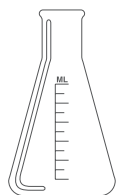


## Biopharmaceutical Section



American Statistical Association

# Biopharmaceutical Report

Volume 14, No. 1

Winter 2006

**Chair:** Leonard Oppenheimer

**Editors:** Demissie Alemayehu, David Giltinan and Richard Caplan



## Letter from the Chair

**Leonard Oppenheimer**

My one year tenure as Chair is rapidly drawing to a close and I'd like to thank all of the dedicated members who volunteered their time and energies in making this section so successful.

The two new initiatives that I wanted to accomplish during my tenure got started but remain in progress.

The first initiative was to effectively exploit our positive financial situation to best support our members and our mission. Should the section be doing more to support students, to help promote the proper use of statistical science to benefit society, to enhance our continuing education and professional meeting opportunities, to assist statisticians in countries or situations with insufficient resources? Please send your recommendations to Stacy Lindborg, the 2006 Chair, for consideration by the Executive Committee.

My second initiative was to increase membership. Even though the section has over 2100 members and is the largest section in the ASA, I'm sure that many of your co-workers are not members of our section, and perhaps not even members of the ASA. Mike Hesney (Chair), Naitee Ting, and Bruce Binkowitz (Membership Committee) will be developing a plan to look into this issue. You can help them get started. The next time you're eating lunch with your statistician co-workers or at a staff meeting, ask how many of them are or aren't section members and why. Encourage them to visit our web site, learn how we contribute to and support our profession, and encourage them to join the section and get involved in our activities.

I'm now looking forward to working with our new section chair, Stacy Lindborg, and her Executive Committee in making the section larger and even more effective. ■

## The Role of Statistics in Medical Devices— The Contrast with Pharmaceuticals

Greg Campbell,  
Director, Division of Biostatistics  
Center for Devices and Radiological Health  
Food and Drug Administration

### Introduction

Medical devices touch all of our lives in many ways as they are revolutionizing medicine with extraordinary advances in detecting and treating disease and in mitigating the ravages of injury and age. With these advances come new challenges

## Contents

### FEATURED ARTICLE

<b>The Role of Statistics in Medical Devices— The Contrast with Pharmaceuticals</b> Greg Campbell . . . . .	1
--	---

### BIOPHARMACEUTICAL SECTION NEWS

<b>Letter from the Chair</b> Leonard Oppenheimer . . . . .	1
<b>2005 FDA/Industry Workshop</b> Ken Koury . . . . .	8
<b>29th Annual Midwest Biopharmaceutical Statistics Workshop</b> Melvin Munsaka . . . . .	10

in evaluation. A growing number of statisticians in the Biopharmaceutical Section of the ASA are becoming involved in research, development and evaluation of these breakthrough products.

Section 2 of the charter for the Biopharmaceutical Section is quite inclusive and allows that the “special interest of the Biopharmaceutical Section is the application of statistics development and use of therapeutic drugs, biologics, and devices in humans and animals”. It continues by saying the primary interests of this Section include: “the biochemical and physical sciences involved with drug, biologics, *device discovery*, formulations, product development and quality control; the biological sciences involved in evaluating drug and *device safety* and efficacy of therapeutic drugs, biologics, and *devices*; experimental human and veterinary medicine in support to therapeutic drugs, biologics, and *devices*... .” (italics added).

What is a medical device? It is defined by law as “an instrument, apparatus, implement, machine, contrivance, implant, *in vitro* reagent, or other similar or related article, including any component, part, accessory, which is 1) recognized in the Official National Formulary, or the United States Pharmacopoeia, or any supplement to them, 2) intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or 3) intended to affect the structure, or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animal and which is not dependent upon being metabolized for the achievement of its primary intended purposes.” (U.S. Code, FD &C Act; 21 US Code 321 (h)). A simpler definition for a medical device is any medical item for use in humans (or animals) that is neither a drug nor a biological product.

Medical devices differ fundamentally from pharmaceutical drugs and biological products in their mechanisms of action. While pharmaceuticals rely on chemical action and the mode of action of biologics is biological, the mechanism of action for a medical device is usually physical. While most pharmaceuticals are for therapeutic uses, *medical devices* can be therapeutic, diagnostic or something else. Examples of devices that are neither for therapy or diagnosis are stair-climbing wheel-chairs, breast implants and facial injections of inert material to minimize wrinkles as well as monitoring devices (whose purpose is not directly to diagnose a disease or a condition) such as uterine monitors, pulse oximeters, apnea monitors, thermometers, and blood pressure machines. There is a staggeringly broad range of products that fall under the definition of medical device. The next time you are in a doctor’s or dentist’s office or in a hospital, look around and observe all the devices that are in use. They range from simple tongue depressors and thermometers to infusion pumps and the latest heart assist and

bypass machines. In the ophthalmic area, devices include contact and intraocular lenses as well as lasers for eye surgery such as LASIK (laser-assisted *in situ* keratomileusis). Audiological devices include hearing aids and cochlear implants. Devices in dentistry include drills, dental amalgam and tooth implants. In neurology, there are deep brain and vagus nerve stimulators. Implants include artificial hips and knees and finger joints as well as spinal fixation devices and breast implants and weight loss devices. There is a wide range of cardiological devices including external

and internal cardioverter defibrillators, stents, endovascular grafts and ablation catheters, to name a few. Diabetics are well aware of the meters and strips to test for blood glucose and of pumps to deliver insulin. There are wound healing devices, latex gloves, condoms, implant material for artificial skin, gastroesophageal reflux disease and for incontinence, as well as injections for knee pain and for facial wrinkles. Monitoring devices include the vast array of machines and alarms used in the intensive care unit (ICU), at

the hospital bedside and at home. Some devices have multiple functions such as an implanted cardioverter defibrillator that monitor the heart for life-threatening arrhythmias and when detected provides a shock to the heart.

There are many *in vitro* diagnostic devices that test patient samples of blood, urine and sputum for conditions such as hepatitis, pregnancy, human papilloma virus, prostate screening antigen, strep and influenza. There are new commercial genetic tests for the oncogene Her-2/neu and for the cytochrome P-450 super-family. There are *in vivo* tests for otitis media and for detecting ovarian cancer in colposcopy. There are a vast number of old and new diagnostic imaging systems including x-ray, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and MR Spectroscopy, not to mention analog and now digital mammography.

