2006 FDA/Industry Statistics Workshop
Session Abstracts

GS1 Statistics in the FDA and Industry: Past Present, and Future

Organizer(s): Ken Koury, Schering-Plough; Estelle Russek-Cohen, FDA; Greg C.G. Wei, Pfizer)
Chair: Mohan Beltangady, Pfizer

In the centennial year of the FDA (the original Food and Drugs Act of 1906), this session reflects on the application of regulation in response to public health issues, and how the recognition of new medical and societal needs has influenced the design, conduct, analysis and interpretation of clinical studies. Initially, manufacturers were required only to show the safety of their products in order to obtain FDA approval for marketing. Perhaps the most substantial influence came in 1962 when the Kefauver-Harris Drug Amendments were passed to ensure drug efficacy and greater drug safety. For the first time, drug manufacturers were required to demonstrate the effectiveness of their products before marketing them. Patient package inserts were first required in 1970, and rapid changes occurred during the 1980s and beyond based on the Belmont Report, the Orphan Drug Act, the Childhood Vaccine Act, provisions for expanded access to experimental drugs, accelerated review of products for life-threatening diseases, improved assessments of response as a function of gender, the Pediatric Rule, and the Pediatric Research Equity Act. The evolution of the regulatory environment as FDA responded to the many clinical research issues has provided a very fertile ground for the application of statistics and statistical thinking, and it has created the opportunity for an enormous number of statisticians to specialize in this area. The importance of statistical contributions and innovation will continue to grow in the future as product development becomes more complex and greater emphasis is placed on flexibility and efficiency. The perspectives of statisticians from FDA, industry and academia, as well as the importance and productivity of the interactions among these three groups will be presented and discussed during this session.

Statistics in the FDA and Industry
David DeMets, University of Wisconsin

The clinical trial, especially the randomized clinical trial, has been the central research method for evaluating new pharmaceuticals, biologics and devices as well as procedures and behavioral interventions. Despite past success, new challenges must be addressed for this success to continue. Examples include the training of the next generation of biostatisticians and clinical trialists, finding collaborative approaches between academia and industry, reducing costs for data collection and management, monitoring for long term safety or adverse events and finding a balance for researchers and sponsors for conflict of interest issues. In addition, the era of genomics and proteomics will require new statistical methods and innovations due to the vast amounts of data and multiplicities. These and other issues for the future of statistics and clinical trials will be discussed.

FDA History and Statistics Intertwined
Mary Foulkes, FDA

Legislation (e.g., the original Food and Drugs Act of 1906) in response to public health issues, in recognition of new medical and societal needs, has influenced the design, conduct, analysis and interpretation of clinical studies. Perhaps the most substantial influence came with the 1962 Kefauver-Harris Drug Amendments, then drug manufacturers were required to prove the

effectiveness of their products before marketing them. The 1980s and beyond brought rapid changes: revise regulations for human subject protections based on the Belmont Report; the Orphan Drug Act, enabling research and marketing of products treating rare diseases; expanded access and accelerated review of products for life-threatening diseases; and the Pediatric Rule, requiring sponsors of selected new and extant drug and biological products to conduct studies to assess their safety and efficacy in children. These changes or evolutions have provided new opportunities for statistical contributions to clinical research advances, or have initiated the development of new methodologies. This history and its interrelation with statistical issues will be considered.

GS2 Flexibility in Clinical Trials: How Do We Deal With It?

Organizer(s): Shein-Chung Chow, Duke; Shuyen Ho, GalxoSmithKline; Qian Li, FDA; Jerry Schindler, Cytel
Chair: Sue-Jane Wang, FDA

The Adaptive Clinical Trial has emerged as a potentially powerful tool to help improve the efficiency and success rate of the drug development process. Recent advances in statistical methodology now permit the modification of specific aspects of the design of an ongoing trial while preserving the overall validity of the analyses. Pharmaceutical companies are planning for the broader implementation of adaptive clinical trials. This session will discuss lessons learned from our experiences implementing these trials. We will review the steps required for statisticians within pharmaceutical companies to design, review, and analyze flexible clinical trials. We will also discuss the need for simulations in the development of adaptive trial protocols, issues and concerns from the FDA in evaluating clinical trials with adaptive designs, as well as problems and their solutions encountered during the adaptive trial process.

Just How Flexible Can a Prospective Clinical Trial Be?
Don Berry, Berry Consultants

I will argue that (i) clinical trials should be prospectively designed and (ii) prospective trials are inherently inflexible. That doesn't mean that they cannot be arbitrarily complicated. But they can be run by automata. I will further argue that clinical trials are becoming and will continue to become more complicated. Trials with simple designs that ignore interactions between treatments and patient covariates are inefficient and their use hinder finding treatments that benefit patients with heterogeneous diseases such as cancer. The path is treacherous with mines every step of the way. Avoiding the mines is challenging, but enormously rewarding ... and fun!

Considerations in Enhancing Flexibility of Clinical Trial
H.M. James Hung, FDA

Enhancing flexibility of clinical trial designs is one of the hot topics nowadays. In past decades, the classical design has been improved to allow the flexibility for terminating the trial early if the experimental treatment is proven effective or deemed harmful or even futile, based on the accumulated data during the course of the trial. Statistical validity of such an enhanced design in terms of type I error is maintained. The operational aspects of this design can still be an issue but, by and large, there have been many good models for how to deal with these aspects. As the flexibility of trial designs is enhanced further, the potential risk of the resulting trial being not interpretable increases. In this presentation we shall share our review experiences, discuss the many issues arising from use of more flexible designs and stipulate some alternative strategies to improve clinical trial designs and programs.

Strategies for Implementing Adaptive Clinical Trials
Jerald Schindler, Cytel

Adaptive drug development has the potential to improve the success rate of the drug development process, reduce the total number of patients required in clinical trials, and improve the quality of the clinical information obtained. Over the past few years, statistical methods have been developed to help preserve the validity and integrity of the analysis of adaptive trials. Now companies are developing strategies to incorporate adaptive methods into their clinical development programs for their entire drug portfolio. This talk will present an overview of some of the steps required in the development of an adaptive strategy for drug development. In particular, the need for careful planning, simulation, and co-ordination of various steps in the clinical development process will be highlighted. Changes that can be made now to implement these strategies today will also be discussed.

GS3 - Surrogate Endpoints and Accelerated Approval

Organizer(s): Alex Bajamonde, Genentech; Lilly Yue, FDA

Clinical trials designed to definitively assess improvements in clinically meaningful endpoints will typically require enrolling many patients and then subsequently following them for a long time. Hence, there has been great interest in developing strategies for reducing the time and cost of drug development. One such strategy is the use of surrogate endpoints either in proof-of-concept trials or in label-enabling trials. The validation of surrogate endpoints has been studied relatively intensively in the literature (eg, Prentice 1989; Freedman, Graubard, and Schatzkin 1992; Fleming and de Mets 1996). Efforts have been made to converge to a common framework, encompassing the wide variety of settings one can encounter. Recent developments including variance reduction factors, sample size assessment methodologies, and surrogate threshold effects will be discussed. Examples of drug development programs that incorporate the use of biomarkers and surrogate endpoints to accelerate the evaluation of new drugs will also be discussed. These examples include: (1) use of a biomarker implicated directly in the purported mechanism of action of the candidate drug to form nested hypotheses to demonstrate efficacy, and (2) planned analyses of a surrogate endpoint for accelerated approval (e.g., progression-free survival in certain oncology settings) with a subsequent planned and appropriately powered analysis on a definitive endpoint (e.g., overall survival) for full approval. Technical, logistical/tactical, and regulatory considerations will be elaborated on. Finally, regulatory perspectives on the use of surrogate endpoints will be discussed.

