA is an end product and not an effort by itself.

The second day started with a session on QA and QC—The Mysterious Black Box. The first presentation, given by Matthew Thomas from the Clinical Investigations Branch in CDER, elaborated on the principles of clinical trial inspections. Pharmaceutical statisticians usually don't get involved in this aspect, but the results will nevertheless impact our activities since investigators identified to have major problems will be excluded from subsequent analyses. Thomas reviewed the history of CDER's Bio-research Monitoring Compliance Program, the procedures of clinical trial inspections and the outcome of an inspection. The outcome is classified into one of three categories—in compliance, deficiency of minor or moderate significance, or major problems. Steve Wilson then gave his perspective on QA and QC of clinical trials as a statistical reviewer. Wilson gave an example where the original efficacy scores in a trial were changed by the sponsor, resulting in an entirely different p-value from the previous analysis. In his opinion, the QA and QC responsibility of a statistical reviewer at CDER includes: assessing compliance with the protocol's analysis plans, assisting in the planning of clinical trial site inspections, checking the appropriateness of statistical models and conclusions, verifying results reported in the NDA, modifying models and assessing robustness/sensitivity of the results, modifying data sets and re-analyzing, examining the trial and data for bias, and assessing impact of audits (if we need to exclude questionable sites). Finally, Nancy Paul Silliman (CDER) illustrated a procedure currently in use in the anti-infective area to evaluate equivalence claims. The procedure resulted from a collaboration among several statisticians in Division of Biometrics IV. Since a medical reviewer in the anti-infective area typically goes through each individual patients records and decides the individual's evaluable and response status, the review can be tedious and time-consuming. The proposed procedure involves taking a random sample (e.g., 25%) from the data-
base and asking the medical reviewer to evaluate each patient in this sample. Once the medical reviewer completes the evaluation of each individual in this sample, the statistical reviewer will construct a bootstrap confidence interval for the difference in the response rates. If the bootstrap confidence interval is already within the allowable interval defining equivalence, equivalence is concluded. Otherwise, increase the percentage of the random sample (e.g., take an additional 25% sample) and repeat the process. Thus one might be able to reach the equivalence conclusion earlier without the medical reviewer's evaluating and classifying every patient. Silliman's presentation generated a lot of interest among the workshop participants.

Following QA and QC, a session was devoted to Pathways to Decision Making. This session included presentations on the NDA review process by Rajaphalan Srinivasan (CDER), preparing for the Advisory Committee Meeting by Paul Flyer (CDER), and preparing for the Advisory Committee Meeting from the industry's perspective by Deborah Shapiro (Merck). Srinivasan outlined analyses typically conducted by the statistical reviewer in an NDA review. Flyer talked about the use of Advisory Committees to discuss general scientific issues, controversial applications and issues of public interest. He gave three examples of recent Advisory Committee Meetings in the Office of Drugs Evaluation IV. Shapiro provided a glimpse of how Merck prepared for the Advisory Committee Meetings. Basically, Shapiro indicated that they practice, practice, and practice until everyone is driven crazy; then they practice more. There is simply no substitution for practice in getting ready, according to Shapiro.

The workshop set aside one session to look at Challenges in the Development of Biologics and Devices. Both biologics and devices are growing areas where statistical evaluations are plowing new ground. Susan Ellenberg from Center for Biologics Evaluation and Research (FDA) covered statistical challenges in areas of blood and blood products, risk assessment of transmissible diseases, vaccines, cellular/gene therapies, and xenotransplantation. Greg Campbell from Center for Devices and Radiological Health discussed the differences between developing a drug and developing a device. Unique challenges in the development of a new device include the needs for new designs, community-based informed consent, statistical means to define substantial equivalence, and post-marketing surveillance. Furthermore, practices such as mid-trial adjustment, tendency to rely on historical data, using a single trial for both safety and effectiveness, and reluctance to perform adequate clinical feasibility and pilot trials pose challenges in the regulatory evaluation of a new device application. The last formal presentation at the workshop was given by James Reimann from Genetech who discussed advances in non-standard controls. Reimann illustrated how far protein-based therapeutics (e.g., erythropoietin, growth hormone, insulin, t-PA, vaccines) have come during the past 20 years.

The last session at the workshop was a panel discussion: How Can We Foster a Mutually Beneficial Working Relationship? Panelists included representatives from the FDA and industry. In an effort to encourage floor participation, no formal presentations were planned for this session. The discussion was brisk and lively. Issues discussed during this session included agreeable modes of communication (e.g., direct or in the presence of a CSO), the likelihood of having two statisticians on an Advisory Committee, and the sharing of FDA’s review of an NDA prior to the dissemination of such information to the Advisory Committee members. Even though no definitive decisions and commitments could be promised at the workshop, the communication issue definitely caught everyone’s attention.

Workshop participants received copies of the presentation materials. For individuals who did not attend the workshop but would like to obtain a copy of the workshop materials (bound together in a booklet), we have limited copies that can be purchased at $25 apiece. Please direct your request to Tracey Dowden at the Meetings Department of the ASA (Phone: (703) 684-1221 x147; E-mail: tracey@amstat.org). Requests will be processed on a first-come first serve basis. Please watch for announcements for future Biopharmaceutical Section workshops at the Section’s Web site, in Amstat News, and in the Biopharmaceutical Report. For those who couldn’t join us at the 1997 workshop, we hope to see you at our future workshops.