There are also a number of products that are device-biologic or device-drug combinations. Some examples include spinal cages which incorporate biological material and drugs used in combination with lasers to enhance cancer therapies. Recently, the newly introduced drug-eluting coronary stents have revolutionized the field of interventional cardiology with a device called a stent that has been coated with a drug. The future of personalized medicine may well be based on the exciting advances in pharmacogenomics, where the use of a diagnostic test developed from genomics indicates whether a person is a good candidate for a particular drug therapy.

## The Nature of Medical Devices—Invention and evolution not discovery

While drugs such as new chemical entities are discovered, by contrast devices are invented. Usually the development of a device is a process that begins with a creative

---

*Imagine the inventor who continually tinkers to make this wonderful product even better. The initial design can change dramatically over time.*

---

idea that is translated into a prototype invention and then it is continually improved over time. Imagine the inventor who continually tinkers to make this wonderful product even better. The initial design can change dramatically over time.

Contrast this with the drug development. Once a drug is discovered it is unchanged in its chemical composition during the entire development process. Of course, the dose and the indications for use, including the dosing regimen, and how drug is prepared and released can change but the chemical entity remains immutable.

A challenge during the development of a device is that it may be changing during the course of its pre-market study. And once a device is on the market the company will usually continue to make changes, hopefully improvements, to the device during its life-cycle. The commercial life-length for marketing for a particular device may only be a couple of years before it is supplanted with a newer model. While both devices and drugs are covered under patent protections over the life of the patent, the device can often change radically and may dictate additional patents whereas the drug is not likely to change. In addition, although the pace of changing device technology is extremely rapid and a new generation of devices such as stents may supplant an older one in less than five years, a specific device may be designed to last a lifetime. The actual length of life of a specific device such as an implant is usually measured in decades or it may be effective for the life of the patient.

Whereas a drug is metabolized, so that discontinuing treatment will address many adverse events, many implants, once in the body, cannot easily be removed. So the risk posed by the device remaining in the body must be weighed against the risk of removal. Implants pose other interesting problems as do other devices that rely on surgical technique to introduce the device onto or into the body. One problem is that the success of the device may depend on the skill of the surgeon to implant it. A similar issue arises in diagnostic imaging devices that rely on skilled radiologists to read and interpret the images.

## The Medical Device Industry

There are many more medical devices than drugs. While the number of drugs is in the thousands, if you count all the types of devices and the various models within each type, there are easily tens of thousands of different medical devices. The companies that make these devices also differ from pharmaceutical firms both in size and number. While pharmaceutical companies tend to be large, the median size of a medical device company is under 50 people total. In contrast to the relatively few pharmaceutical firms, the

number of medical device firms is in the tens of thousands; for example, there are over 25,000 medical device firms registered with the FDA.

## Medical Device Regulation in the U.S. and Worldwide

There is a different law in the U.S. for the regulation of devices than for drugs. Whereas the law for pharmaceuticals is valid scientific evidence based on well-controlled trials, the law for the regulation of medical devices is much more flexible, as one might expect given the different nature of devices compared to drugs. In the U.S., there are three classes of devices, Class I (general controls), Class II (special controls) and Class III, (Premarket Approval) (21 CFR 860.3). The device analog of the New Drug Application (NDA) is the Premarket Approval (PMA) Application that is required for Class III devices. Whereas the determination of safety and efficacy of a new drug has historically relied on valid scientific evidence from well-controlled trials, the FDA directive for PMAs for medical devices is to "rely upon valid scientific evidence to determine whether there is reasonable assurance that the device is safe and effective". And the source of valid scientific evidence includes not only evidence from well-controlled studies but also "partially controlled studies, objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience". (21 CFR 860.7) While for drugs in the U.S. the usual

standard is two Phase III trials, a single confirmatory study is often sufficient for devices.

While the development of drugs is characterized by phases, I through IV, device development is not so easily compartmentalized. Whereas initial drug studies are pre-clinical studies for toxicity, there is usually no analog for devices. There are nonetheless often animal studies for non-diagnostic devices, particularly implants, which would

precede the first-in-man study. In contrast to pre-clinical drug studies to estimate parameters such as ED50, for some medical devices the clinical studies are often preceded by considerable bench and animal testing for reliability and biocompatibility. There are certainly pilot and feasibility studies that serve as first-in-man studies. The analog in the device world of the Phase III drug trial is the confirmatory device study. And of course both drugs and devices have post-approval studies, Phase IV studies in drugs and condition-of-approval and other post-market studies for devices.

The diagnostic area offers the possibility of completely different phases for the development paradigm; see for example books by Pepe (2003) and Zhou et al (2002) for

---

*In contrast to the relatively few pharmaceutical firms, the number of medical device firms is in the tens of thousands; for example, there are over 25,000 medical device firms registered with the FDA.*

---

different characterizations of the diagnostic device development process.

Just as there is a requirement for drug manufacturers to obtain FDA approval before conducting drug trials on humans through the Investigational New Drug (IND), there is a similar provision for significant-risk devices that is called the Investigational Device Exemption (IDE). A significant-risk device is a device that presents a potential serious risk to the health, safety, and welfare of a subject and is 1) an implant or 2) used in supporting or sustaining human life or 3) of substantial importance in diagnosing, curing, mitigating, or treating disease or preventing impairment of human health. Many diagnostic devices may be exempt from IDE provided the testing 1) is non-invasive; 2) does not require invasive sampling presenting significant risk; 3) does not introduce energy into a subject; and 4) is not used as a diagnostic procedure without confirmation of the diagnosis by another, medically established diagnostic device or procedure.

There is regulatory submission in devices for which there is no drug analog, namely, the Premarket Notification or, as it is often called, the 510(k), named after the relevant paragraph in the Food, Drug, and Cosmetic Act. For a device that is not in class III, this provision allows a company to notify the FDA that it intends to market a device that is "substantially equivalent" to a product already on the market and the agency must decide to either clear it (allow its marketing) or to find that it is not substantially equivalent (NSE). Modifications to these devices must also be cleared by the FDA through this 510(k) process. The 510(k) paradigm particularly highlights the difference between device invention and refinement in stark contrast to drug discovery and its subsequent development.