Biomarkers and Surrogate Endpoints in Drug Development: Technical and Regulatory Considerations
Gracie Lieberman, Genentech

Clinical trials designed to definitively assess improvements in clinically meaningful endpoints typically require enrolling many patients and then following them for a long time. Hence, there is great interest in developing strategies for reducing the time and cost of drug development. One such strategy is the use of surrogate endpoints either in proof-of-concept or label-enabling trials. The surrogate endpoints used can be those already well defined in clinical practice and used by clinicians to monitor and treat patients, or new mechanism-driven biomarkers. This is particularly applicable to Oncology applications where the number of molecular-targeted agents under clinical development keeps growing. We will discuss two Oncology examples (Herceptin and Iressa) in which efficacy surrogate endpoints were used for approval; we will discuss the challenges and risks. We will discuss how the same surrogate endpoints are being used in proof of concept trials

and we will assess the limitations of such strategies. We will then assess how mechanism-based biomarkers and new imaging techniques can be integrated into development programs; oncology and non-oncology examples will be discussed.

Surrogate Endpoints, A Regulatory Perspective
Greg Campbell, FDA

With the proliferation of biomarkers in this age of genomics and increasingly sophisticated imaging modalities, it is not surprising that the FDA has turned its attention to both surrogate endpoints and biomarkers. While many biomarkers are mentioned on the FDA's Critical Path Opportunities List, the lone mention of surrogates is in Surrogate Outcomes for Cardiovascular Drug Eluting Stents. While it is clear that if one were to impose the criteria of Prentice (1989) for the validation of surrogate endpoints very few such surrogates would be eligible, the use of late loss for coronary drug eluting stents as a surrogate for target vessel revascularization may be a particularly interesting example. Different levels of generalizability for a potential surrogate are mentioned, from the global surrogate that serves for all medical products and all indications to the very narrow that may only be appropriate for a single medical products company and a single indication. In addition, different types of surrogates are identified. Of particular interest are intermediate temporal endpoints that can in some cases be surrogates for longer term clinical primary endpoints.

Developments in the Statistical Evaluation of Surrogate Endpoints in Clinical Trials
Geert Molenberghs, Universiteit Hasselt, Diepenbeek, Belgium
Ariel Alonso, Universiteit Hasselt, Diepenbeek, Belgium

The validation of surrogates has been studied by Prentice who presented a definition of validity as well as a formal set of criteria that are equivalent if both the surrogate and true endpoints are binary. Freedman et al supplemented them with the proportion explained, the fraction of the treatment effect mediated by the surrogate. Noting operational difficulties with the PTE, Buyse and Molenberghs proposed to use jointly the within-treatment partial association of true and surrogate responses, and the treatment effect on the surrogate relative to that on the true outcome. In a multi-center setting, these quantities can be generalized to individual-level and trial-level measures. Buyse et al proposed a meta-analytic framework. Several extensions have been formulated and unifying attempts have been made. This includes a so-called variance reducation factor and an information-theoretic approach. Work has been done to convert the evaluation methodology to sample size assessment methodology.

GS4 Interpreting Subgroups for Regulatory Purposes

Organizer(s): Jeff Helterbrand, Genentech; Lakshmi Vishnuvajjala, FDA-CDRH

A single primary analysis population, often the intention-treat-population, is typically pre-specified for label-enabling clinical studies. However, at the time of analysis, many subgroup analyses are performed and the information they provide can impact regulatory decisions. In this session, the speakers will review many issues/controversies surrounding the interpretation of subgroups and review some case studies. The session will close with panel Q&A.

Frequentist and Bayesian Subgroup Analysis at the Center for Devices and Radiological Health
Gene Penello, FDA-CDRH
Harry Bushar, FDA-CDRH

Subgroup analysis, whether pre-specified or post hoc, can be both a bane and a virtue. It can be bane because an analysis plan to adjust for multiplicity is required, which can siphen away power to detect important effects. It can be virtue in that the strengths and weaknesses of a medical product's safety and effectiveness are usually be best revealed by interrogating the data with a variety of analyses, including subgroup analysis. In the Center for Devices and Radiological Health (CDRH), subgroup analysis has been considered from both a frequentist and Bayesian perspective. An example of a post hoc frequentist subgroup analysis will be given. The analysis failed when the overall analysis was included as one of the subgroups tested. In the second part of the talk, Bayesian subgroup analysis will be discussed, in which subgroups are considered a priori exchangeable. The discussion will encompass (1) Bayesian trials for overall and subgroup testing, (2) operating characteristics, and (3) how a post hoc finding can be utilized as prior information for a new study designed to confirm that finding.

Subgroup Analyses: Can we 'Smooth' out the Rough Edges?
Daniel Sargent, Mayo Clinic

Subgroup analyses in clinical trials are simultaneously essential for scientific discovery (hypothesis generation), required for ethical and administrative reasons (differential efficacy or toxicity in subsets), and vexing (extreme difficulty in interpretation). The greater understanding of biology, in cancer in particular, is leading to greater segmentation of the population. New agents may be tested based on biology, not tumor site, or have efficacy in only a portion of a population (defined by a biologic marker). Additionally, technology is allowing massive multiplicity in markers explored, further complicating the issue. After motivating with several examples, in this talk I will present a) practical considerations regarding the interpretation of subgroup analyses, and b) statistical methods based on hierarchical models (allowing an overall mean and a subgroup specific effect), and/or shrinkage estimators (allowing interactions in the model but then shrinking them) that may assist in the analysis and interpretation of subgroups questions.

Perils of Subgroups: Concerns, Examples, Alternatives
Andreas Sashegyi, Eli Lilly

While the statistical limitations of subgroup analysis have been well-characterized and widely discussed for some time, the focus on subgroups remains a common and prominent component in the presentation and interpretation of clinical trials. The danger of subgroup analysis to help target treatment is pitted against the reluctance to apply overall results of large trials to individual patients, without regard to determinants of individual response. We present a number of recent examples that illustrate the challenges and argue for a shift in focus away from subgroup analysis toward patient-focused risk/benefit analysis.

GS5 Data Monitoring Committees: Getting a New Perspective on an Old Issue

Organizer(s): Paul Gallo, Novartis; Karen Kesler, Rho, Inc.; Matilde Sanchez, Merck; Jason Schroeder, FDA-CDRH
Chair: H.M. James Hung

Data Monitoring Committees (sometimes known as Data and Safety Monitoring Boards) provide the vital service of objectively representing the patients' interest in clinical trials. However, this responsibility often raises tough ethical and statistical questions that concern industry, regulatory, and academic statisticians. In this session, the speakers will present these issues in detail, drawing on their personal experiences. We will also discuss evolving approaches for dealing with

these issues in the area of adaptive designs. In a discussion after the presentations, Susan Ellenberg will provide her views on points made by the speakers.

DMC: The View from the Academic and Consultant Statistician
Ralph D'Agostino, Boston University

Data Monitoring Committees (DMC) have become a major component of clinical trials. Statisticians have played an important role in their development and implementation. They have been, in part, responsible for the success and influence of the DMC. The present talk reviews the role of the academic statistician in the various and ever changing roles regarding the DMC ranging from being a member, to being the non-member independent statistician, to being asked to review critically the DMC's decisions, and to being on steering committees dealing with the DMC. Personal examples from the FDA, NIH and Industry including cardiovascular and Vioxx trials will be used for illustration. Personal opinions of how to deal with these roles and how useful and successful statisticians have been will be presented.

Confidentiality and Trial Integrity Issues for Monitoring Adaptive Design Trials
Paul Gallo, Novartis

Adaptive design trials raise important issues with regard to the processes for review of interim data and implementation of adaptation decisions, while avoiding bias and maintaining interpretability of trial results. We discuss the issues, distinctions versus more familiar monitoring situations, various aspects of operational models for data monitoring, and types of adaptations which may be more or less prone to concerns about bias.

DMCs: An FDA Division of Cardiovascular Devices Perspective
Bram Zuckerman, FDA

Data Monitoring Committees (DMC) are an important part of many clinical trials. FDA has recently published a final guidance document on Data Monitoring Committees that applies to all centers within the agency. This talk will review recent application of this guidance to a review division in the Center for Devices and Radiological Health (CDRH). Specific DMC challenges for device clinical trials as well as generic issues will be commented on.