1998 Biopharmaceutical Section/FDA Workshop
The Biopharmaceutical Section and the FDA Statistical Association will co-sponsor a workshop on Current Statistical Issues in Drugs, Biologics, Medical Devices, and Risk Assessment on September 24 and 25, 1998. The meeting will take place at the Crystal City Hyatt in Arlington, Virginia.

Registration information and other details will be available in June. The workshop will feature speakers from the FDA, industry, and academia. Planned sessions include Modernization Act, Use of GEE models, Bayesian Methods, Risk Assessment, Diagnostic Devices, Computer Intensive Methods, and Use of Non-Standard Controls.

For more information, contact Bob Small at 919-286-8917 or small003@mc.duke.edu.

Errata in Membership Survey
Several errors have been identified in the tables in the Report of October 1996 Survey of Biopharmaceutical Section Members published in the Fall/Winter 1997 issue.

Table 2: The column headings of Male and Female were reversed.
Table 52: The N for Bachelor’s degree is 46.
Table 61: The word ‘Meetings’ should be deleted from the title.
Appendix Table 17: The second category should be 1-5.
Biopharmaceutical Section JSM Scientific/Technical Program
Tom Capizzi, 1998 Program Chair

The ASA Biopharmaceutical Section has planned a comprehensive and exciting program for the August 9–13 JSM in Dallas. I want to thank all of the individuals who volunteered to plan and to participate in the section’s scientific activities. The section will sponsor four invited paper sessions and four special contributed paper sessions. The section will sponsor/co-sponsor two short courses and will have six roundtable luncheon topics. In addition, there will be eight regular contributed paper sessions and one poster session.

Invited Paper Sessions

Immunogenicity and Efficacy Issues in Vaccine Clinical Trials:

Brian Wiens (Merck) and Tony Lachenbruch (FDA) have organized a session that focuses on issues that occur in the development of vaccines that may not have received the attention that is deserved. The organizers believe that statisticians are able to make immediate contributions to developing new and effective vaccines against common illnesses. This session will bring together statistical experts from academic, industrial and regulatory backgrounds to discuss common issues. Some of the papers that will be presented in this session may also be applicable to problems outside the immediate domain of vaccine development. Correlates of protection (i.e., surrogate markers) are very important in a variety of settings in the pharmaceutical industry. Some papers are utilizing random effects models in vaccine development problems; these models have become quite popular in many areas of statistical research and application in recent years. All three papers involving random effects models will present new and exciting models for vaccine development problems. The authors and titles of the presentations are as follows:

- Norm Bohidar, Merck; Joe Pigeon, Villanova University; and Brian Wiens, Merck: “Issues in Predicting the Duration of Vaccine Induced Protection.”

The discussant will be Tony Lachenbruch of the FDA.

Assessment of the FDA Draft Guidance on Statistical Procedures for Bioequivalence Studies:

J. P. Liu of the National Cheng-Kung University is the organizer of a session that examines the draft FDA guidance entitled “Statistical Procedures for Bioequivalence Studies: Population and Individual Bioequivalence.” The draft guidance asks sponsors to provide evidence of individual bioequivalence for all generic drug products as well as the innovative drug products whose post-approval changes need bioequivalence testing as required by SUPAC. For the changes of formulations prior to the NDA submission, the draft guidance asks for population bioequivalence. Because the current requirement for average bioequivalence will be replaced by either individual bioequivalence or population bioequivalence, this sweeping change required by the draft guidance and its impact on the pharmaceutical industry will be tremendous. In addition, the statistical designs and analysis of individual bioequivalence are much more complicated with a lot of unresolved issues. The speakers are as follows:

- Shelly Zacks, SUNY Binghamton: “Dose Escalation Schemes for Cancer Phase I Trials.”

The discussant will be Naitee Ting, Pfizer.

Biopharmaceutical Report, Spring 1998

Permutation Tests in Clinical Trials:

The organizer of this session is Vance Berger of the FDA. Permutation or exact tests are often considered for use in the evaluation of clinical trials and epidemiologic studies. In this session, examples are used to illustrate the utility, strength, and limitations of permutation tests. Presentations will describe the dependence of such tests on the observed test statistics and methods for ameliorating this dependence, the utility of rank-based alternatives for analysis of ordinal outcomes, the advantages of multiple outcome permutation tests in identifying important individual outcomes, and the development of exact tests for detecting familial disease clusters and risk factors. The session will include the following papers:

- Weng Kee Wong, UCLA: “Designing Studies for Dose-Response.”
- Nancy Flourny, American University; Stephen Durham, University of South Carolina: “Ehrenfest Urn Designs for Quantile Estimation.”
- Shelly Zacks, SUNY Binghamton: “Dose Escalation Schemes for Cancer Phase I Trials.”
- Discussant: Naitee Ting, Pfizer.

