Statistics in Pharmaceutical Development and Manufacturing

JOHN J. PETERSON

GlaxoSmithKline Pharmaceuticals, Collegeville, PA 19426

RONALD D. SNEE

Tunnell Consulting, King of Prussia, PA 19406

PAUL R. McALLISTER

GlaxoSmithKline Pharmaceuticals, Upper Merion, PA 19406

TIMOTHY L. SCHOFIELD

Merck & Co., West Point, PA 19486

ANTHONY J. CARELLA

Pfizer Inc., Groton, CT 06340

The pharmaceutical industry is undergoing rapid change and facing numerous challenges, including the demands of global competition, the need to speed up the drug-development process, and the Food and Drug Administration's (FDA's) expectations for the incorporation of the principles of quality by design (QbD) and process analytical technology (PAT) in process and analytical development. Statistical thinking and methods play a significant role in addressing these issues. This article provides an overview of the use of statistical thinking and methods in the R&D and manufacturing functions of the pharmaceutical industry. The exposition includes the history of pharmaceutical quality and regulation, phases of pharmaceutical development and manufacturing and the basic quality and statistical tools employed in each, emerging statistical methods, the impact of statistical software and information technology, and the role of statisticians in pharmaceutical development and manufacturing. Four case studies are included to illustrate how these issues play out in actuality. A summary provides a succinct synopsis of those issues and concludes that the complex, technical nature of pharmaceutical development and manufacturing offers many opportunities for the effective use of statistical thinking and methods and that those who use these methods can become catalysts for both process-development understanding and product-quality improvement. Additional details can be found in our technical report (Peterson et al. (2009)).

Key Words: Design Space; Drug Manufacturing; FDA; ICH; Opportunity; PAT; Pharmaceutical Industry; Quality by Design; R&D; Statistical Methods; Statistical Thinking; Technical Change.

Dr. Peterson is Director in the Research Statistics Unit. He is a Senior Member of ASQ and an ASA Fellow. He is the corresponding author. His email address is john.peterson@gsk.com.

Dr. Snee is Principal of Performance Excellence and Lean Sigma Initiative Leader. He is a recipient of the ASQ Shewhart and Grant Medals, and is an ASQ Fellow. His email address is Snee@TunnellConsulting.com.

Dr. McAllister is Director of the Statistical Sciences Unit, North America. He is a Member of ASQ. His email address is paul.r.mcallister@gsk.com.

Mr. Schofield is currently a Senior Consultant with the Biologics Consulting Group, Inc. He is a member of the ASA. His email address is tschofield@bcg-usa.com.

Mr. Carella is a Director in the Nonclinical Statistics Group, Global Research and Development Division, Pfizer, Inc. He is a Senior Member of ASQ. His email address is anthony.j.carella@pfizer.com.

STATISTICAL concepts and tools have been successfully applied for decades in such sectors as chemicals, automobile manufacturing, and computer-chip manufacturing, but their use in the far more regulated pharmaceutical industry presents some unique challenges. For example, many pharmaceutical companies hesitate to invest heavily in large-scale manufacturing-process quality before a drug is approved for marketing because failure in the clinic means product failure. Because time from inception to clinical approval may span 12–15 years, 60–75% of product patent life may have expired by the time the Phase III (confirmatory) trials have been completed. Even after a successful set of Phase IIA (dose determination) and Phase IIB (proof of concept) clinical trials are completed, 40–50% of product patent life may have expired. In addition, one in three drugs is expected to fail in Phase I (first time in man) trials for assessing safety, tolerability, and drug blood levels. Pharmaceutical companies are therefore under considerable economic pressure to file for approval of the new drug application (NDA) with regulatory authorities as soon as possible after completing successful Phase III clinical trials, when a considerable amount of money must be invested in product launch. As a result, quality and process understanding initiatives must compete for time and money with potential losses from considerable 'at-risk' development activities.