Discussant: Susan Ellenberg, University of Pennsylvania

PS01 The Role of the Statistician in Post-Marketing, Including Surveillance

Organizer(s): Harry Bushar, FDA-CDRH; Greg Campbell, FDA-CDRH; Gosford Sawyer, InnovaStat LLC

The statistician's role in clinical development in industry and regulatory settings is generally well understood. However, when it comes to post-marketing surveillance or the medical device industry, it is unclear what role industry and FDA statisticians play. Recent issues with approved pharmaceuticals and medical devices highlight the importance of the statistician's role as part of a multi-disciplinary team in assessing post-market data for signal/noise inference. Risk/benefit and health outcomes data are becoming increasingly important components in determining public health/reimbursement policy. This session will provide some perspectives on the role of statistics

in post-marketing pharmaceutical and medical device trials and will challenge statisticians to broaden their sphere of influence in post-marketing surveillance issues.

CDRH Postmarket Transformation and the Role of Statistician
Danica Marinac-Dabic, FDA

The CDRH postmarket responsibilities include monitoring of the devices and radiological products for continued safety, effectiveness and reliability, helping provide science based information to the public and remedial action as needed. In continuing efforts to improve its postmarket programs, CDRH has initiated postmarket transformation and identified specific steps to increase its ability to identify, analyze and act on postmarket information in order to improve the safety and effectiveness of medical devices and radiation emitting products. One of the goals of the postmarket transformation is to "develop a world class data sources and systems" related to the field of medical devices. The CDRH Epidemiology Program and CDRH Biostatistics Program play essential role in achieving this transformation goal. The strong collaboration between the two programs is essential in ensuring the Total Product Life Cycle approach, effective post-approval studies program, development of enhanced and active surveillance programs and data mining. The best practices, opportunities and the challenges of the multi-disciplinary approach to the postmarket transformation will be discussed.

Statisticians in Medical Device Post-Market Studies
Andrew Mugglin, University of Minnesota

Post-market studies of medical devices have become much more common in recent years. It is difficult to speak for the entire world of medical devices, so this talk will focus specifically on the speaker's experience with studies of implantable cardiac pacemakers and defibrillators. Post-market studies of these devices have rapidly become much larger and more complicated than they previously were, falling into several recognizable categories. In this talk we will discuss this particular field of medical devices, the various categories of post-market studies, and the roles that statisticians play in each. It is apparent that statisticians can play a larger and more strategic role in this arena than they often do.

Discussant: Jay Herson, Johns Hopkins University

PS02 - Biomarker Analysis

Organizer(s): Aloka Chakravarty, FDA; Mark Chang, Millenium; Kalyan Ghosh, Merck

Biological marker or biomarker, as more commonly known, refers to a variety of physiologic, pathologic, or anatomic measurements that are thought to relate to some aspect of normal or pathological biologic processes (Temple 1995; Lesko and Atkinson 2001). In recent years, it has received wide interest as a crucial element of the FDA Critical Path initiative. The session will cover the statistical basis of biomarker theory - why it is considered in clinical trials, its advantages and pitfalls, and situations in which use of a biomarker should be considered (or not). In particular, validation of a biomarker and its escalation potential to a surrogate marker will be of special interest. Issues such as biomarkers for patient selection as well as endpoint evaluation will be covered. Experiences in pharmaceutical industry as well as regulatory issues will be discussed with case examples. At the conclusion of this session, the attendees are expected to

have a broad-based theoretical introduction to biomarker use, practical and logistic knowledge of issues in using a biomarker in a clinical trial, and an orientation to the regulatory implications of biomarker use.

Opportunities and Challenges in Utilizing Biomarkers for Drug Development
Mark Chang, Millenium

Compared to a gold standard endpoint, such as survival, a biomarker can often have following characteristics: (1) being measured earlier, easier, and more frequently, (2) less subject to competing risks, less affected by other treatment modalities, (3) a larger effect size. The utilization of biomarker could lead to (1) better target population, (2) larger effect size, (3) smaller sample size, (4) faster decision-making.
The talk will cover the following:

- Opportunities of Enrichment Strategies with Biomarkers
- Prognostic and Predictive Biomarkers
- Challenges in Biomarker Validations
- Adaptive Design using Biomarkers
- Optimization using Bayesian Utility Theory
- Summary & Discussion

New Paradigms for Clinical Drug Development in the Genomic Era
Richard Simon, NIH

Genomic technologies today provide unprecedented opportunities to develop effective therapeutics. They offer opportunities to enhance the efficiency of clinical development and to target drugs to patients most likely to benefit from them. Targeting drugs to appropriate patients provides a rational strategy for limiting adverse effects and limiting escalating medical costs that threaten the medical and economic well being of individuals, companies and society. The effective utilization of genomic technologies in therapeutics development entails significant challenges and requires re-examination of familiar paradigms. I will try to present a roadmap for the development and validation of robust genomic biomarker diagnostics in conjunction with development of new drugs and will try to clarify many of the misconceptions that I believe exist about drug development in the genomic era.

Discussant: Sue-Jane Wang, FDA

PS03 Bridging Studies, Migration Studies, and Related Topics

Organizer(s): Jingyee Kou, FDA-CBER; Barry Schwab, Johnson & Johnson; Daniel Wang, Johnson & Johnson

In 1998, the ICH published its E5 Guideline for Industry on 'Ethnic Factors in the Acceptability of Foreign Clinical Data.' Since then, about 20 pharmaceutical products were approved in Japan using bridging strategies. In the US, use of bridging has been encountered in the submission of preventive vaccines and the extension of approved products to the pediatric population. For example, in the clinical development of influenza vaccines, a population might not have been included in the clinical endpoint efficacy study. Immunogenicity bridging studies can then be conducted to compare the immune response observed in the clinical endpoint efficacy study to

that elicited in other populations. E5 provides a guideline for using the bridging strategy to minimize duplication of clinical data and to allow for extrapolation of foreign clinical data to a new region. In practice, there are successful and failed cases as well as unsettled issues (clinical and statistical) when implementing the bridging strategy. In this session, the speakers will share their experiences, knowledge, and advances in this area.

Clinical Evaluation of Preventive Vaccines: Use of Bridging Studies
Marion Gruber, FDA-CBER

A bridging study is defined as a clinical trial in which a parameter of interest, e.g., a study population, manufacturing process, formulation, or dosing schedule, is directly compared with a different version of that parameter to assess the effect of the change on the product's clinical performance. A clinical bridging study may be performed to provide scientific support that efficacy data obtained from a particular study population can be extrapolated to a different population. In clinical bridging studies, the immune response and safety endpoints observed in the clinical efficacy studies are usually compared to corresponding endpoints in the bridging study. Immune response endpoints often serve as primary endpoints. This session will discuss regulatory considerations and approaches concerning bridging studies. In addition, examples of vaccines licensed using clinical bridging studies will be provided.

Transition between Bridging and Global Clinical Trials - T2DM as a Model for Development
Eric Lewis, GalxoSmithKline

In 1998, the ICH E5 document reached Step 4. This document provides guidance for the development of medications in ethnically diverse populations and the application of those data to "New Regions." The intent of the guidance was to decrease redundancy in clinical development. Bridging has its hurdles but it has basically been successful as demonstrated by the large number of medications "bridged" into Japan. Japan is but one example, there are many nuances depending upon the countries involved. Between 7 and 8 years after the implementation of ICH E5, there is now an excellent opportunity to design and implement truly global development programs. Type 2 diabetes mellitus provides a model for this development. This presentation will address some of the factors and opportunities for global diabetes development.

Bridging Studies and Global Drug Development
Weichung Joe Shih, UMDNJ

It has been over a decade since the proposal of ICH-E5 was initiated in 1992, and about eight years since E-5 was ¡¥finalized¡¦ in 1998. Issues around E-5 have started to surface since its publication. Statisticians have offered various proposals regarding the interpretation and methodology needed to implement E-5 at the symposiums held in Taiwan, Thailand, Japan, USA and elsewhere. In this talk, I will review some of the major methods and argue that these methods are respectively suitable for three different situations of the cross-country drug development.