Formal Design Considerations for Phase I Clinical Trials:

This session organized by William Rosenberger of the University of Maryland will address recent advances in the design of Phase I studies. Phase I clinical trials are of critical importance, particularly in cancer research. The goal is to find a maximum tolerated dose (MTD) that can be passed to later phases of testing. This goal leads to a conflict between individual and collective ethics. On the one hand, it is important to protect volunteers from highly toxic dose levels. On the other hand, it is important to estimate the MTD with precision; underestimating will potentially lead to an ineffective dose being selected for later testing. Most designs in the past have been ad hoc without much consideration for formal estimation procedures. In this session state-of-the-art designs, using optimal design theory and sequential methods, will be presented. This session will include the following speakers and discussant:

- Weng Kee Wong, UCLA: “Designing Studies for Dose-Response.”
- Nancy Flourny, American University; Stephen Durham, University of South Carolina: “Ehrenfest Urn Designs for Quantile Estimation.”
- Shelly Zacks, SUNY Binghamton: “Dose Escalation Schemes for Cancer Phase I Trials.”
- Discussant: Naitee Ting, Pfizer.
**Roundtable Luncheon Topics:**

Richard Entsuah (Wyeth-Ayerst) has organized six roundtable luncheon topics. The topics and discussion leaders are as follows:

1. **Statistical Issues in Clinical Trials for Medical Devices**
   - Gregory Campbell, Ph.D., Food and Drug Administration

2. **Summarization of Adverse Events in Clinical Trials**
   - Brian K. Varney, Trilogy Consulting Corporation

3. **Statistical Analysis in Outcomes and Pharmacoeconomic Studies**
   - Christopher M. Barker, Ph.D., Principal Pharmacoepidemiologist, Roche Pharma Business-Palo Alto

4. **General Linear Mixed Models that are Well Formulated for Hypothesis Testing**
   - Professor John Overall, University of Texas, Houston

5. **Flexible Clinical Trial Designs**
   - Alexandra D. Carides, Ph.D., Merck Research Laboratories, Blue Bell, PA

6. **Dealing with Dropouts in a Clinical Study**
   - William Myers, Ph.D., The Procter & Gamble Company

A description of the topics can be found on the Section's Web site.

---

**Special Contributed Paper Sessions:**

These sessions consist of five contributed papers that are organized around a particular topic. This format allows for a more focused discussion of relevant statistical issues. Gregory Campbell of the FDA has organized a session on statistical issues in therapeutic and diagnostic devices research. Statistical issues in the evaluation of health-related quality of life will be addressed in a session organized by Juan Sanchez, Mary Lynn Brecht and Geri Padilla of UCLA. Ron Helms of the University of North Carolina has organized two sessions that showcase talented students who are conducting research in the area of incomplete data.

**Short Courses:**

The section will sponsor a short course on the Design and Analysis of Clinical Trials. The instructors will be J. P. Liu and Shi-Chung Chow. This course will be based on their forthcoming book. In addition, the section will be a secondary sponsor of short course given by Darlene Stangl on Bayesian Methods in Biostatistics. The primary sponsor of this course is the Bayesian Statistics section.

---

**Active Control Equivalence Trials**

**Leader: James B. Whitmore**

Immunex Corporation

The use of active control equivalence trials is becoming more common in the development of new medical therapies. Whether conducted due to ethical concerns over a placebo control, due to the desire to obtain comparative data for marketing purposes, or in order to demonstrate “bioequivalence” for therapies with negligible blood levels, these trials play an increasingly important role in the assessment of new therapies. The planning, conduct and interpretation of results from these trials, however, can be fraught with difficulties. Participants in this roundtable discussion were from the pharmaceutical industry and academic medical research centers. The group discussed a number of theoretical and practical issues which are summarized below.

Active control equivalence trials are fundamentally different from traditional, placebo-controlled trials (Garbe 1993, Jones 1996, Senn 1993). Many of these differences can be attributed to the objective of “proving the null hypothesis.” For example, there is the necessity for a high standard of conduct with rigorous attention to follow-up and protocol deviations because increased noise may favor the demonstration of no difference between treatments. Although the equivalence problem can be formulated in the framework of conventional hypothesis testing by reversing the roles of the null and alternative hypotheses, the assessment of equivalence is best made through confidence intervals. If the confidence interval includes only values deemed to be clinically insignificant, then equivalence has been demonstrated. In this way, one is not misled by the p-value from an over- or under-powered test.

It is the determination of this largest clinically insignificant difference (Δ) that poses a major problem in the planning of an equivalence trial. While clinicians often treat the choice of “Δ” for a placebo-controlled trial as a statistical issue, in an equivalence trial it is imperative that Δ be acceptable to the medical community and, if the trial is for registration purposes, to regulatory agencies, as well. Generally, Δ will be smaller than that observed in the placebo-controlled trials of the active comparator. How much smaller may be a function of the magnitude of responses one would expect in the target patient population, and may also depend on the anticipated toxicities of the active control and test treatments. Based upon the bioequivalence criteria of ±20%, one might suggest Δ should be 20% of the treatment effect (difference between active control and placebo responses). However, this might be too strict a requirement, especially if the treatment effect is small. Clearly, the value of Δ should be such that it rules out the possibility that the test treatment is more like placebo than it is like the active control. So, a rule of thumb might be that Δ be no more than 50% of the active control’s treatment effect.