Nevertheless, a number of factors are converging to increase the need for sophisticated, statistics-driven approaches to quality and process understanding in the pharmaceutical industry, including:

- *Regulatory Trends.* Recent regulatory guidelines from the Food and Drug Administration (FDA), the European Medicines Agency (EMEA), and the International Conference on Harmonisation (ICH) encourage scientifically based approaches to quality and compliance. Implementing the concepts embodied in those guidelines—process analytical technology (PAT), quality by design (QbD), and design space—will require new, more statistically rigorous and risk-based ways of doing things.
- *Inherent Characteristics of Pharmaceutical Manufacturing.* Many of these inherent characteristics—and the challenges they present—clearly call for the increased use of sophisticated statistics-driven approaches. For example, in most other manufacturing industries, product specifications are clearly tied to product performance. In pharmaceuticals, however, it is difficult to tightly connect, say, 'tablet dissolution' rates to drug efficacy and safety over a vast array of potential product users, each with different body size, age, lifestyle, genetics, and drug-metabolism chemistries. Furthermore, pharmaceutical companies must maintain quality in a many-step production process that creates a complex molecule that must have the proper molecular structure and be free of serious chemical impurities or biological contaminants. In addition, up to now, there has been a lack of incentive for continuous improvement in pharmaceutical manufacturing *after* regulatory approval. This is due primarily to the fact that substantial changes in the manufacturing process or recipe required formal regulatory approval. But new regulatory guidance has been recently introduced to provide more flexibility with regard to continuous improvement in manufacturing. However, pharmaceutical manufacturers will have to show clear process understanding and prediction ability in order to be granted such flexibility. To meet all these complex challenges, pharmaceutical companies need more, not less, statistical thinking and practice.
- *Economic Pressures.* Many companies, faced with thin product pipelines, major patent expirations, and downward pressure on pricing, now need to cut their manufacturing costs, improve yield and productivity, and generate bottom-line savings that can be used to drive growth and innovation. Statistically driven improvement methodologies found in QbD are critical for success in these efforts.
- *Increased Need for Effective Technology Transfer.* Virtually every drug at some stage of its development or manufacture must be transferred from one site to another. Furthermore, mergers, acquisitions, the rise of "global" generics, the ongoing rationalization of manufacturing, and other factors have increased the frequency with which pharmaceutical manufacturing organizations must effectively and efficiently transfer products and manufacturing processes from one location to another. Cost pressures, market needs, government regulations, tax benefits, and logistic issues have also greatly magnified the importance of efficient, compliant, and cost-effective technology transfer—whether it's to a nearby plant or a site a world away. Success-

ful transfer requires a degree of understanding of products and processes that can be greatly improved by statistical techniques.

As these trends continue and converge, the role of statistics and statisticians will only grow larger in the industry. To provide readers with a wider perspective on the status and use of statistical tools and methods in the pharmaceutical industry—now and in the future—the sections that follow treat these key issues:

- The history of pharmaceutical quality and regulation
- The phases of pharmaceutical development and manufacturing and the basic quality and statistical tools employed in each
- Emerging statistical methods
- The impact of statistical software and information technology
- The role of the statistician in pharmaceutical development and manufacturing

The article concludes with four case studies that illustrate how these issues play out in actuality, and an article summary provides a succinct synopsis of the issues and their implications for the future. This article will not address statistical methods or quality in the conduct of clinical trials. The interested reader should see Cleophas et al. (2006) for statistical methods in clinical trials and Griffin and O'Grady (2006) for a review of quality in the conduct of clinical trials. Two books that review statistics in the pharmaceutical industry more broadly are by Millard and Krause (2001) and Buncher and Tsay (2005). A list of key abbreviations and acronyms are given in the Appendix at the end of this article.

History

To understand the forces that have shaped the use of statistics in pharmaceutical development today, it is essential to understand the history of the regulation of the industry. Although bills to regulate food and drugs were introduced in the U.S. Congress as early as 1879, the modern history of pharmaceutical regulation may be divided into two distinct eras: (1) the period of largely reactive legislation that lasted through most of the 20th century and (2) the period of science-based regulatory initiatives that began with the dawn of the 21st century and continues today. Though the two periods overlap to some extent, the new era, with its emphasis on increased process understanding, promises to be markedly different in terms of industrial statistical focus. Table 1 provides a summary of the key regulatory events.