Not all devices undergo randomized clinical trials as the pathway to regulatory approval in the U.S. or overseas. Sometimes historical controls are used rather than a randomized clinical trial, especially if the control is well characterized or studied.

Just as there are Advisory Committees for the FDA for drugs (16 in number) and for biologics (5), for medical devices there are 18 panels for the Medical Device Advisory Committee, one of which is a Dispute Resolution Panel to resolve scientific disputes associated with device applications.

While the ICH, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, is a global effort for pharmaceuticals, there is a similar global organization for medical devices called the Global Harmonization Task Force (GHTF), working toward the harmonization of medical device regulation with website [www.ghrf.org](http://www.ghrf.org) (accessed 11/05). At this time there are no international statistical device efforts that correspond for example to ICH's E-9 (Statistical Principles for Clinical Trials).

### Statistical Issues in Medical Device Studies

The statistical issues in medical device studies include almost all those confronted by clinical trials of pharmaceutical products (except pharmacokinetic or pharmacodynamic considerations or preclinical toxicity studies). For a general

discussion of the similarities and differences in the clinical trials of devices in contrast to those of drugs, see Wittes (2001). Statistical issues that are identical include: sample size issues, handling of missing data, survival analysis, quality-of-life variables, interim analysis, multiplicity of primary and/or secondary endpoints or of hypotheses, sample size re-estimation and meta-analysis. DeMets (2000) considers surrogate endpoint evaluation for medical devices and Scott and Campbell (1997) device subset analyses. With regard to Data Monitoring Committees (DMCs), drugs and devices share many common issues; the FDA Draft Guidance document on DMCs applies to both drugs and devices (FDA, 2001). For a discussion of the independence of the statistician in DMC analyses, see Siegel et al (2004). For devices, non-inferiority studies are of particular interest not only in PMAs but also in Premarket Notifications (510(k)s) where the focus is directly on "substantial equivalence". For non-inferiority device studies, of particular interest the choice of the non-inferiority margin as well the concern (called assay sensitivity in the pharmaceutical world) that the new device is superior to a sham or an inactive control; see Yue (2001) for a discussion of some design issues.

There are a number of statistical issues that are fairly unique to medical device studies.

A very important one concerns design. It is sometime impossible to have a blinded study. Both the patient and the investigator may know exactly what treatment they are receiving. In some cases the treatment is also impossible to hide from the third party evaluator. There are of course sham devices in some cases and there is often a so-called placebo effect associated with the sham (see Kaptchuk et al, 2000). However, in some cases it may be impossible to construct a sham and in other cases it may be unethical. The concern in such instances is the size of the bias associated with lack of blinding.

In other cases the study may have no randomization, either because there is a single concurrent group with an historical control or there are two concurrent arms but no randomization. Not all devices undergo randomized clinical trials as the pathway to regulatory approval in the U.S. Even when a device is Class III, involving a PMA submission, sometimes historical controls are relied upon. Without randomization, the bulwark of statistical inference is undermined and much attention needs to be paid to bias. Causal inferential methods, particularly propensity scores, have proven helpful. The propensity score methodology has been applied in a number of medical device studies to provide insight into how comparable the historical control data is compared to that of the current new device (see Yue, 2003). This is in contrast to the drug world where randomized trials are the norm; however, even in such well-designed trials if there is disproportionate drop-out or non-compliance, causal inferential techniques may prove helpful.

In some cases the performance of the control group is so well-characterized that rather than point to a particular source of data for the historical control, instead an Objective Performance Criteria (OPC) is established by experts using publicly accessible data so that if the new device performs according to the specifications in the OPC then its safety and/or effectiveness can be demonstrated. The statis-

tical issues relate to the sample size necessary for such an assessment and incorporation of variability into the OPC.

Implants pose a number of statistical challenges. For all sorts of surgical trials including those for implants, a major concern is the ability of the surgeon to perform the operation successfully. It is well-known that surgeons often differ in their abilities and that for many procedures the surgeon improves with experience (the learning curve). For a novel treatment such as a new implant, one concern is whether there should be a “burn-in” period that is not part of the study or, if not, to perform an analysis that allows for some improvement as a function of experience for surgeons. For a statistical treatment of learning curves in surgeries, see Cook et al (2004). A statistical issue of particular interest in multicenter trials is the interaction between treatment effect and center or clinic or investigator or surgeon can be large and of great significance. The design challenge is to select wisely the number of sites to enable speedy recruitment yet have power detect any important treatment-by-center interaction. Sometimes a large treatment-by-center interaction in a device trial is indicative of a gap in the training or how the protocol has been implemented at different sites, whereas such an interaction in a drug trial may be more a reflection of a compliance problem.

Another concern with implant and other device trials is that the clinical endpoint is often evaluated over many time points and hence the importance of repeated measures analyses and longitudinal techniques is underlined. Other challenges are posed by the types of devices and the corresponding available data. For a recent example of multivariate mixed modeling involving coronary stents, see O'Malley et al (2003).

It is fairly common in device trials that changes to the protocol or even to the device occur during the conduct of the confirmatory trial. While protocol changes are a problem in pharmaceutical trials, they are more prevalent in device studies. Further, in some cases, an improvement to the device is discovered during the course of the study and the company then wants to modify the device on the fly during the course of the study and then market the new version. In fact, the 1997 FDA Modernization Act allows for this eventuality. As one can imagine this poses very challenging statistical problems for evaluation of the performance of a device.

## Diagnostic Devices

It has been estimated that about 30% to 35% of devices are diagnostic in contrast to the drug world where probably a lot less than 5% would be so considered. While most pharmaceutical drugs are therapeutic, their trials (and

therapeutic device studies as well) often rely on diagnostic and monitoring tests for inclusion/exclusion criteria and for disease progression or recurrence. Most drugs that are used in diagnostic situations are in the context of contrast agents to enhance imaging in conjunction with devices such as MRI. In the world of biologics there are considerably more diagnostic test kits that rely on biological agents such as some HIV and Lyme disease tests and tests for blood type. However the diagnostic device world encompasses a very large proportion of the *in vitro* tests for blood and urine specimens as well as all the imaging devices from MRI, CT and MR spectroscopy to mammographic imaging, not to mention other *in vivo* tests.