PS04 - Statistical Issues in Medical Device Trials

Organizer(s): Greg Campbell, FDA-CDRH; George Koustenis, FDA-CDRH; Gosford Sawyer, InnovaStat, LLC

Clinical trials that evaluate the safety and effectiveness of medical devices offer unique challenges to the clinical and statistical communities, as well as the medical device industry. These devices include drug eluting (coated) stents, cardiac pacing systems, heart valves, cardioverter defibrillators, cochlear implants, ophthalmic laser systems, total hip replacement systems, robotic surgical systems, left ventricular assist devices, transmyocardial revascularization lasers, and many other applications across all fields of medicine. Topics related to the design, analysis and interpretation of medical device clinical trials will be discussed in this workshop including: pooling of foreign and domestic data, the challenges of historical controls and the use of Objective Performance Criteria (OPC) in non-randomized studies, and the incorporation of prior information (Bayesian approaches). A general question and answer session will follow the presentations.

The Use of Objective Performance Criteria in Medical Device Studies
David Breiter, Guidant

Objective performance criteria are commonly used in non-randomized trials to assess the safety and effectiveness of medical devices. In general, reasons such designs are incorporated include ethical considerations; the lack of an appropriate established therapy or diagnostic test for comparison; and the availability of data to establish an industry wide bar for performance, such as with prosthetic heart valves. In addition, the nature and pace of medical device innovation, rapidly turning out improvements and next generation devices, lends itself to the establishment of objective performance criteria. Usage and difficulties encountered with respect to objective performance criteria will be discussed in this presentation.

The Use of Objective Performance Criteria (OPC) in Medical Device Evaluation
George Koustenis, FDA-CDRH

The purpose of this presentation is to provide a brief introduction to the possible use of Objective Performance Criteria (OPC) in the evaluation of medical devices during the regulatory approval process. Advantages and disadvantages of OPC's will be discussed along with general criteria as to when OPC's might be used, and how they should be developed. The talk will conclude by calling for a dialogue between FDA, the medical device industry, clinical medicine, and other interested parties to develop a general guidance for the use of OPC's in medical device evaluation.

Evaluating Poolability of Continous and Binary Endpoints
Roseann White, Abbott Vascular

The traditional methods of assessing poolability generally looks for a significant interaction effect between site and treatment. In the case of binary effects, many times too little data is available in order to assess a clinically meaningful interaction effect. In the case of some continuous endpoints, there is often too much data and an effect size that is clinically not meaningful is statistically significant. This paper proposes a boot strap method aimed at assessing whether there is a statistically significant and clinically meaningful interaction effect.

Discussant: George Woodsworth, University of Iowa

PS05 - High Dimensional Expression Data: Consistency Across Platforms and Statistical Prediction Modeling

Organizer(s): Fred Immerman, Wyeth; Sue-Jane Wang, FDA-CDER

High-dimensional technologies for measuring gene and protein expression are being widely used in the quest to develop predictive biomarkers. Data quality and cross-platform consistency issues associated with these rapidly evolving technologies impact their ability and utility for predicting ultimate therapeutic responses. This session will address current efforts to characterize cross platform consistency and data quality, and examine their impact on the development of predictive statistical models. The utility will be illustrated by bridging the development from non-clinical discovery to clinical exploration.

An Analysis of the Microarray Quality Control Data

James J. Chen, FDA-NCTR

Recent completion of the MicroArray Quality Control (MAQC) project provides a unique opportunity to assess reproducibility of gene expression data across multiple sites and the comparability across multiple platforms. In the MAQC project, three metrics were considered in the analysis: differential gene list overlap, log ratio compression, and log rank correlation. The analysis of inter platform comparability was limited to the fold change between two samples, namely Samples A and B. We analyze the MAQC data using several measures to evaluate inter platform comparability and intra platform performance. The analysis include: 1) cross platform concordance and consistency; 2) platforms' variability, discriminability, and consistency; 3) sensitivity, specificity and accuracy in gene selection; 4) self-consistency of titration mixture; and 5) TaqMan and microarray platforms comparability.

Challenges In Progressing Biomarkers To Clinical Use - Proteomic Experiences

Chris Harbron, Astra Zeneca

The development of -omics technologies has led to the publication of a large number of predictive biomarkers in the literature. However this has not been reflected in a correspondingly large number of biomarkers being approved for clinical use. There are many contributory reasons for this discrepancy, including the necessity for high quality, reproducible data and the importance of carefully avoiding over-fitting within the modelling process. This presentation will illustrate some of these issues using a real example from an LC-MS proteomics platfrom.

Prognostic Model building with Biomarkers in PGx Trials: Lessons Learned from Oncology Case Studies

Li-an Xu, Bristol-Myers Squibb
Douglas Robinson, Bristol-Myers Squibb

There is great interest in using biomarkers, including genetic, genomic and proteomic data, to enhance understanding of the mechanism of action, pharmacodynamic properties, or ability to predict safety and response efficacy of investigational drugs. It is hoped that such biomarkers can be used to build prognostic models to select the patient population which may best benefit from investigational drugs. There are many statistical challenges for building prognostic models with clinical data, including 1) the data quality, 2) how to deal with the so called p>>n problem, where p is the number of predictors or biomarkers and n is the number of patients in the trial, 3) how to select biomarkers to be included in prognostic model building, 4) how to validate the sensitivity and specificity of the discovered biomarkers, and 5) how to link the pre-clinical findings to clinical results. In this presentation, we will use two examples from oncology clinical trials conducted at Bristol-Myers Squibb to address the challenges that accompany multi-gene predictive model building and some innovative approaches to handle them.

PS06 - Advantages and Challenges of Bayesian Clinical Trials

Organizer(s): Suman Bhattacharya, Genentech; Telba Irony, FDA

The use of Bayesian statistics has conspicuously increased in clinical trial research because it is ideally suited to adapting to information that accrues during a trial, potentially resulting in smaller, shorter, and more informative trials and for allowing patients to receive better treatment during the course of a trial. In this session we will hear from experts in academia, industry and FDA how to design, analyze, and interpret a Bayesian clinical trial in various settings and how the resulting design can be used in the development of a novel health care product. We will hear the advantages and challenges of using Bayesian methodologies, followed by several case studies. We will also hear about the use of the Bayesian approach in the regulatory setting.

Bayesian Adaptive Dose Finding Studies: Smaller, Stronger, Faster

Scott Berry, Berry Consultants

Bayesian adaptive phase II studies are a powerful alternative to the standard fixed designs. I will discuss modeling the dose-response relationship, adaptively allocating the doses, and adaptive stopping rules in a dose-finding study. Simulations will be presented showing that huge improvements can be made over the fixed designs in terms of sample size, time of the trial, and the power of trial increases. The ability to put subjects where they are needed brings about improvements for the study design, as well as the subjects in and out of the trial.

Increased Bayesian Submissions: A Future State of Drug Applications?

Stacy Lindborg, Eli Lilly & Company
John Seaman, Baylor University
Melissa Spann, Eli Lilly & Company

Historically, classical statistical inference has been used almost exclusively in clinical research – for both internal decision making within companies and for drug approvals by external agencies. The continued increase of the cost of drug development and the emergence of the FDA's Critical Path initiative has created an environment in which there is increased opportunity to challenge "business as usual". While most companies are moving forward with novel Bayesian prospective trials, one complimentary step to build confidence in methods with decision makers is to explore historical examples. In this presentation we will review a classic Phase III registration study conducted using a frequentist analytical approach. Following the completion of the trial, the trial was re-designed using an adaptive Bayesian approach. At the end of a rigorous study design process, the study was reanalyzed according to the Bayesian design. We will use this study to explore practical advantages/challenges of Bayesian inference, the prior elicitation process, sensitivity analyses conducted to evaluate the impact of the prior, computational issues and lessons learned.