The 20th Century: Regulation and Reaction

In the United States, the precedent for federal regulation of biological products was first established in the Biologics Control Act of 1902, which came in response to the deaths of 13 children caused by a contaminated diphtheria vaccine and the deaths of 9 children caused by a contaminated smallpox vaccine. The act created the Center for Biologics Evaluation and Research (CBER, which became one of the centers of the FDA in 1972).

In 1906, the United States enacted the Food and Drugs Act, known as the "Wiley Act." Although the law focused largely on adulterated food, it also sought to prevent false claims on product labels and to force the acknowledgement of product ingredients such as alcohol, opium, and morphine.

In 1927, the precursor to the current agency dropped all research functions and the current name, FDA, was applied in 1930. In response to 107 deaths from a poisonous solvent used in the manufacture of a sulfa drug, the 1938 Federal Food, Drug, and Cosmetic Act (FDCA) required sellers to prove that their products were safe. The Act also authorized factory inspections and supplemented penalties to include injunctions as well as seizures of product and criminal prosecution. The early 1940s saw the death of some 300 people from a contaminated sulfa drug, leading to stricter manufacturing standards, which would eventually come to be known as the Good Manufacturing Practice (GMP) standards. Also during this period, the FDA enacted batch certification for some products. This required producers to submit a sample of every batch of product to the FDA for testing.

The 1960s witnessed the birth of some 10,000 deformed infants in Europe from a compound called thalidomide. The compound was not marketed in the United States, but resulted in the Kefauver–Harris Drug Amendments of 1962 passed to ensure drug efficacy and greater drug safety. For the first time, drug manufacturers were required to prove to the FDA the effectiveness of their products before marketing them. In addition, testing in animals was required before any human dosing and the FDA's power to inspect manufacturing facilities was expanded.

The Prescription Drug Marketing Act (PDMA)

TABLE 1. Key Regulatory Events

Year	Event	Purpose
1906	Pure Food and Drug Act	Prevent false claims
1930	FDA takes its current name	Agency is purely regulatory—no research functions
1938	Federal Food, Drug, and Cosmetic Act	Require proof of safety before marketing
1949	First publication of FDA "Guidance to Industry"	Address the appraisal of toxic chemicals in foods
1962	Kefauver–Harris Drug Amendments	Require proof of efficacy and safety before marketing
1987	Prescription Drug Marketing Act	Ensure that pharmaceutical products purchased by consumers are safe and effective, and free from counterfeit, adulterated, misbranded, subpotent, or expired drugs
2004	Pharmaceutical cGMPs for the 21st Century—A Risk-Based Approach	Emphasize risk-based approaches to development and manufacturing
2004	PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance	Achieve greater understanding of drug development and manufacturing processes. Data acquisition and multivariate analysis cited as important tools
2005	ICH Harmonized Tripartite Guideline: Pharmaceutical Development, Q8	Foster quality by design and the understanding of design space—emphasis on design of experiments to define interactions and work in multidimensions
2005	ICH Harmonized Tripartite Guideline: Quality Risk Management, Q9	Encourage the use of quality risk-management tools in all phases of a product's lifecycle
2007	ICH Harmonized Tripartite Guideline: Pharmaceutical Quality System, Q10	Enhance science- and risk-based regulatory approaches

of 1987 was designed to halt the sale of counterfeit, adulterated, and expired drugs. The FDA moved to expedite the approval of drugs for life-threatening diseases such as AIDS and eased access to drugs for patients with limited possibilities for treatment.

This history of pharmaceutical regulation in the 20th century could be viewed as a succession of quality regulations, with heavy legislative involvement deriving from the potential for pharmaceuticals to cause harm to human life. In parallel, the quality of products and processes in nonpharmaceutical industries has evolved largely as a result of market forces.