While the randomized clinical trial is more than fifty years old and sophisticated statistical machinery has developed to design and analyze such trials, the field of diagnostic evaluation is much younger and the statistical problems in a sense much more daunting. The statistical tools for the design and analysis of diagnostic test evaluation are different from those in therapeutics. In many situations random-

ization may be abandoned in favor of a much more efficient design that compares two diagnostic tests by applying both tests to the same individuals or samples. If the order of testing might be thought to be important then the design could prescribe a randomization for the ordering. But the paired design is often preferred. In diagnostic test evaluation the different sorts of bias become of enormous concern to statisticians; a good discussion of bias in

---

*For a novel treatment such as a new implant, one concern is whether there should be a “burn-in” period that is not part of the study or, if not, to perform an analysis that allows for some improvement as a function of experience for surgeons.*

---

this context is Begg (1987).

In terms of analysis, it is important to distinguish whether the true disease state can be known or not. Some people call this the question of whether or not there is a *gold* or *perfect standard*. In cases where the truth is known and the test is dichotomous, the analysis reduces to two-by-two tables and the inference concerns confidence intervals and/or hypothesis tests for sensitivity, specificity and predictive values. For continuous or ordinal data, the usefulness of Receiver Operating Characteristic (ROC) methodology is particularly relevant for the evaluation of diagnostic products when there is a reliable reference standard; see Zweig and Campbell (1993). In imaging where the reading and interpretation of the images can be quite difficult, as in mammography, an important statistical issue concerns the incorporation of the variability of the readers' skill into the design and analysis. The bootstrap can often play a very important role in assessing the performance of diagnostic tests, as in Beiden et al (2000). The exciting area of the evaluation of computer aided diagnosis (CAD) systems poses interesting statistical challenges in both design and analysis. ROC methodology is often quite useful in investigating the adjunctive claim that such a CAD system

improves the performance of a human reader.

In many situations there is no gold standard. Then the ideal would be to calibrate the reference standard but this is often impossible. Without such calibration, the indiscriminate use of terms such as *sensitivity* and *specificity* is fraught with danger due to the bias that an imperfect standard might introduce. There are also situations where truth may be known only for a subset of subjects. In such cases one would have to worry about how to adjust for verification bias because the establishment of the truth of the condition of the patient depends inherently on the test or tests applied to the patient. An example of this would be the use of discrepancy resolution where discordant results receive further testing. When the true disease state is unknown, latent variable analysis may be used to estimate the performance of the tests under study as well as the unknown prevalence of disease. In agreement studies the focus is on the evaluation of a diagnostic test by assessing its agreement in performance with some other well-understood but imperfect test. For more detail on reporting the results of such diagnostic tests the reader is referred to the FDA's "Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests; Draft Guidance for Industry and FDA Reviewers" (FDA, 2003). In the case of *in vitro* diagnostic tests, there is often a concern with establishing the limit of detection; a nonparametric statistical approach is Linnet and Kondratovich (2004).

In the comparison of two continuous diagnostic tests in the absence of truth, techniques called method comparison are used. This includes techniques, like Deming or orthogonal regression, that allow for errors in both test measurements. However, the inferential question in terms of comparability or equivalence of two such diagnostic tests is not just whether the slope is one and the intercept zero, but whether the total error is well-controlled or not. Note that the use of such methods to establish the (two-sided) equivalence of two methods is very similar to the bioequivalence in the study of generic drugs, using the concept of bioequivalence.

The area of genomics and proteomics in the fast growing field of genetics poses a number of statistical challenges. It also includes genetic tests based on SNPs (single nucleotide polymorphisms), complimentary DNA (cDNA) and oligonucleotide microarrays and proteomics. The design and analysis challenges from a statistical perspective are fairly daunting. Good design depends on intimate knowledge of how microarrays are produced. The analysis is confounded with the challenge of how to adjust for the enormous multiplicity concerns; one approach is the use of the concept of False Discovery Rate (FDR) of Benjamini and Hochberg (1995).

The ability of firms to produce useful tests based on these technologies may rest fundamentally on the ability to produce tests that are reliable. See Campbell (2004) for a discussion of the statistical and regulatory issues associated with microarrays and other genetic tests. The implications of the growth of this area are enormous not only for the development of new diagnostic tests and the new statistical methods to evaluate their performance but also for the pharmaceutical world and how clinical trials of new drugs

may be devised to establish drug efficacy on subgroups identified by a genomic or genetic test.

## Bayesian Statistics for Medical Devices

In 1998 the FDA's Center for Devices and Radiological Health co-sponsored a Bayesian workshop for the device industry concerning the use of Bayesian methods in medical device clinical trials. Device companies have pioneered the design and analysis of Bayesian clinical studies, using hierarchical models to borrow strength from prior quantitative studies as well as building Bayesian predictive models. Irony and Pennello (2001) consider sources of valid prior information for medical device studies. The recent revolution in computing and in algorithms such as Markov Chain Monte Carlo (MCMC) has made such Bayesian computations feasible. More recently, the FDA and Johns Hopkins University sponsored a more general workshop "Can Bayesian Approaches to Studying New Treatments Improve Regulatory Decision-Making?" in May, 2005. The purpose of this workshop was to explore the feasibility of Bayesian methods in a more expansive regulatory manner than just as it relates to medical devices. The proceedings have been published in the journal *Clinical Trials* 2, pp. 271-378 and include a recent recounting of CDRH activity concerning medical devices and lessons learned in Campbell (2005). There is also a website with the video of the program <http://webcasts.prous.com/fda/article.asp?AID=208> (accessed 11/05).

## Post-Market Issues

The post-market arena is particularly challenging for medical devices. There are many more devices and models of devices than there are drugs. The fact that every year or two there may be another model also has implications in terms of the timing to pick up safety issues in the post-market. In addition there is no analog of the pharmaceutical prescription databases that could be used to estimate for example the number of devices implanted in the U.S. in a year. For a glimpse of some of the statistical issues in the device post-market world see Lao (2000). More recently, work has begun to apply the Bayesian Gamma-Poisson mixture modeling techniques of DuMouchel (1999) that have been used for pharmaceutical surveillance to medical device post-market data.