Another issue is whether equivalence should be one- or two-sided. Clearly, bioequivalence should be two-sided, in order to balance efficacy and toxicity. When dealing with a therapeutic equivalence trial, however, a case could be made for a one-sided confidence interval. On the other hand, there
may be a reasonable likelihood going into the trial that the test treatment could be superior to the active control. As regulatory agencies prefer the use of two-sided tests for placebo-controlled, superiority trials, and in order to allow for the possibility of the test treatment being superior to the reference, it is advisable to assess equivalence with a two-sided confidence interval. There is a growing body of literature concerning the simultaneous testing of equivalence and superiority (Dunnett 1996, Morikawa 1995).

Because the monitors, investigators, and patients know all participants are receiving an active treatment, there is the concern that results of an equivalence trial could be biased. Assuming the trial is randomized and blinded, this should not introduce a bias in the treatment comparison, however the degree of efficacy relative to that seen in previous (placebo-controlled) trials with the active control may be elevated. This should not invalidate the conclusion of equivalence—as it might if the level of efficacy had decreased—but does illustrate one of the frailties of active control equivalence trials: without a placebo control, they depend on an external validity check. Hence, it is critical that an equivalence trial’s active control dosing, its patient population, its trial design, and its endpoints be patterned on the previous trials of the active comparator. Unfortunately, this may not be possible if the patient population or standard of care has changed substantially since these previous trials were conducted. If this is the case, additional equivalence trials, perhaps with multiple active controls, may be needed to verify the results.

References:

Interim Analysis and Early Termination in Clinical Trials
Leader: Sung C. Choi
Virginia Commonwealth University

Several subjects related to the discussion topic were addressed by the 11 attendees who came from a variety of different types of organizations (e.g., academia, pharmaceutical company, CRO, and medical research centers).

Purpose and aims
The purposes of the interim analysis include: chance for early termination of clinical trials, futility analysis, and possible changes in protocols such as the sample size. Further, specifying the interim analysis in the protocol can prevent irrational stopping of trials.

Type of interim analysis
The general consensus was to use group sequential tests (GST) with boundaries given by O’Brien and Fleming or other similar boundaries.

Frequency of interim analysis
The number of interim analyses would vary among trials. However, discussants generally agreed that the number should be no more than five. Some published results seem to indicate that it might not be advantageous in terms of the power function to plan more than five interim analyses for most trials in practice.

Other analyses following the interim analysis
Determination of p-values, confidence intervals for the endpoint parameters, and analysis of the secondary endpoints when trials are based on group sequential tests were discussed. In particular, the needs and difficulties of such analysis were discussed.

Problems with interim analysis
One problem is that the sample size might be insufficient to analyze other endpoints if trials are terminated early using GST. Another problem occurs when the accrual rate is high while the response time to observe outcomes is long. In such trials there would be little or no benefit of performing an interim analysis. In some trials, only the interim analysis in the early stages of the trials can still be performed. The spending function can be useful in designing unevenly spaced interim analyses.

Futility analysis
The futility analysis in clinical trials refers to testing for possible early stopping to accept the null hypothesis. The participants generally agreed that some type of futility analysis should be incorporated in the design of trials. Such analysis can prevent prolonging the trials when evidence of the null hypothesis is strong.

Considerations in Designing Dose Response Studies
Leader: Naitee Ting
Pfizer

The following outline was used to help guide the discussion. Since the focus of the discussion was on design issues, analysis of dose response studies was not addressed extensively.

Questions which need to be considered in designing these studies:
- In which doses or dose regimens are we interested?
- How many treatment groups should we include in the first study?
- Do we need placebo as well as active control groups?
- Is it necessary to estimate a minimum effective dose (min ED)? If so, how to define min ED?
- How to define a dose response relationship for the drug being studied?
- Sample size estimation
- How many studies do we intend to include so that a dose response curve can be characterized?
- In the case of multiple studies, do we consider designing a combination of studies ahead of time and analyzing them together when they are done? If so, how?
Factors which should be considered before designing Phase II dose response clinical studies:

- Regulatory environment
- Therapeutic area
- Characteristics of the test drug
- Prior knowledge
- Type of endpoint
- Method of analysis

One of the challenges is to develop a drug for the control of a life-threatening disease when the drug is known to be active, but the dosage is unclear. A possible solution to this problem is to design a Phase II/III trial with a few fixed doses against placebo for long term use. There are ethical concerns to this type of design, but in some special cases these designs are useful. For study designs of this nature, we will have to rely on a good understanding of pre-clinical results and Phase I clinical findings.

Another problem we discussed is how to design a study for OTC usage of a drug already on the market. In one example, the study drug includes three ingredients. The challenge was to find the optimal combination of these three ingredients, assuming the dose range of each ingredient is known. The question is how to combine them within the known dose ranges. One way to deal with this problem is to consider response surface methodology and design the study accordingly. In this type of study design, placebo needs to be included to examine inactive combinations. For studies of OTC drugs, cost is always an important concern. Various cost-effective dose combinations should be considered in these designs.

We spent some time discussing concentration-response designs. In this type of study, we measure the drug concentration levels from each patient's serum and analyze the relationship between these drug levels and efficacy responses. Knowledge of concentration-response relationship helps to fine tune the dosage of study drug for individual patients. Designs for concentration-response studies are, in general, similar to designs for dose response studies.