The 21st Century: The Rise of Science- and Risk-Based Approaches

The focus of the FDA and regulatory agencies in many other parts of the world is currently expanding to include a greater emphasis on fundamental understanding of manufacturing processes as the basis for a knowledge-driven, risk-based approach to quality. In the early 2000s, the FDA began a program to focus on what it called manufacturing science. The FDA defined manufacturing science as encompassing knowledge about products and processes, technology used to manufacture and control these processes, and the underlying foundation of a robust quality system at the manufacturing site. This represented a clear shift away from a long-standing position of rigid regulation and inspection to achieve quality standards.

The knowledge base to support these changes begins in research and development and continues though technology transfer and commercial manufacturing. Information related to the active pharmaceutical ingredient (API) and drug product formulation, manufacturing processes and analytical methods, critical-to-quality parameters and attributes, and product specifications are all key elements of the knowledge base. With this knowledge, the firm and the FDA can determine the potential for events to affect fitness for use (i.e., risk). A product's risk potential can be assessed through a mutually developed risk classification system. By sharing risk-mitigation strategies with the FDA, a manufacturer may have a product reclassified to a lower risk class. This is one of the key benefits of a science- and risk-based approach to GMPs.

These new approaches to regulation, compliance, and quality were embodied in a series of guidelines issued to the industry by the FDA and The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). The ICH is a unique project that brings together the regulatory authorities of Europe, Japan, and the United States. The last five rows of Table 1 summarize some of the key documents. They are readily accessible on the FDA and ICH websites.

While regulatory submissions relating to drug product manufacturing have traditionally focused on prescriptive validation and product testing, and less on process design, the increased process understanding pointed to by these 21st century guidelines could lead to increased regulatory flexibility for drug manufacturing. In an effort to streamline costs and improve product quality, the pharmaceutical industry has become increasingly interested in working with regulators to modernize drug development and quality programs. Science- and risk-based regulatory processes will ensure that FDA resources are focused on the highest risk areas and firms are encouraged to use innovative technology to mitigate risk. However, it is incumbent upon the firms to ensure that low- and medium-risk areas remain in an appropriate state of control because these risk classes will receive less regulatory attention.

Additional historical evidence of the growing importance of statistics to the pharmaceutical industry may be found in the eventual creation of a Biopharmaceutical Section within the American Statistical Association (ASA). Initially, a subsection of the more general Biometrics Section, full section status was driven by the need to have a professional organization to bring together industry, academia, and regulators to help meet the specialized needs of the industry. The early leaders in the creation of the subsection were Joe Dresner, Charlie Dunnett, Mike Free, Joe Ciminera, Ron Gauch, Marti Hearron, and Joe Meyer. Early in the life of the Biopharmaceutical Section, the primary preclinical topics were those related to toxicology—especially carcinogenicity studies. Process development and manufacturing issues have only come to the forefront with recent FDA initiatives. Professional organizational homes for these topics are currently spread across ASA and the American Society for Quality (ASQ).

Below is a short summary of some of the early contributors to the area of "industrial" statistics in the pharmaceutical industry. (This is not meant to be a definitive list and we apologize for any inadvertent omissions of names or contribution areas.)

Shein-Chung Chow	Product stability
Joseph Ciminera	Control charting, process monitoring
Mike Free	Product stability
John R. Murphy	Screening designs, content uniformity
Earl Nordbrock	Product stability
Charles B. Pheatt	Product dissolution
Steven Ruberg	Product stability
David Salsburg	Methods comparison
Charles B. Sampson	Quality control
Jun Shao	Product stability
Wayne A. Taylor	Sampling plans
Lynn Torbeck	Process validation

Drug Development, Manufacturing Processes, and Analytical Methods

Following the discovery of an API, pharmaceutical development takes place along five parallel paths: (1) clinical trials, (2) preclinical assessment, (3) API development, (4) drug product development (final dosage form), and (5) analytical method development (Figure 1). The objective of this work is the submission and approval of an NDA. Once clinical development begins, it usually drives the time lines for the other four development paths.