Bayesian Design and Analysis (Clinical Regulatory Perspective)

Celia Witten, FDA-CBER

Bayesian study design is an option for clinical studies in support of product approval. As with any study design, the clinical interpretation for the modelling and study design choices in Bayesian studies is important. In this talk, several examples of the use of Bayesian statistics, and the importance of clinical interpretation of study design, will be reviewed.

PS07 – Standards and Processes for Effective Communication with the FDA

Organizer(s): Cathy Barrows, GalxoSmithKline; Pratap Malik, Waban Software; Shailaja Suryawanshi, Merck; Stephen Wilson, FDA

Effective communication with the FDA is an essential part of the drug development and approval process. This session will present suggestions and examples of how to improve communication through the use of standards. The speakers will focus on improving communication with the FDA during drug development and on streamlining the processes involved in submission and review through the use of standards like pre-NDA meetings, CDISC and controlled computing environments.

Improving FDA/Industry Interactions: Suggestions from FDA/CDER Statisticians

Rafia Bhore and Janice Derr

FDA scientists are in frequent communication with the regulated industry over development issues during reviews of drug submissions. Although the regulatory communication process is well established, statistical communications between FDA/CDER reviewers and industry counterparts during the drug submission process can be challenging. The emerging emphasis on streamlining clinical trials, as described in the FDA Critical Path Opportunities List, calls for effective communication at all stages of drug development. This talk will provide an overview of how FDA/CDER statistical reviewers communicate internally with their review team colleagues and externally with industry statisticians and clinical researchers. We will give suggestions for improving statistical communication based on drug applications in different therapeutic areas in CDER.

The KISS of CDISC A New Approach to Study Reporting

Pam Ryley, Vertex

Many companies currently use a mix of legacy and current tools to process and report study data. This may result in processes that resemble projects that are reminiscent of the old Mousetrap® by Milton Bradley game. This presentation shares a plan to step beyond the 'Rube Goldberg' approach to simplify and streamline clinical study reporting by optimizing the use of the CDISC model.

SDTM/ADaM Pilot Project Update

Cathy Barrows, GalxoSmithKline

The SDTM/ADaM Pilot Project was initiated by CDISC to illustrate how the various CDISC components can be used to result in a submission of electronic data that are in a format which is acceptable to the FDA and meets the needs of both medical and statistical reviewers. Legacy clinical trial data were converted into a submission package that included tabulation datasets following the SDTM model, analysis datasets following the ADaM model to support the analyses summarized in the submission report, and all relevant metadata as dictated by these models. A group of FDA medical and statistical reviewers then evaluated the submission package to assess whether or not the CDISC-compliant datasets and accompanying metadata were in a format compatible with the tools and processes used for a review and provided the level of detail needed for a review. This presentation will provide an overview of the project as well as a summary of the FDA reviewers' evaluation of the submission package. The objective will be to illustrate the use of standards like CDISC in the communication of drug development work for submission and review.

Communication and Partnership with the FDA- A Recent Experience
Ram Suresh, Schering-Plough

Effective communication and partnership with the FDA is a very important component of any drug development process. End of Phase II meetings, Pre-NDA meetings and formal requests for review of Data Analysis Plans are important modes of communication that facilitate obtaining scientific input from the FDA. A recent experience in the development of an antifungal product for prophylactic use, that fully utilized all these avenues for communication, will be discussed. In particular, evaluation of clinical studies based on the endpoints and statistical methods specified in the protocols that needed to be adapted based on evolving scientific and regulatory considerations, will be highlighted.

PS08 - Diagnostic Medical Imaging

Organizer(s): Josy Breuer, Schering AG; Gene Pennello, FDA

Diagnostic imaging is an increasingly vital tool for identifying and treating disease and, more recently, for monitoring effectiveness of therapy. Examples include digital and three-dimensional tomosynthesis mammography, computer-automated detection and diagnosis (CAD or CADx), and flourescence-based enhanced imaging techniques (e.g., fluorodeoxyyglucose-positron emission tomography or FDG-PET). In general, imaging modalities can present challenges in both study design and analysis. In this session, noted researchers in imaging will discuss approaches to evaluating the diagnostic performance of specific imaging modalities.

Assessing Computer Algorithms: Study Designs and Performance Metrics
Nicholas Petrick, FDA
Brandon Gallas, FDA
Sophie Paquerault, FDA
Frank Samuelson, FDA

Computer algorithms have growing use in clinical medicine. The clinical utility of new imaging modalities and many computer software tools rely primarily on the ability of the clinician to interact with these medical devices. This interaction complicates the evaluation of these devices because reader skill and variability are integral to the evaluation process. In this presentation, we will provide an introduction to some different types of computer software devices available to the clinician. We will discuss some basic pre-clinical and clinical studies utilized in evaluating these devices. We will also introduce the different performance metrics utilized for evaluating these medical devices, including receiver operating characteristic (ROC) and free-response ROC performance metrics, sensitivity and specificity. Finally, we will address how reader variability can be incorporated into the overall statistical assessment methodology for some of these common performance metrics.

Design Issues for Evaluating the Utility of FDG-PET as Patient Management and Drug Development Tools
Lori Dodd, NCI/NIH

As part of a demonstration project under the Oncologic Biomarkers Qualification Initiative, FDG-PET was identified as an imaging technology with the potential to improve patient management and to assist with oncologic drug development. In this talk, I will discuss design issues that should be considered when evaluating PET for these potential goals. An example of a trial in lymphoma will be considered.

Overview of Current and New Contrast Agents for Medical Imaging
Kohkan Shamsi, Berlex, Inc.
Suming Chang, Berlex, Inc.

Various types of contrast agents are used that complement and overcome the limitations of imaging devices like MRI and CT. Complementary development of machines and contrast agents leads to optimization of images and help in diagnosis and management of the patients. An overview of current and new contrast agents will be presented. Iodinated contrast agents are used for CT and are available in various concentrations. MRI contrast agents are either gadolinium based or iron based agents and can be grouped into various categories. MRI contrast agents like Magnevist, Omniscan etc are commonly used. Agents with higher concentration (Gadovist) has also been developed. Liver specific agents include compounds that target RES cells (Resovist/Feridex) and hepatocyte specific agents (Primovist and Teslascan). Blood pool contrast agents like MS325, Gadomer and P792 have long intravascular half life and provide prolonged imaging window that allow acquisition of high resolution images. Targeted agents that are potentially useful for lymph node imaging, plaque imaging and nerve imaging will be presented. Issues that are specific to contrast agents development will also be discussed.

Simultaneous Comparisons of Accuracy, Sensitivity, and Specificity in Diagnostic Studies
Suming Chang, Berlex, Inc.
Joerg Kaufman, Schering AG

A diagnostic test is a procedure to increase the probability of a correct diagnosis in diseased and non-diseased patients or patients with diseased or non-diseased observational units, e.g., vessel segments. The choice of one or more performance measurements for a diagnostic test remains a challenging topic. The most widely used performance measurement is the accuracy, an index of validity for the total study population. It is fully recognized that the sensitivity and specificity play an important role in demonstrating the efficacy of a new diagnostic agent in confirmatory trials. But, the claims based solely on subgroup analyses may not be accepted in the absence of a significant effect for overall study population. Often, the sample sizes of diseased and non-diseased patients or vessel segments are markedly unbalanced. Depending on the weighting schemes used for the accuracy index, different results may be observed in a treatment comparison. We propose to use a weighted accuracy along with simultaneous comparisons of sensitivity and specificity for a primary endpoint. The problem of adjusting for multiple comparisons will also be discussed.

PS09 - Guidance and Standards for Diagnostic Devices

Organizer(s): Vicki Petrides, Abbott Laboratories; Lakshmi Vishnuvajjala, FDA-CDRH

There are many aspects of diagnostic assays that need to be evaluated by manufacturers, the FDA, and laboratories. Fortunately, there are also guidance documents and standards that have been published through the joint efforts of these interested parties to assist in evaluating many of these characteristics. This session will address and highlight some of the guidances and standards that are available.