Dose response trials are typically designed to study the relationship between efficacy response and active drug dose (or drug concentration levels). However, the relationship between dose and drug safety or drug toxicity are also studied. In some special cases, trials are designed to study the dose and toxicity response. But often these studies are conducted at a later stage, not at early Phase II. Indeed the relationship between dosage and safety is an important concern for both the sponsors and the regulatory agencies.

One of the primary statistical concerns in designing a dose response study is sample size estimation. Two types of sample sizes can be considered: number of subjects per treatment group or number of treatment groups for the study. The number of subjects per treatment group is estimated based on the statistical method used for data analysis. However, the number of treatment groups is usually selected based on the study objectives and practical limitations.

We feel that, in general, dose response studies should be designed to compare fixed dose regimens instead of titration or flexible dose regimens. When there are only a few treatment groups in a study, people usually analyze the data using techniques to handle discrete treatment groups, such as ANOVA or ANCOVA. In cases, where more dose groups are tested in one trial or when the primary interest is the concentration-response relationships, regression methods may also be used to analyze these results.

If results from the first dose response study help identify a range of active doses, we often will need additional dose response studies to help fine tune the dose regimens. We may need to identify the minimum effective dose, maximum effective dose, and characterize the dose response curve, etc. One question is how can we optimize this process? Should we begin by considering a sequence of studies to address these issues over time? Should we start with one single study with lots of treatment arms and few patients per arm, perform an interim analysis, and then use the interim analysis results to produce the number of arms? In other words, should the strategy be to design a sequence of many studies or to design a sequential study which captures most of the information for our decision making process? Or should we do both?

We discussed the effort we should spend on characterizing dose-response relationships. Do we need to determine the minimum effective dose? How much effort can we afford to build a complete understanding of dose related issues? What about maximum effective dose? It can take a large amount of effort to quantify completely the dose-response relationships. It might be better to delay these additional efforts until after a preliminary set of dose response findings have been established. On this basis, we can further build up more details regarding the characterization of dose response relationships.

Another issue is the case where the dose response relationship is not monotonic. When this is suspected, the design is more difficult. We did not spend much time discussing this issue.

In general, the transition from Phase I to Phase II in drug development is very critical. At this stage of drug development, there is some information regarding maximum tolerable dose, but there are no clinical efficacy data. Hence the design of the first Phase II dose response study is a major challenge. How many doses should we study? Which specific doses should we select? Do we need a placebo and/or active control? How many patients should we include for each treatment arm? Do we use an equal number of patients per group? As more and more information is collected from Phase II studies and we know more about the study drug, this knowledge will help design future studies, including additional dose response studies.

**Multiple Endpoint Issues in Clinical Trials**

**Leader: George Chi**

**FDA**

The motivation for the problem which was stated as follows was described. If a primary endpoint and also a secondary endpoint are defined in the protocol, and if the primary endpoint fails to achieve statistical significance, can one test the secondary endpoint and make a claim based on the apparent significance of the p-value calculated?

Strictly speaking, in the Neyman-Pearson hypothesis testing framework, one has expended all the alpha (0.05 in this case) in testing the primary endpoint. Any further testing would result in an inflation of type I error unless one had prospectively defined a way for doing this so that the overall type I error is under control at the α-level.

The following proposal was discussed at the roundtable:

1. Define what is meant by a primary endpoint, a co-primary endpoint, and a secondary endpoint. This would help eliminate many incidental endpoints that should not have
been considered as primary endpoints. It also helps to focus on the clinical decision rule which will be defined in terms of the endpoints. These definitions have a regulatory perspective.

**Definition 1:** An endpoint is a primary endpoint if it satisfies the following conditions:

- It is a clinical endpoint that provides a measure of clinical benefit realized in the patient that is acceptable by the clinician as a meaningful measure of the drug effect for the disease under treatment, and furthermore
- It is an endpoint such that a positive finding in this endpoint alone is sufficient to result in the claim, regardless of what other endpoints show.

There can only be a few such primary endpoints, such as mortality.

**Definition 2:** An endpoint is a co-primary endpoint if it satisfies the following conditions:

- It is a clinical endpoint that provides a measure of clinical benefit realized in the patient that is acceptable by the clinician as a meaningful measure of the drug effect for the disease under treatment, and furthermore
- It is an endpoint such that a positive finding in this endpoint alone is sufficient to result in the claim, provided any primary endpoint(s) or other equally important co-primary endpoints are not showing a negative effect.

**Definition 3:** An endpoint is a secondary endpoint if it satisfies the following conditions:

- It is a clinical endpoint that provides a measure of clinical benefit realized in the patient that is acceptable by the clinician as a meaningful measure of the drug effect for the disease under treatment, and furthermore
- It is an endpoint such that a positive finding in this endpoint alone is not sufficient to result in the claim. However, it together with other secondary endpoints may provide sufficient evidence to result in a claim or it may provide supportive evidence to the primary or co-primary endpoints.

However, in the presence of a positive finding in the primary endpoint or co-primary endpoints, whether positive findings in the secondary endpoints can be used in a labeling claim need further discussion. If they can, at what level they should be tested is not clear.