Except in the smallest companies, no individual statistician is responsible for supporting all five paths. This section treats API development, drug product formulation, and analytical development, including the statistical methods and tools that are applied in pharmaceutical development and manufacture. It also demonstrates how these activities fit into a QbD paradigm. Although API development, drug product formulation, and analytical development are presented sequentially here for the sake of clarity, they usually proceed in parallel and they also support the manufacture of clinical trial supplies. Furthermore, all three development paths are periodically affected by how the drug product is faring in the clinical trials. This section, and indeed the entire article, focus primarily on the activities in the gray arrows in the box in Figure 1.

Table 2 shows how API, drug product development and manufacturing, and analytical-method de-

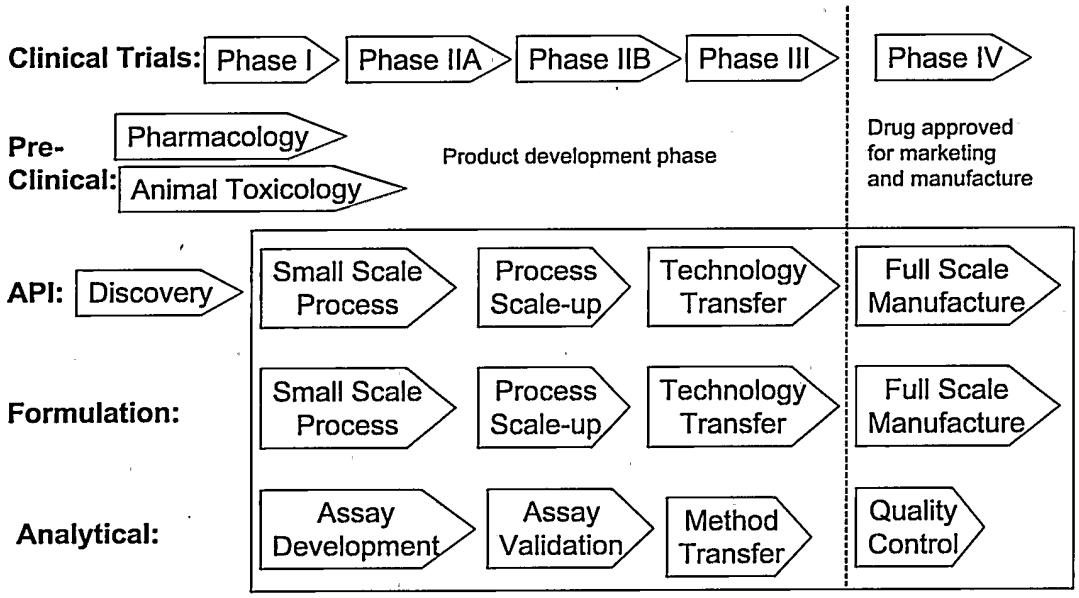


FIGURE 1. Schematic of Pharmaceutical Development and Manufacturing Processes.

TABLE 2. Phases of Development for API, Drug Product, and Analytical Methods

	API	Drug product	Analytical methods
Definition	The active ingredient in the medicine	The medicine as administered to the patient (e.g., a pill or injection)	Procedures for quantifying the amounts of active ingredient or impurities in either the API or drug product
Stages:			
Preclinical	Molecule selection, synthetic route development, creation of supplies for animal testing	Animal studies	Determine measurement needs, select analytical methods, develop standards
Phase I	Start GMP manufacturing process for clinical trial investigation of physical properties	Consider formulation, investigate supplies, stability requirements	Analytical-method optimization
Phase II	Scale-up of chemical synthesis	Drug-product formulation optimized, shelf-life studies begin	Analytical-method validation
Phase III	Clinical-trial supply manufacture, identification of critical process characteristics, process validation	Manufacture of clinical trial supplies, scale-up of drug product process, shelf-life determination, process validation	Product and process specifications finalized, process monitoring (e.g., control charts) established
Phase IV	Technology transfer from R&D to commercial manufacturing facility	Technology transfer from R&D to commercial manufacturing facility	Product and process monitoring for statistical control and quality improvement

velopment for small-molecule therapeutics interlink with the phases of clinical studies. The relationships are detailed and discussed in this section of the article.