Revising FDA's "Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests"
Kristin Meier, FDA-CDRH

FDA released “Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests; Draft Guidance for Industry and FDA Reviewers” for public comment on March 12, 2003. The guidance is intended to describe some statistically appropriate practices for reporting results from different studies evaluating diagnostic tests and identify some common inappropriate practices. The recommendations in this guidance pertain to diagnostic tests where the final result is qualitative (even if the underlying measurement is quantitative) and focuses special attention on the practice called discrepant resolution and its associated problems. Following publication of the draft guidance, FDA received 11 comments. Overall, the comments were favorable and requested that additional information be included in the final guidance. Some respondents requested greater attention to the use of standard terminology. This presentation will highlight some of FDA’s current thinking on revising this guidance document including the definition for reference standard and how to report results when there is no reference standard.

Total Error Evaluation in Qualitative Diagnostic Assays
May Mo, Abbott Laboratories

Total analytical error is a useful metric to evaluate assay’s performance and define its product requirements. The CLSI EP21-A guideline released in 2003 provides useful protocols and procedures to analyze and report analytical error f quantitative assays. This presentation will profile the methods that one diagnostic manufacturer has used to gradually introduce the total analytical error concept to all phases of its assay development process since the first release of the draft guideline. For most newly developed quantitative assays, the product requirements for random measurement error (precision) and systematic biases (method comparison, anticoagulant, interferences, etc) are all traceable to the allowable total analytical error from the clinical perspective. However, when extending this new practice to the qualitative assays, the development teams are facing many challenges. The main questions are: How to determine the allowable total error for qualitative assays? How can total analytical error be evaluated? By sharing our exploratory thoughts and efforts in this area, we are hoping to generate discussion and learn from our industry and regulatory colleagues.

Consensus Standards for Method Evaluation: ISO TC69 and CLSI EP
Daniel Tholen, Daniel Tholen Statistical Consulting

International consensus standards are developed with the expertise of all stakeholders in a technical issue, assembled in an open environment where all concerns are addressed. No single interest can dominate, and experts from a variety of different interests may contribute. International acceptance is essential for global trade and understanding. ISO TC 69, SC 6 addresses problems of Accuracy of Measurement Methods and Results, with working groups in Accuracy, Limits of Detection, and Measurement Uncertainty. Other subcommittees deal with acceptance sampling, quality control charts, process capability, and terminology. The CLSI Area Committee on Evaluation Protocols has an extensive series of guidelines for laboratories and manufacturers for determining the capabilities of clinical laboratory measurement procedures. Where they cover the same issue, the CLSI evaluation protocols are consistent with the ISO protocols, but are focused on the needs of the medical laboratory industry and are written for the clinical laboratory worker.

PS10 - FDA's Quality by Design Initiative

Organizer(s): J. David Christopher, Schering-Plough; Yi Tsong, FDA

FDA's Quality by Design (QbD) initiative emphasizes building quality into a product during development rather than attempting to "test" quality into manufacturing of the product. In this

session perspectives from both industry and FDA will be presented, discussing pilot programs, risk assessment, design space and the role of FDA. The QbD approach to pharmaceutical development relies heavily on statistical methodologies, but the roadmap to its implementation is still in progress. Clarification of objectives and chemistry review considerations must be understood to ensure the use of appropriate statistical approaches. The goal of this session is to heighten awareness of the potential benefits to both industry and FDA and to identify important issues for consideration, including Q&A from the audience.

Use of Quality Risk Management and Statistics in Quality by Design - Industry Perspective
Vincent McCurdy, Pfizer

Among the new goals for pharmaceutical manufacturing in the 21st century will include measurement of process capability. In order to achieve a high level of process capability (approach 6 sigma), Drug Substance and drug product manufacturing processes will need to be designed (redesigned) with much greater process understanding. A successful Quality by Design program will encompass both R&D and Manufacturing alignment around a single philosophy and work process. It will begin with an appropriate risk management strategy to identify the key process parameters and quality attributes. While little experimental data are required to initiate a risk assessment, the risk assessment process will identify the most important experiments needed to gain the process understanding. These experiments will be prioritized in a logical sequence that may take years to complete. Likewise, a prioritization of the high value PAT opportunities will be another deliverable from the risk assessment exercise. Extensive use of DOEs, 6 Sigma and science of scale tools is warranted to obtain process understanding.

Role of Statistics in Pharmaceutical Development using Quality-by-Design Approach - an FDA Perspective
Chi-Wan Chen, FDA-CDER

In FDA's view, quality by Design is a system approach to pharmaceutical development where (1) the product is designed to meet patient requirements; (2) the process is designed to consistently meet product critical quality attributes; (3) the impact of formulation components and process parameters on product quality is understood; (4) critical sources of process variability are identified and controlled; and (5) the process is continually monitored and updated to assure consistent quality over time. Statistics can play an important role in this science- and risk-based QbD approach. Various statistical tools, e.g., design of experiments, model building and evaluation, and statistical process control, can be utilized, as appropriate, during product and process design/development and throughout product lifecycle. FDA is inviting participants in the CMC Pilot Program to share, in a summary manner in their applications, their experience with applying these statistical tools to product and process design/development.

PS11 - Smart Choices: Decision Analysis Approaches to Clinical Trials

Organizer(s): Suman Bhattacharya, Genentech; Telba Irony, FDA; Andreas Sashegyi, Lilly

Decision Analysis (DA) is methodology for defining and analyzing a decision problem and making optimal decisions. The basic idea is to delineate the decision problem by quantifying uncertainties and preferences, and then to solve the resulting optimality problem by finding the best possible decision. The use of decision analysis approaches has been gaining increasing importance in

making clinical trial decisions for novel treatments. The use of this approach is becoming more and more relevant in the process of making stage-gate decisions during drug and devices development program. In this session we will hear from experts in the academia, industry and FDA about the possibilities of using DA in various clinical trial settings and in general how it can be used while developing a novel health care product. Case studies will illustrate the use of the approach. Finally, we will also hear the regulatory perspective on the use of DA.

A Quantitative Approach to Clinical Development

Carl Burman, Astra Zeneca

Decision Analysis (DA) can be applied in a range of situations related to the development, regulation and use of drugs. Examples are: dose optimization based on the efficacy-safety tradeoff; regulatory decisions about approval; the prescribing physician's tailoring of the treatment; the ethical requirements on trials; and health economics evaluations. From the sponsor perspective, all these DAs from different stakeholder perspectives must be factored into a commercial optimization, when optimizing trials and trial programs. The DA approach to all problems is essentially the same. Taking study design as the example, a model-based drug development could involve the systematic

1. gathering of relevant information
2. integration of information into quantitative models
3. generation of alternative study designs
4. outlining of objectives and utility specification for different trial outcomes
5. use of clinical trial simulation based on the model to predict the outcomes
6. application of decision analysis to maximize the project value

This talk will mainly focus on the commercial optimization of clinical trial designs and also suggest the use of DA in regulatory applications.

Challenges in Using a Formal Decision Analysis Approach to Medical Device Approval

Larry Kessler, FDA

The Center for Devices and Radiological Health of the FDA regulates an enormous diversity of medical products. These products include items as diverse as latex gloves, infusion pumps, implantable cardioverter defibrillators, magnetic resonance imaging machines, and in vitro diagnostic tests. For products that pose some risk to health, the Agency requires some sort of pret notification or approval before the product can be marketed in the U.S. It is at this time one might imagine a formal decision analysis calculating some sort of risk-benefit profile or index that could apply to the approval process. We will discuss why this is a difficult process from a regulatory/legal and technical perspective. In addition, we will discuss how the collection of post market data can be integrated into such a process that will allow an improved understanding of the risk-benefit balance as the product moves from discovery to delivery to common medical practice.