2. Assuming that we accept these definitions of primary, co-primary, and secondary endpoints, then the question is if one has such a set of endpoints, then how do we handle the multiple endpoints testing problem. The key point is that generally one needs first to define exactly the clinical decision rule that will result in a claim. The hypothesis and type I error should reflect the clinical decision rule.

The actual discussion considered a very simple situation involving two endpoints, one primary and one co-primary. There are various clinical decision rules that may actually be used to test the hypothesis of no drug effect using these two primary endpoints depending upon the situation. One way is to consider both endpoints simultaneously as a bivariate distribution and perform a global test based on the joint distribution. Examples are the method proposed by Geller, Gnecco, and Tang or the O'Brien method. Of course, how one estimates the unknown correlation between the two distributions is unclear, and how one can obtain an \( \alpha \)-test for the test statistic derived from these two endpoints is not clear, particularly when the distributions are non-normal. Since the correlation needs to be estimated from the data, it seems that inflation in type I error may be expected.

In Vitro-In Vivo Relationships for Extended Release Drug Products

**Leader:** David Mauger

Penn State University

Most of the attendees were associated with the pharmaceutical industry, and several were with academic centers; none had current affiliations with regulatory agencies. There was general recognition that the statistical literature is relatively void of applied or methodological work on in vitro-in vivo correlations compared with the pharmaceutical literature, in contrast to related areas of biopharmaceutical research such as dissolution testing, bioequivalence, and bioavailability. The discussion leader used a draft of FDAs guidance for industry document entitled “Development, Evaluation, and Application of In Vitro-In Vivo Correlations” to stimulate discussion. The discussion focussed primarily on Level A in vitro-in vivo correlations as defined by the USP subcommittee on biopharmaceutics.

The first issue discussed was the use of in vitro-in vivo correlations to obtain waivers for demonstrating in vivo bioavailability or bioequivalence in the NDA/ANDA review process. There was general consensus that simply demonstrating a strong association, as evidenced by a high correlation coefficient, may not be adequate and that a predictive relationship which could be validated is necessary. There was not general agreement as to whether an in vitro-in vivo correlation could be validated internally (i.e., with the same test statistic) or if it should require external validation with other data. It was felt that a minimal requirement would be internal validation via cross-validation or similar measure, as opposed to a simple goodness-of-fit.

This led to a discussion of which measures of in vivo bioavailability should be predicted from in vitro dissolution. Possibilities considered included: 1) in vivo absorption as
estimated via deconvolution, which is the in vivo measure most frequently used and probably the most directly related to in vitro dissolution, but it may not be the most relevant measure of bioavailability; and 2) the plasma concentration profile, which has a clear relationship with the usual summary measures of bioavailability (AUC, CMAX), but requires sophisticated modeling strategies and assumptions which may be difficult to evaluate. This appears to be an area in which a study of the statistical properties of various methodologies for constructing in vitro–in vivo correlations would be most useful.

A question arising from this discussion was whether or not in vitro–in vivo correlation studies of sufficient strength to merit granting of a waiver for demonstrating in vivo bioavailability are an efficient use of resources. There was agreement that an in vitro–in vivo correlation study designed for use in obtaining such waivers would need to be more extensive than a study designed for use as an internal research tool. Several participants commented that the FDA seems to be encouraging the use of in vitro–in vivo correlations to obtain waivers even though the criteria used to evaluate in vitro–in vivo correlations seem to be less well defined than the criteria used to evaluate bioavailability.

ASA Fellows Committee

Larry Gould (Chair), Charles Goldsmith, & Bruce Rodda

The Fellows Committee is charged with recognizing the service and accomplishments of members of the Biopharmaceutical Section through election as Fellows of the American Statistical Association. This charge is carried out by encouraging the nomination of members of the section deserving of this recognition. Last year (1997), the Committee encouraged the nomination of Biopharmaceutical Section members for the first time. Six members of the Biopharmaceutical Section were elected to fellowship; most of the nominations resulted from proactive action by the Committee.

To follow up on this encouraging result, the Committee obtained guidance from members with experience in the Fellowship election process and combined this information along with ASA recommendations into a package to guide potential nominators as to what was needed for a good nomination package. The package was distributed to the Section membership represented on the electronic mailing list and was announced at the 1997 Annual Business Meeting of the Section. In addition, there are people (e.g., members of the Fellows Committee) who can provide help and answer questions about the process.

A number of nominations have been submitted for the current year. At least two of these have come about as a direct result of the membership being made aware of the fact that Section members can be (and have been) elected to fellowship and the provision of explicit guidance as to how to proceed.

We probably could have had more nominations (and nominators) had the information been available sooner, but lead times for dissemination via newsletter or mailings turned out to make this impractical. However, the package for 1998 nominations should work for 1999 nominations, too, with the only change probably being the name of the Chair of the ASA Fellows Committee.

There certainly are members of the Section whose contributions to Statistics in terms of service to the profession, research, applications, teaching, and consulting deserve to be recognized by nomination to fellowship. There is a track record of successful nominations of worthy candidates. All that's needed to succeed this year, and in succeeding years, are recommenders willing to identify (and, even better, pursue the nomination of) deserving candidates.

Congratulations 1997 ASA Fellows!