## A Challenge for Medical Device Statisticians: Little Statistical Infrastructure

At the present time the same kind of statistical infrastructure that exists for pharmaceutical statisticians is not available for statisticians in the medical device industry. For example, statisticians who are associated with pharmaceutical or biological products have the support of organizations such as Drug Information Association (DIA) and Pharmaceutical Research and Manufacturers of America (PhRMA), which frequently offer statistical courses, programs, workshops and sessions. In contrast, while device statisticians have Advanced Medical Technology Association (Advamed) and Medical Device Manufacturers Association (MDMA), National Electrical Manufacturers Association (NEMA) and

Association of Medical Diagnostic Manufacturers (AMDM), these associations have only rarely had meetings that are analogs of the DIA or PhRMA workshops on statistical topics or statistically focused sessions at annual meetings. The good news is that Advamed is now considering an Interest Group for Statistics.

One outlet for statisticians interested in medical devices might be through the American Statistical Association. In August, 2005 at the Joint Statistical Meetings, there were nine well-attended Topic Contributed Sessions on medical devices, all of which were sponsored by the Biopharmaceutical Section and six of which the section was the primary (first) sponsor (thanks to Kannan Natarajan for his support as Biopharm program chair). There was also an organizational meeting for statisticians interested in statistics for medical devices that was attended by over 70 people, including most graciously by Len Oppenheimer, the current Chair of the Biopharmaceutical Section. At this meeting it was pointed out that the charter for the Biopharmaceutical Section explicitly includes not only pharmaceutical drugs and biologicals but also medical devices. While a fraction of those in attendance were members of the Biopharmaceutical Section, it is possible that other statisticians with medical device expertise could be attracted to join the Section.

### The Future: New Challenges Bringing Device and Drug Statisticians Closer

The future is bright for medical devices. The crucial role that statisticians have played in the past is sure to continue to expand. Furthermore, the association between pharmaceutical drugs and medical devices will continue to strengthen. There has been a noticeable trend to movement between statisticians from drugs to devices and vice versa. There are many clear areas of overlap. One is in the area of non-inferiority where major similar challenges in design and analysis are being faced by pharmaceutical and medical device statisticians. Another is in the diagnostic area where having a good diagnostic test is absolutely essential in conducting therapeutic drug or device trials. There are now many instances where statistical methodology to evaluate diagnostic devices is just as helpful for contrast agent drugs for imaging and for all kinds of biological tests. An exciting area that I think is a portent of the future is the very rapidly growing area of drug eluting stents. This area is a very good example of a combination product, a device that has been coated with a pharmaceutical product. The effectiveness of the resulting product relies on both the design and placement of the metal stent as well as the particular drug coating on the product. Such products require the best talents of pharmaceutical scientists as well as device scientists and engineers. The design and analysis of studies to demonstrate safety and effectiveness of these products require statisticians who have familiarity with both the principles of good drug studies as well as device evaluation.

Yet another very exciting and newly developing area is that of pharmacogenomics. Again this relies on the combined efforts of scientists and statisticians who are knowledgeable with good design and analysis of drug trials as well as with diagnostic test evaluation. In a sense this is a new area and will require innovative thinking to come

up with research designs to allow for the simultaneous investigation of a new drug and a new diagnostic test. The FDA has recently released "Drug-Device Co-Development Concept Paper" on how to prospectively develop a pharmaceutical or biological product and a diagnostic device in a scientifically robust and efficient manner.

[www.fda.gov/cder/genomics/pharmacoconceptfn.pdf](http://www.fda.gov/cder/genomics/pharmacoconceptfn.pdf) (accessed 11/05). ■

### Acknowledgement

The author is grateful to acknowledge the helpful comments of Kay Barrick on a draft of this article.

### References

1. Begg CG. (1987). Biases in the assessment of diagnostic tests. *Statistics in Medicine* 6:411-423.
2. Beiden SV, Wagner RF & Campbell G. (2000). Components-of-variance models and multiple bootstrap experiments: an alternative methodology for random-effects, receiver operating characteristic analysis. *Academic Radiology* 7, 341-349.
3. Benjamini Y. & Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statistical Soc., Ser. B-Methodological* 57, 289-300.
4. Berry D. "Using a Bayesian approach in medical device development," Technical Report 97-21; Duke University Institute of Statistics and Decision Sciences: Durham, North Carolina, 1997. Web address <http://ftp.stat.duke.edu/WorkingPapers/97-21.ps> (accessed 11/05).
5. DuMouchel W. (1999). Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *The American Statistician* 53, 177-202.
6. Campbell G. (2004). Some issues in the statistical evaluation of genetic and genomic tests. *Journal of Biopharmaceutical Statistics* 14, 539-552.
7. Campbell G. (2005). The experience in the FDA's Center for Devices and Radiological Health with Bayesian strategies. *Clinical Trials* 2, 359-363.
8. Cook JA, Ramsay CR & Fayers P. (2004). Statistical evaluation of learning curve effects in surgical trials. *Clinical Trials* 1, 421-427.
9. Food and Drug Administration. (2001). Draft Guidance for Clinical Trial Sponsors on the Establishment and Operation of Clinical Trial Data Monitoring Committees. Released 11/15/2001. [www.fda.gov/cber/gdlns/clindatmon.htm](http://www.fda.gov/cber/gdlns/clindatmon.htm) (accessed 11/05).
10. Food and Drug Administration (2003). Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests; Draft Guidance for Industry and FDA Reviewers. Released 3/12/2003. [www.fda.gov/cdrh/osb/guidance/1428.html](http://www.fda.gov/cdrh/osb/guidance/1428.html) (accessed 11/05).
11. International Conference on Harmonisation (1998). E9: Statistical Principles for Clinical Trials [www.fda.gov/cber/gdlns/ichclinical.txt](http://www.fda.gov/cber/gdlns/ichclinical.txt) (accessed 11/05).