Linking Statistical Analysis and Decision Analysis in Drug Development and Approval

Dalene Stangl, ISDS, Duke University

Historically, quantitative analysis in clinical trials has served to present statistical summaries of research data. Primarily this means providing point and interval estimates for quantities of interest related to effects of treatment and prevention interventions. The focus is on an estimate of effect size and a hypothesis test of whether we could observe, by chance alone, effects at least as large as those observed. While this research process is itself a sequence of decisions, using

the statistical research results within a formal decision theoretic framework to guide subsequent decision making rarely occurs. However, nowhere is the need for a formal link between statistical analysis and decision analysis clearer than in drug development and approval.

PS12 - Use of Historical Control Data in the Development of Medical Products

Organizer(s): Victor Hasselblad, Duke University; Pat O'Meara, O'Meara Associates, Inc.

This session will explore the use of historical controls in the development of medical products using case studies. The session will include a case study that uses the doubly robust methods described by Lunceford and Davidian. The second speaker will describe his company's experience with historical controls in pivotal trials leading to registration. The third speaker will describe of the use of historical controls in a preclinical setting.

Use of Historical Control Data in the Regulatory Approval of Myozyme
Ron Knickerbocker, Genzyme Corporation
Nancy Sillman, Genzyme Corporation

This presentation will discuss the statistical, logistical, and regulatory issues in the design and approval of Myozyme, an enzyme replacement therapy to treat a rare Lysosomal Storage Disorder. This discussion will include issues in the design and analysis of a company sponsored retrospective historical control study and the use of this study in the interpretation of results from an 18 patient pivotal study and a 21 patient supportive study in an orphan disease. The focus will be on statistical and regulatory issues surrounding the submission and approval of the product.

The Use of Historical Control Data in Evaluation of Oncogenicity of Pharmaceuticals
Karl Lin, FDA-CDER
Atiar Mohammad Rahman, FDA

Historical control data can also be used as a quality control mechanism for a carcinogenicity experiment by assessing the reasonableness of the spontaneous tumor rates in the concurrent control group, and for evaluation of disparate findings in dual concurrent controls.

Formal statistical procedures have been proposed that allow the incorporation of appropriate historical control data in tests for trend in tumor rate. In this study, the empirical Bayes procedures using the beta-binomial distribution to model historical control tumor rates and to derive approximate and exact tests for trend (Tarone; 1982) was applied to some real data sets. The formal statistical procedure works well in situations in which historical data from a large number of studies with relatively large control groups are available to provide reliable estimations of the parameters of the prior distributions. The results from this study show that the incorporation of the historical control data improves the power of the tests. The greatest improvement of power is shown in the tests of rare tumors.

PS13 - Classifiers in Combination Rx/Dx Submissions

Organizer(s): Gene Pennello, FDA; Nusrat Rabbee, Genentech

Recent advances in diagnostic technologies are enabling opportunities to revolutionize treatment decision making by identifying variability among patients in drug response, both in efficacy and in toxicity. Examples of FDA-approved diagnostic tests to indicate the eligibility of patients for

particular therapies include the Dako immunohistochemistry (IHC) test for Her-2/neu overexpression to select metastatic breast cancer patients for treatment with Herceptin (trastuzumab), the Dako IHC test for epidermal growth factor receptor expression to select metastatic colorectal cancer patients for treatment with Erbitux (cetuximab), and the Roche Amplichip to classify a patient's rate of metabolism of therapeutics metabolized primarily by the gene products of two cytochrome P450 genes (CYP2D6 or CYP2C19). Pharmaceutical and medical device companies are partnering to submit in parallel new Dx classifiers to determine patient eligibility for new Rx therapies. FDA guidances relevant to the submission of such combination Rx/Dx products include the CDRH guidance Multiplex Tests for Heritable DNA Markers, the CDER guidance Pharmacogenomic Data Submissions, and the inter-center Drug-Diagnostic Co-Development Concept Paper. In this session, participants will discuss validation of Dx classifiers of combination Rx/Dx products in Phase III trials for simultaneously validating both the Rx therapy and the Dx classifier used to select it. Topics that will be discussed include study design challenges, regulatory schemes for analytical and clinical validation of biomarkers and for combination Rx/Dx products, and business costs and benefits to inclusion of putative predictive biomarkers in clinical trials of pharmaceuticals.

Multiple Comparison Issues in the Use of Biomarkers in Clinical Trials
Stan Letovsky, Millenium

When does it make sense to include a biomarker-related hypothesis as part of the main study objectives of a phase III clinical trial? For a fixed study size, the requirement to adjust for multiple comparisons implies that additional hypotheses reduce the probability of success for the primary hypothesis of drug efficacy. This reduction in probability of success can be justified under certain conditions, from the standpoint of a drug company, based on the possibility of achieving approval in a stratified market where they might otherwise achieve no market at all. However the conditions under which this hedging of bets is justified on cost/benefit grounds are surprisingly rare. Adaptive clinical trial designs may reduce the cost of biomarker hypothesis, though not to zero.

Clinical Trial Designs for Marker Validation
Daniel Sargent, Mayo Clinic

Increasing scientific knowledge is generating promising marker candidates for diagnosis, prognosis assessment, and therapeutic targeting. To apply these results to maximize patient benefit, a disciplined application of well-designed clinical trials for assessing the utility of markers should be employed. We review issues to consider when designing a clinical trial assessing the usefulness of a predictive marker. We then present two classes of clinical trial designs: the Marker by Treatment Interaction Design, and the Marker Based Strategy Design. In the first design, we assume that the marker splits the population into groups for whom the efficacy of a particular treatment differs. This design can be viewed as a classical randomized clinical trial with upfront marker stratification. In the second design, each patient is randomized to have therapy determined by marker status, or to receive therapy independent of marker status. The predictive value of the marker is assessed by comparing the outcomes of the two 'strategy-based' arms. We discuss the advantages and disadvantages of the two trial designs, and their appropriateness to specific clinical situations.

PS14 - Case Studies in Modeling and Simulation

Organizer(s): Alan Hartford, Genentech; Junfang Li, Sanofi-Aventis; Matilde Sanchez, Merck

Modeling and simulation continue to gain importance as efforts are made to streamline drug development. The areas of drug development incorporating these tools are as varied as the statistical methods needed to accomplish them. Two case studies will be presented to give examples of how modeling and simulation are used in different phases of development. The first has a large scope by dealing with optimizing clinical development as a whole and the second is more focused on a specific aspect of drug development by exploring relationships between pharmacokinetics and pharmacodynamics. A third speaker, yet to be identified, will describe the viewpoint of a reviewer of submissions involving modeling and simulation.

Modeling and Simulation: Tool for Optimized Drug Development

Martin Roessner, Sanofi-Aventis

Modeling and Simulation (M&S) is a tool for quantitative decision-making in Drug Development. It provides an integrated framework to use pre-clinical, clinical, PK/PD, epidemiological, financial and marketing information to optimize the drug development process. M&S uses mathematical/statistical models and Monte-Carlo simulation in a predictive manner to address critical development questions. Use of this tool can help reduce late-stage attrition, identify unpromising programs early, reduce time to launch, or support expanded labeling.

Models of drug effects vs. dose considering efficacy of a drug measured usually by surrogate endpoints in early development as well as safety considering e.g. adverse events, laboratory or ECG data, can be used to explore the relationship between the clinical utility and dose. The results can be compared to standard treatments available and the quantitative comparison can be used for further development decisions. Development of a Clinical Utility Index (CUI) requires close collaboration between expert functions. An example is presented to demonstrate the use of a CUI and how it can support go no-go decisions in drug development.

Exposure-Response (PK-PD) Applied to Model-Based Drug Development: A Case Study of Drug X

Matthew Riggs, Metrum Research Group

A brief overview of the role of pharmacokinetic-pharmacodynamic (PK-PD) modeling and simulation (M&S) will be provided, with particular emphasis on continued model development and application throughout the entire drug development process. PK-PD models should capture both the known and unknown (uncertainty) characteristics of the system, where ideally, as information is gained about individual compounds and more generally about entire therapeutic areas, the descriptive and predictive capabilities of the models should expand, as well. A blinded example (Drug X) will be provided to showcase the continuum of M&S through a drug's development. PK and PK-PD modeling for Drug X, as well as comparator information for Drug Y, are being used to support Drug X development decisions. A review of the progression and application of these models, beginning with surrogate markers and expanding into clinical response measures, will be provided in the context of supporting dose & formulation selections for subsequent clinical development phases.