Congratulations to the six members of the Biopharmaceutical Section who were elected as Fellows in 1997! The new fellows and their citations are:

- **Charles Ralph Buncher**, Professor of Biostatistics and Epidemiology, University of Cincinnati College of Medicine. For three decades of creative statistical applications in the environmental, imaging, and pharmaceutical areas; for administration of major statistics studies; and for excellence in teaching.
- **Vernon M. Chinchilli**, Professor of Biostatistics, College of Medicine, Pennsylvania State University: For insightful contributions to the analysis of longitudinal data; and for excellence in statistical education and service to the profession, both in the American Statistical Association and to government agencies.
- **Alan B. Forsythe**, Senior Director of Corporate Biomedical Information, Amgen, Inc.: For applications of statistics in medical and biotechnological research; for innovative contributions to linear models; and for excellence in teaching and the guidance of students.
- **Joseph F. Heyse**, Senior Director, Vaccine Biostatistics and Research Data Systems, Merck Research Laboratories: For excellence in the application of statistical methods to preclinical and clinical trials; for seminal work in the application and dissemination of statistical methods to health economics and outcomes research; and for service to the profession.
- **Robert L. Obenchain**, Research Scientist, Statistical and Mathematical Sciences, Eli Lilly and Company: For excellence in consulting in the telecommunications and pharmaceutical industries; for contributions to the maximum likelihood theory of shrinkage estimation; for development of freeware statistical software systems; and for distinguished service to the profession.
- **Robert R. Starbuck**, Assistant Vice President, Clinical and Biostatistics and Data Management, Wyeth-Ayerst Research: For significant contributions to the management of statistical and allied disciplines in the pharmaceutical industry; for leadership in developing alliances among industry, academia, and government statisticians; and for service to the statistics profession and the pharmaceutical industry.
Section News

Minutes of ASA
Biopharmaceutical Section

Executive Committee Meeting

October 29, 1997, Bethesda, Maryland

Bob Davis introduced Ralph Harkins, who was newly appointed to the Executive Committee. Davis reviewed 1998 appointments. The Associate Editor of the Biopharmaceutical Report has to be appointed, and appointments to the Student Paper Awards Committee have to be finalized. Davis reviewed the volunteer list. Davis announced that the section made a $500 donation to the student analysis competition to be held at Winona State University. It was decided we would try to schedule the next meeting of the Executive Committee on Tuesday morning, March 31, 1998, during the ENAR meeting in Pittsburgh, Pennsylvania. Monday morning is the second choice.

The Section voted to thank Gary Neidert for his contributions to the Section during his time on the Executive Committee. Specifically, he:
- Rewrote the Manual of Operations to make it a useful guide for the Executive Committee,
- Reestablished the Continuing Education Committee and appointed an outstanding leader, who launched the successful fall workshops,
- Established the Fellows Committee,
- Put together an outstanding slate of candidates for 1998 Section officers,
- Served as a valuable mentor to Bob and the rest of the Executive Committee during 1997.
- The Section also recognized Lianng Yuh for his contributions to the 1997 programs and Sheila Chow for handling the Best Contributed Paper awards for the last two years.

JSM Meeting Minutes
The minutes of the August 11 meeting held at the Joint Statistics Meetings in Anaheim, California were approved.

Treasurer’s Report
Jeff Meeker distributed a draft statement of income and expenses of the Section through September, 1997.
Several solicitations for donations were discussed, including a Young Researchers luncheon at WNAR, the Student poster/competition, and an informal contact for Statistics Day. The Section decided against making any further contributions.
Meeker presented the current proposal for the 1998 budget. It was decided the price of the 1998 Proceedings would be increased to prior levels, so the Section would again show a profit. Jeff indicated his budget would show a profit from the Workshop in 1998. Also discussed was an increase in dues from $1 to $2. It was decided not to pursue an increase for 1998, but an increase for 1999 will be placed on the agenda for the March, 1998, meeting.
Assignment: Meeker will finalize the budget and submit it to ASA.

Manual of Operations Update
Meeker presented the proposed update in the Manual of Operations for the Biopharmaceutical Section associated with the addition of the Student Awards Committee. The changes were approved.
Assignment: Meeker will finalize the changes and distribute the Manual of Operations.

1997 Workshop
Christie Chuang-Stein reported on the 1997 workshop held in a Young Researchers luncheon at WNAR, and an informal contact for Statistics Day. The Section decided against making any further contributions.
Meeker presented the current proposal for the 1998 budget. It was decided the price of the 1998 Proceedings would be increased to prior levels, so the Section would again show a profit. Jeff indicated his budget would show a profit from the Workshop in 1998. Also discussed was an increase in dues from $1 to $2. It was decided not to pursue an increase for 1998, but an increase for 1999 will be placed on the agenda for the March, 1998, meeting.
Assignment: Meeker will finalize the budget and submit it to ASA.

1997 Best Contributed Papers
Heft announced the Biopharmaceutical Section Best Contributed Paper Awards:
- First Place: Devan V. Mehrotra. ANOVA with Unequal Variances—Correcting a Popular Strategy.