12. DeMets DL. (2000). The role of surrogate outcome measures in evaluating medical devices. *Surgery* 128, 379-385.
13. Irony TZ & Pennello GA. (2001). Choosing an appropriate prior for Bayesian medical device trials in the regulatory setting, in *American Statistical Association 2001 Proceedings of the Biopharmaceutical Section*. Vol. M, #85. American Statistical Association, Alexandria, Virginia.
14. Kaptchuk TJ, Goldman P, Stone DA & Stason WB. (2000). Do medical devices have enhanced placebo effects? *Journal of Clinical Epidemiology* 53 (8) 786-792.
15. Lao CS. (2000). Statistical issues involved in medical device post-marketing surveillance. *Drug Information Journal* 34, 483-493.
16. Linnet K & Kondratovich M. (2004). A partly non-parametric approach for determination of the limit of detection. *Clinical Chemistry* 50; 732-740.
17. O'Malley AJ, Normand ST & Kuntz RE. (2003). Application of models for multivariate mixed outcomes to medical device trials: coronary artery stenting. *Statistics in Medicine* 22, 313-336.
18. Pepe, M.S. (2003). *The Evaluation of Diagnostic Tests and Biomarkers*. London: Oxford Press.
19. Scott PE & Campbell G. (1997). Interpretation of subgroup analyses in medical device clinical trials. *Drug Information Journal* 32, 213-220.
20. Siegel JP, O'Neill RT, Temple R, Campbell G & Foulkes MA. (2004). Independence of the statistician who analyzes unmasked data. *Statistics in Medicine* 23:1527-1529.
21. U.S. Code, Title 21, Federal Food, Drug and Cosmetic Act, Part 321(h) Web address [www.fda.gov/opacom/laws/fdcact/fdcact1.htm](http://www.fda.gov/opacom/laws/fdcact/fdcact1.htm) (accessed 11/05).
22. U.S. Code of Federal Regulations, Title 21 (Food and Drugs), U.S. Government Printing Office, Washington DC, 2001, Part 860.7 Web address [www.gpoaccess.gov/cfr/index.html](http://www.gpoaccess.gov/cfr/index.html) (accessed 11/05).
23. Wittes J. (2001). Clinical trials of the effectiveness of devices: An analogy with drugs. *Surgery* 129, 517-523.
24. Yue LQ. (2001). Design issues in non-inferiority medical device clinical trials, in *American Statistical Association 2001 Proceedings of the Biopharmaceutical Section*. Vol. M, #451. American Statistical Association: Alexandria, Virginia.
25. Yue LQ. (2003). The Application of Propensity Score Analysis to Non-randomized Medical Device Clinical Studies: A Regulatory Perspective. In *American Statistical Association 2003 Proceedings of the Biopharmaceutical Section*, Toronto, Canada. American Statistical Association: Alexandria, VA.
26. Zhou X.-H, Obuchowski NA & McClish DK. (2002). *Statistical Methods in Diagnostic Medicine*. New York: John Wiley and Sons.
27. Zweig MH & Campbell G. (1993). Receiver operating characteristic plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39, 561-577.

## 2005 FDA/Industry Workshop

Ken Koury, Schering-Plough

The 2005 edition of the FDA/Industry Workshop was among the finest in this long series of enormously successful workshops. The elegant Marriott Wardman Park Hotel in Washington D.C. provided the ideal setting for an outstanding program and the nearly 500 attendees. The weather was beautiful again this year, encouraging participants to walk through the extensively landscaped gardens surrounding the hotel and to explore the historic neighborhoods and attractions nearby.

Four short courses were offered on Wednesday September 14, the day before the workshop. These courses have become an attractive feature of the meeting, providing an excellent opportunity for specialized training in current topics by highly regarded speakers at a very modest cost. The format remained unchanged from 2004, with two parallel sessions in the morning and two in the afternoon.

Once again, these presentations were well attended and well received. In the morning sessions, Javier Cabrera from Rutgers University described the *Analysis of Micro-arrays*, while Estelle Russek-Cohen (CDRH) and Steve Wilson (CDER) discussed *Communicating with FDA Statisticians*. In the afternoon, Jason Hsu from Ohio State University presented *Multiple Comparisons for Making Clinical and Genomic Decisions*, and Joe McKean from Western Michigan University described *Robust Nonparametric Methods for the General Linear Model*.

On Thursday, September 15 five general sessions were featured, and all were rated highly by workshop attendees. The opening session provided an update on the FDA's "Critical Path Initiative", focusing on the role of statistical science in enhancing more efficient, more effective and safer product development. Walter Offen, Lilly, described what this initiative means for pharmaceutical industry statisticians, while Richard Simon, NCI, outlined a new paradigm for using a pharmacogenomic classifier to design and analyze a clinical trial to evaluate the effectiveness of a pharmaceutical product in an overall population, as well as in pre-defined subsets determined by the classifier.

Robert O'Neill, FDA, elaborated on three approaches that could provide tangible statistical contributions to achieving the goals of this initiative: (1) development of guidelines for the statistical handling of missing data, multiple endpoints, and non-inferiority studies; (2) new tools for the design, planning and interpretation of clinical studies, including adaptive/flexible designs, modeling and simulation techniques, and the application of Bayesian methods; (3) addressing other needs such as improving safety and quantitative risk assessment, pharmacovigilance, and product quality.

This presentation provided an excellent framework for the rest of the workshop, as many of the important topics identified by Dr. O'Neill were addressed in more detail in the sessions that followed.

The second general session, for example, illustrated the benefits and challenges of implementing Bayesian approaches in the design and analysis of clinical trials by discussing case studies for both medical devices (Andrew Mugglin, Medtronic) and drug products (Donald Berry, MD Anderson). The goals of the Bayesian approach are to learn faster in early trials (get the dose right!), to develop drugs/devices more efficiently, and to offer better treatment of patients enrolled in clinical trials.

The general session on Data Monitoring Committees incorporated outstanding presentations that focused on ensuring patient safety during the drug development process. Steven Snapinn, Amgen, discussed DMC issues from the pharmaceutical industry perspective, including assessing the need for a DMC, independence of the DMC, the scope of its responsibilities, setting stopping boundaries, issues related to decision-making, and stopping for futility.