Impact of Pharmacometrics Reviews on Drug Approval and Labeling Decisions

Jogarao Gobburu, FDA

The value of quantitative thinking in drug development and regulatory review is increasingly being appreciated. Modeling and simulation of data pertaining to pharmacokinetic, pharmacodynamic and disease progression is often referred to as the pharmacometrics analyses. Such analyses have been the basis for regulatory decisions pertaining to drug approval, labeling and trial design.

The objective of the current report is to assess the role of pharmacometrics, at FDA, in making drug approval and labeling decisions. Seventy four new Drug Applications (NDAs) submitted between 2000-2006 and needed pharmacometrics reviews were surveyed. Pharmacometric analyses were important for approval related and labeling decisions in 90% of the NDAs. Collaboration among the FDA clinical pharmacology, medical and statistical reviewers and effective communication with the sponsors was critical for the impact to occur. The survey and the case studies emphasize the need for early interaction between the FDA and sponsors to plan the development more efficiently, by appreciating the regulatory expectations better.

Discussant: Stella Machado, FDA

PS15 - Assessing Agreement

Organizer(s): Gregory Campbell, FDA; Shein-Chung Chow, Duke University

In method comparison studies, we are often interested in whether two methods can be used interchangeably or a new method can replace an existing method. Traditionally, method comparison has been evaluated through assessment of agreement between methods. Assessing agreement has been based on indices such as intraclass correlation coefficient (ICC) or concordance correlation coefficient (CCC) for continuous data and kappa index for categorical data. We note that interchangeability between methods is similar to bioequivalence between drugs in bioequivalence studies. In 2001 FDA guidance for industry, FDA adopted an individual equivalence criteria for assessing individual bioequivalence between a test drug and a reference drug. Therefore, it is of great interest to investigate how the traditional agreement indices are related to the individual bioequivalence criteria and whether new coefficients of agreement can be developed that are similar to individual bioequivalence criteria. We propose to assess individual agreement among methods between measurements by different methods. New coefficients of individual agreement (CIA) are proposed to link the individual bioequivalence criteria with existing agreement indices. A unified approach that includes CCC and kappa as special cases will be presented for multiple (k) raters (instrument/methods) each with multiple (m) readings. We hope that these recent advances will stimulate discussion that can lead to future guidance for industry that develops new method, device, assay, or instrument for clinical use.

Assessing Individual Agreement
Huiman Barnhart, Duke University
Michael Haber, Emory University
Andrzej Kosinski, Duke University

Evaluating agreement between measurement methods or between observers is important in method comparison studies and in reliability studies. Often we are interested in whether a new method can replace an existing invasive or expensive method, or whether multiple methods or multiple observers can be used interchangeably. Ideally, interchangeability is established if individual measurements from different methods are similar to replicated measurements from the same method. This is the concept of individual agreement that is similar to the concept of individual bioequivalence. We distinguish two situations where there is a reference method and there is no reference method. We propose coefficient of individual agreement (CIA) for both cases and compare the CIA with the existing agreement indices such as concordance correlation coefficient (CCC) for continuous data and Kappa for categorical data. Several examples are used for illustration.

Assessing Agreement for Qualitative Tests with no Perfect Standard

Bipasa Biswas, FDA

Method comparison studies for qualitative tests are discussed. Population based equivalence test like McNemar's test for addressing equivalence or the Kappa Statistics for inter-rater agreement are addressed with some problems with these measures for assessing agreement for qualitative tests in the absence of a perfect standard. Agreement studies for tests with qualitative outcome (present/absent) in the absence of a gold standard are generally assessed by positive percent agreement and negative percent agreement in stead of sensitivity and specificity.

A Unified Approach for Assessing Agreement for Continuous and Categorical Data
Wenting Wu, Mayo Clinic
A.S. Hedayat, University of Illinois at Chicago
Lawrence Lin, Baxter

This paper proposes a series of Concordance Correlation Coefficient (CCC) indices to measure the agreement among k raters, with each rater has multiple readings (m) from each of the n subjects for continuous and categorical data. In addition, for normal data, this paper also proposes the coverage probability (CP) and total deviation index (TDI). Those indices are used to measure intra rater, inter rater and total agreement, precision and accuracy. Through a two-way mixed model, all CCC, precision, accuracy, TDI, and CP indices are expressed as functions of variance components, and GEE method is used to obtain the estimates and perform inferences. Most of the previous proposed approaches for assessing agreement become one of the special cases of the proposed approach. For continuous data, the proposed estimates degenerate to the overall CCC (OCCC) independently proposed by several authors. When m=1 and k=2, the proposed estimate degenerates to the original CCC. For categorical data, when k=2 and m=1, the proposed estimate and its inference degenerate to the Kappa for binary data and Weighted Kappa for ordinal data.

Discussant: Michael Haber, Emory University

PS16 - Rare Events Estimation Using Insurance Claims Databases

Organizer(s): Gary Aras, Amgen; George Rochester, FDA-CDER

Insurance claims databases offer the ability to study very large numbers of subjects to assess safety, particularly for rare adverse events, although issues concerning the validity of covariates, confounding by indication and outcome ascertainment must be considered carefully. Related topics such as registries, prospective cohort studies to assess drug safety, will also be discussed.

Evaluating Adverse Events After Vaccination in the Medicare Population
Robert Ball, FDA-CBER

Post-licensure observational studies using large linked databases can provide important knowledge about whether adverse events are associated with vaccines, but databases that have been used for these purposes to date still may not have sufficient statistical power to examine rare events, and may underrepresent the elderly. Data from the Medicare program (> 40 million covered persons), can help fill this gap. To assess the utility of Medicare data for evaluating adverse events after influenza (FLU) and pneumococcal polysaccharide vaccine (PPV), we used data derived from the Medicare National Claims History File and Enrollment File to determine

whether hospitalization for cellulitis and abscess of the upper arm and forearm (CAUAF), ICD-9 code 682.3, might be associated with vaccination. For PPV, deviation from a uniform distribution was statistically significant (chi2=231; 53 degrees of freedom; p<.001). Forty-two hospitalizations occurred in the 3 days after vaccination, corresponding to an incidence rate for this diagnosis code of 2.5 per 100,000 persons. Medicare data provide the capability of evaluating even rare adverse events, and include an enormous elderly population.

An Emerging Role for Analyses of Secondary Data to Assess Drug Safety in Post-Market Commitments?

Cathy Critchlow, Amgen

Analyses of data from secondary data sources, e.g., medical claims, are useful in generating hypotheses to be tested in further studies. However, the potential utility of such data analyses to assess drug safety in post-market commitment studies, either alone or in conjunction with clinical registries, remains largely unexplored. Strengths of clinical registries include the ability to estimate absolute event incidence rates, regardless of whether the events were of apriori interest. Strengths of analyses of secondary data sources include large sample size although issues concerning the validity of ascertainment of disease severity, drug exposure, covariates, and outcome events must be considered. Of primary concern in either type of study is the selection of appropriate comparator groups to guide the determination of whether a potential safety signal is present. Study design issues, including the strengths and limitations of each study approach will be explored, with the specific objective of considering how analyses of secondary data sources might supplement, or in some cases be conducted in lieu of, clinical registries in assessing drug safety in post-market commitment studies.

Registries versus Databases for Assessing Rare Events: Points to Consider

Mary Lou Skovron, Bristol Myers Squibb

An overview of the characteristics, strengths and limitations of different types of registries and databases will be presented. Points to consider when selecting a registry or an epidemiological database will be discussed, in the context of the types of events to be estimated (eg short or long latency event, ‘hard’ vs ‘soft’ endpoint). Both practical and methodological considerations will be discussed. Examples of the considerations to be discussed include generalizability, population size, information on patient history, validity of data, duration of follow-up, completeness of event capture, ability to identify salient potential confounders.