1998 Workshop
Bob Small reported that the FDA has an internal 2-day seminar they hold each fall which is attended by 80–100 FDA statisticians. They proposed merging their seminar with the Biopharmaceutical Section workshop for 1998. They have already scheduled meeting rooms on the nearby NIH campus for September 17–18. They also expect to be able to obtain a grant to pay for registration of FDA statisticians, and expenses and honoraria for big name academic speakers. The program committee would include Small, Ralph Harkins, and Sandy Heft from the Biopharmaceutical Section plus others appointed by the FDA.
Note: Since the Executive Committee Meeting, the date of the seminar has changed. Please see page 3 of this newsletter for the current information and dates.
Third Place: Ronald W. Helms. Baseline Values are Random, Too: Using Baseline in Mixed Models.
Sandy also presented Guidelines for Selection of the Best Biopharmaceutical Contributed Paper Award at the Joint Statistical Meetings. The guidelines were approved.
Assignment: Sally Greenberg will post the guidelines on the Section Web page.

1998 ENAR program
Tom Capizzi reported the sessions for 1998 ENAR meeting are those reported at our August 11 meeting. The deadline for abstracts to be submitted is November 10.

1998 Joint Statistical Meeting Program
Capizzi reported the three invited sessions for the 1998 Joint Statistical Meetings are those reported at our August 11 meeting. There are two additional sessions which he intends to submit to the competition: one on drug stability organized by David Pack and one on permutation tests organized by Vance Berger. New rules were instituted by the Committee on Meetings: 1) an individual can only speak once, whether a contributed or invited paper or as a discussant, 2) contributed paper sessions will include seven 15 minute papers and papers from another section could be added to complete the session, and 3) invited paper sessions must have four participants while special contributed paper sessions will consist of five papers and an optional discussant.
Capizzi indicated he has three proposed special contributed paper sessions, including two from the FDA.
It has been reported that the number of invited papers allocated to a section at the Joint Statistical Meetings is based on section membership. However, there is little correlation between the number of sessions allocated for the 1998 meetings and membership data.
Assignment: Davis will draft a letter to Robert Mason, chair of the Committee on Meetings, with a copy to Ray Waller raising this issue and asking for clarification. The letter is to be co-signed by Capizzi and Chuang-Stein.

1998 Short Course
Capizzi has submitted the course Design and Analysis of Clinical Trials based on the text by J. P. Liu and Shein-Chung Chow.
The Committee discussed and approved a request from the Section on Bayesian Statistics that we co-sponsor a one-day short course Introduction to Bayesian Methods in Biostatistics based on the text Bayesian Biostatistics co-edited by Darlene Stangl and Don Berry. The course would be taught by Stangl.

1998 Best Student Paper Awards
Letters announcing the 1998 awards need to be sent to schools. Capizzi agreed to handle this until a Committee is appointed.
Assignment: Davis will contact Denise Roe to make sure an announcement of the Student Paper Contest is submitted to Amstat News by November 1 for both the December and January issues.

1998 Round Tables
Richard Entsuah submitted a list of 15 proposed topics for round table discussions at the 1998 Joint Statistics Meetings. Attendees were asked to indicate five topics to delete to narrow the list to 10.

Biopharmaceutical Report
Curtis Wiltse reported one more issue is planned for this year. The lead article will be on the Section membership survey by Phil Pichotta. The first issue of next year will include a summary of the workshop on FDA/Industry Partnership. An Associate Editor for next year is still to be appointed.

Council of Sections
Chuck Davis sent a letter to Mason on our proposal to referee abstracts for the Joint Statistical Meetings. There wasn't any support for this proposal in the Council of Sections.

Fellows Committee
Larry Gould has developed a package for nominee sponsors. The package should be made available on the Web site. Final packages for this year's nominations are due March 1.

Submissions to Amstat News
Roe has developed an annual schedule for proposed articles for Amstat News.

Web Site
Greenberg reported she is up to date with material that has been submitted except for one old issue of the Biopharmaceutical Report. She will be adding photographs of Executive Committee members, as well as some photographs from the workshop. She asked if anyone wanted to provide a brief write-up of the workshop to go along with the pictures.

Electronic Mail List
Greenberg reported there are currently 110 subscribers. There was a discussion as to whether the section would allow postings from recruitment companies. It was decided the ads should contain enough information to be useful to members, the same position could not be posted more than once, and the words “job announcement” would be in the title. The ads would be placed through the mail list moderator.
Assignment: Greenberg will draft a policy to include specific information to be included for the March meeting.

Membership Committee
Phil Pichotta reported the survey report has been completed and will appear in the next issue of Biopharmaceutical Report. An updated mailing list was obtained from ASA, which indicated a net increase of 61 members since the last list. Pichotta is currently sending welcome letters and a brochure to all new members.
Pichotta contacted Boris Iglewitz concerning the proposed salary survey by the Career Development Committee. The survey is now under the SPRAIG Initiative. Pichotta is now chair of the committee to do that survey.
Pichotta is currently developing plans for a membership drive for the section.

**Assignment:** Pichotta is to develop a proposal for attracting current ASA members into the Section for the March meeting.

**Midwest Biopharmaceutical Statistics Workshop**

Meeker distributed for liaison Jim Bergum the current program for the Midwest Biopharmaceutical Statistics Workshop.

**New Business**

The Executive Committee thanked Bob Davis for his leadership as chair during 1997.