Patrick O'Meara, Pat O'Meara Associates, described the role of the independent statistician on the DMC and the importance of clearly establishing the responsibilities and procedures for all parties involved in documents such as the DMC charter and in the contract with the independent statistician. The importance of obtaining the complete set of study documents, access to all relevant databases, and the required support from the sponsor was also emphasized.

Janet Wittes, Statistics Collaborative, highlighted the challenges facing a DMC in making an accurate evaluation of safety based on the typical data displays presented to them. Current approaches include pre-specifying targeted events based primarily on the presumed mechanism of the drug (appealing to statistical conservatism in order to avoid data-dredging), classifying events precisely but narrowly (thereby diluting potential signals that may be important), presenting a long list of data displays (which tends to cloud the real issues), and censoring the follow-up too early (before important trends might become apparent). Suggested alternatives are to be an empiricist rather than rely on mechanism; reclassify and reorganize events in order to increase sensitivity; and identify sentinel events to overcome the dredging issue. Mary Foulkes, FDA, provided a regulatory perspective on these issues and the related 2001 FDA draft guidance.

The final two general sessions dealt with Pharmacogenomics and Multiplicity Issues in the Analysis of Clinical Trial Data. In the Pharmacogenomics session Javier Cabrera, Rutgers University, discussed tests of significance for small samples in the setting of micro-array experiments and differential gene expression, while Michael Ostland, Genentech, described the use of patient selection markers (biomarkers) in drug development programs in order to improve the benefit/risk ratio in targeted patients.

Sue-Jane Wang, FDA, defined genomic drug trials, summarized the current experience using genomic/SNP biomarkers in randomized controlled trials, and described genomic composite biomarkers and their use in designing a targeted sub-trial within an overall trial.

The Multiplicity session provided an excellent overview of this important topic with Alex Dmitrienko, Lilly, presenting gatekeeping and branching procedures for clinical trials

with multiple endpoints and objectives.

Mohammad Huque, FDA, described the sources of multiplicity in clinical trials and the statistical issues and considerations associated with implementing various procedures and strategies for multiple endpoint testing, while Jason Hsu, Ohio State University, defined the generalized familywise error rate and illustrated the use of the partitioning principle to control this error rate.

For 2005, the number of parallel sessions held on the second day of the workshop, Friday September 16, was expanded to 16 (four sessions in each of four time slots). Most attendees appreciated the additional choices and the diversity of the program. Although certain sessions drew a larger audience than others, all of the parallel sessions were considered as among the most useful of the workshop by some attendees. In fact, a recurring theme in the workshop evaluations was that the parallel sessions were useful because they dealt with specific issues or problems, they were directly applicable to the workplace, and because many of the presentations were excellent. The parallel sessions on FDA reporting and surveillance systems for post-marketing adverse events, topics in drug safety, non-inferiority studies, and flexible designs expanded on the workshop themes of safety and the critical path initiative.

A novel and particularly highly rated session was the development and use of patient-centered questionnaires in clinical research. Other sessions highlighted by the attendees included all of the general sessions and the parallel sessions on non-inferiority studies, CDISC initiatives, flexible design, and topics in drug safety.

The roundtable luncheons were continued from 2004, and the response by attendees was very positive despite the inherent noise in the large room setting. The "birds of a feather" sessions at the end of the final day were another feature that brought smaller groups of statisticians with similar, more focused interests together. Topics included statistical issues in medical device trials and re-randomization.

The energetic and committed workshop organizing committee deserves our thanks for developing an excellent program, attracting talented speakers, and producing a first-rate meeting. Special thanks to Mary Bartholomew, our FDA co-chair from the Center for Veterinary Medicine (CVM), for her hard work and the long hours spent keeping the program on track and the organizers organized. Kathleen Wert, from ASA, did an outstanding job of planning the workshop and ensuring that all of the details that were so important to its success were taken care of in an efficient and pleasant manner.

Planning for the 2006 workshop is underway, and fortunately for the Section, Kathleen will continue as a key member of the team. In fact, the workshop will be held again at the Marriott Wardman Park on September 27-29, 2006, and the co-chairs for 2006 are Richard Kotz from FDA-CDRH and Lee Kaiser from Genentech. Stay tuned for more details! ■

# THE 29th ANNUAL MIDWEST BIOPHARMACEUTICAL STATISTICS WORKSHOP

MAY 22 – 24, 2006 • BALL STATE UNIVERSITY, MUNCIE, INDIANA

*Preliminary Program*

## MONDAY, MAY 22

### 8:30 AM – 4:30 PM WORKSHOP REGISTRATION

Fee: \$145 until May 1 (\$40 for students), \$175 after May 1

### 9:00 AM – 1:00 PM SHORT COURSE (Separate Registration Fee: \$55)

Presenters: BRENDA GILLESPIE, University of Michigan  
Topic: Survival Analysis Using Cox Regression

### 2:15 PM – 2:30 PM INTRODUCTION AND WELCOME

Brian Wiens, Myogen  
DR. JOHN W. EMERT, Chair, Department of Mathematical Sciences, Ball State University

### 2:30 PM – 4:30 PM PLENARY SESSION

Speaker: RAYMOND CARROLL, Texas A&M  
Topic: Genetic Epidemiology (tentative)

## TUESDAY, MAY 23

### 8:30 AM – 11:30 AM CONCURRENT SESSIONS

#### A. MULTIPLICITY

*Organizer/Chair:* Alex Dmitrienko, Eli Lilly

1. "Topic to be announced", Ajit Tamhane, Northwestern University
2. "Topic to be announced", Abdul Sankoh, Sanofi-Aventis
3. "Topic to be announced", Mohammed Huque, FDA
4. "Topic to be announced", Devan Mehrotra, Merck

#### B. ASSAY VALIDATION AND OPTIMIZATION

*Organizer/Chair:* James Schwenke, Boehringer Ingelheim Pharmaceuticals

1. "Assessment of Equivalence Using a Concordance Correlation Coefficient in a Repeated Measures Design", Jorge Quiroz, Wyeth
2. "Use and Misuse of the Gage R&R Study in Pharmaceutical Applications", Stan Altan, Areti Manola, and Jyh-Ming Shoung, RW Johnson R&D
3. "Analyzing Immunoassay Data Using Nonlinear Mixed Effects Models", Brian J. Fergan, Xia Xu, and Phillip Dixon, Center for Veterinary Biologics, USDA (BJF), Iowa State University (XX, PD)

4. "Interval Estimates for Assay Response Curves Using Nonlinear Mixed Models", Erin Blankenship, Jacqueline Wroughton, Walter Stroup, and James Schwenke, University of Nebraska, Lincoln (EB, JW, WS) and Boehringer Ingelheim Pharmaceuticals (JS)

#### C. ASSESSING SAFETY IN DRUG DISCOVERY

*Organizer/Chair:* Kimberly Crimin, Pfizer

1. "Use of Micro Arrays to Identify Biomarkers for Predicting Toxicity", Jack Liu, GSK
2. "Hepatotoxicity: Relationships Between Nonclinical Models and Human Hepatotoxicity", Speaker to be announced
3. "Arrhythmia Liability Models or QT", Speaker to be announced

### 11:30 AM – 1:00 PM LUNCH BUFFET

## TUESDAY AFTERNOON, MAY 23

### 12:00 PM – 1:30 PM POSTER SESSION

*Chair:* Caroline Lee, Pfizer, Co-Chair: Kim Perry, Pfizer  
Posters will be accepted on any biopharmaceutical statistical topic. Abstracts must be received by April 28, 2006. Students may qualify for the Charlie Sampson poster award if abstract, poster panels, and a paper briefly describing the poster are received by May 1. For more information contact: Caroline Lee at (734) 622-3085 or Kim Perry at (269) 833-8341.

### 1:30 PM – 4:30 PM CONCURRENT SESSIONS

#### A. SURVIVAL ANALYSIS

*Organizer/Chair:* Alexandra Carides, Merck

1. "Adventures in the Simulation of Clinical Trials With Time-to-Event Endpoints", Brent Blumenstein, TriArc Consulting
2. "Informative Noncompliance: An Overlooked Issue in the Calculation of Sample Size for Survival Trials", Qi Jiang, Amgen
3. "Analysis of Recurrent Episodes Data: the Length-Frequency Tradeoff", Jason Fine, University of Wisconsin, Madison
4. "Discussant", Steve Snapinn, Amgen

**B. HIGH THROUGHPUT SCREENING OF DRUG CANDIDATES***Organizer/Chair:* Ferdous Gheyas, Schering Plough

1. "Virtual Screening with QSAR: Fishing Expedition?", Andy Liaw, Merck
2. "Find that Hidden Treasure – Smooth the Data", Brad Evans, Pfizer
3. "OBSTree: a New Method of Recursive Partitioning", S. Stanley Young, NISS
4. "Testing for Excess Over Highest Single Agent With an Application to High Throughput Screening of Pairs of Compounds", John Peterson, GSK

**C. CURRENT ISSUES IN TRADITIONAL DRUG SAFETY STUDIES***Organizer/Chair:* Wherly Hoffman, Eli Lilly

1. "Respiratory Safety Pharm Studies", Lori Mixson, Merck
2. "CV Safety Pharm Studies", Ron Menton, Wyeth
3. "CNS Studies", Lonnie Grantham, Pfizer
4. "FDA Perspective", John Koerner, FDA
5. Outsourcing, Wherly Hoffman, Eli Lilly

**TUESDAY EVENING BANQUET**

Announcement of Student Winner of Charlie Sampson Poster Award

Speaker: Bruce Rodda, University of Texas SPH

Topic: Natural Selection or Intelligent Design: The Evolution of the Statistician in the Pharmaceutical Industry

**WEDNESDAY, MAY 24****8:30 AM – 11:30 AM CONCURRENT SESSIONS****A. POST-RANDOMIZATION EVENTS***Organizer/Chair:* Robb Muirhead, Pfizer

1. "Treatment Effect Estimators in Longitudinal Clinical Trials With Concomitant Medication", Neal Thomas, Sam Dickson, Naitee Ting, and Donald Rubin, Pfizer (NT, NT), NC State (SD), Harvard (DR)
2. "Evaluating Dose Response From Flexible Dose Trials", Ilya Lipkovich, Eli Lilly
3. "Tackling Selection Bias in HIV Immunotherapy Trials", Robin Mogg, Merck
4. "Discussant", Jeremy Taylor, University of Michigan

**B. STATISTICAL ISSUES IN COMBINATION DRUG DISCOVERY***Organizer/Chair:* John Peterson, GSK

1. "Mixture-Amount Experiments with Applications to Combination Drug Studies", Qing (Kathy) Liu, Wyeth
2. "A Semiparametric Model for Assessing Drug Interaction for Combination Therapy in in vitro and in vivo studies", Maiyong Kong, MD Anderson Cancer Center, University of Texas
3. "Response Surface Methods and Their Application in the Treatment of Cancer With Drug Combinations", Kathryn Dawson, Novartis
4. Discussant, John Peterson, GSK

**C. UNDERSTANDING POST-MARKETING SAFETY DATA***Organizer/Chair:* Alan Menius, GSK

1. "Assessing Safety Signals in Spontaneous Event Databases", Robertino Mera, GSK
2. "Understanding Safety Signals Using Health Care Data", Bob Obenchain, Eli Lilly
3. "Benefit/Risk Modeling", Speaker to be announced

**11:30 AM – 1:00 PM LUNCH BUFFET**

Closing Remarks: Brian Wiens, Myogen

## Let's Hear from You!

If you have any comments or contributions, contact **Editor:** Richard Caplan, Director of Statistical Science, AstraZeneca, 1800 Concord Pike, Wilmington, DE 19850; Phone 302-885-5915; email: *Richard.Caplan@AstraZeneca.com*, **Associate Editor:** Philip J. Pichotta, Statistical Consultant, 879 North St, Milford, CT 06461-2018, Phone: 203-882-9321, email: *pichottapm@optonline.net*; or **Associate Editor:** David Giltinan, Staff Scientist, Biostatistics, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080; Phone 650-225-2296; email: *giltinan@gene.com*.

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

© 2006 *The American Statistical Association*