API Development: R&D, Tech Transfer, and Manufacturing

The API in a medication or vaccine is the substance or organism that fights the symptom or illness that is being treated or induces immunity to the pathogen. After the discovery of a molecule that has pharmaceutical activity, work is done on this candidate API to further refine the material, to create the manufacturing process, and to scale it up to the desired production levels for full-scale manufacturing.

In the case of small-molecule pharmaceuticals, the first step is the chemical synthesis of the API and the testing of candidate synthetic routes to determine the most effective (chemistry) route from a manufacturing perspective. Next, work is done to verify what was made and how to make it at larger scale—for example, a move from 2–3-gm to 50-gm batches, which are needed to supply preclinical studies in animals. Work is then done to move to bigger batches (1–5 kg) using small-scale commercial equipment or scalable equipment (not beakers). With Phase III clinical studies, API manufacturing eventually grows to its largest scale to date, at least one tenth of commercial batch size.

Much of the statistical thinking and many of the methods and tools used in the development and manufacture of an API are the same as those used for chemical processes. Also, much of process and product development involves experimentation—the key tool, of course, being design of experiments (DoE). Experimentation includes screening experiments, product/process understanding studies, regression modeling, process optimization, and robustness studies. In manufacturing, statistical process control (SPC) is used extensively to monitor and improve processes. On the improvement side, approaches using statistical techniques such as Six Sigma, Lean Manufacturing, PAT, Design for Six Sigma, and QbD are increasingly being used.

It should be noted that the preceding discussion of API applies to drugs developed from synthetic chemicals, the so-called "small molecules". Biologics and vaccines that are produced in biological systems are referred to as "large-molecule" drugs (e.g., proteins), as they are typically much larger molecules. Large- and small-molecule drug production methods have a lot in common, with the principal differences being that large molecules are associated with many raw materials, numerous upstream and downstream process steps, a variety of operating conditions, and numerous types of equipment. Because large-molecule drugs are produced from living organisms, variability is also higher and viral contamination can be an issue. Nevertheless, the associated statistical methods and approaches used are similar to those used for small-molecule drug production. Additional discussion of drug products from small and large molecules is included in this section.

Drug Product Development and Commercial Manufacture (Small Molecules)

Table 3 describes the important formulation and process development activities that must be performed and the milestones that must be achieved in preclinical assessment and in clinical trials to demonstrate acceptable drug product safety and efficacy to regulatory authorities.

The last column of Table 3 lists standard statistical techniques and methodologies that are routinely applied in support of the corresponding stage of drug product development.

Figure 2 illustrates a typical tablet manufacturing process. The API is blended with excipients to form the tablet (i.e., the drug product). The excipients are ingredients that help to keep the tablet intact in storage and then to dissolve at a particular rate after ingestion by the patient.

Drug Product Development and Commercial Manufacture (Large Molecules and Vaccines)

The primary differences in drug product development between small molecules and large molecules (biologics and vaccines) relate to the route of administration of the product and to drug product stability. Most biologics and vaccines are injectables and must be formulated to cause a minimum amount of discomfort to the patient. Some formulation components, such as salts, produce stinging. These are necessary, however, to help stabilize the molecule throughout product shelf-life. Some vaccines, such as live attenuated-virus vaccines, must be stored refrigerated or frozen, and many undergo a formulation process called lyophilization, which is freeze drying under carefully controlled time, temperature, and pressure conditions. Similar to the development of a small-molecule formulation, a large molecule or a vaccine formulation is expedited through the use of

TABLE 3. Drug Product Development and Statistical Support

Milestones	Activities	Statistical support
Nominate an API for clinical development	Discover the API and perform various preclinical studies	Multiple comparison techniques for combinatorial chemists; analysis of genomic data; design and analysis of animal safety studies, etc.
Perform Phase I clinical studies	Determine Phase I dosage type (e.g., liquid, capsule, tablet [or new technology])	Analysis of historical data; statistical thinking (design and analyze experiments)
	Excipient compatibility studies	Design and analyze experiments
	Accelerated stability studies	Regression analysis
Perform Phase IIA (dose ranging) and IIB (proof of concept) clinical studies	Determine Phase II dosage type (new technology)	Analysis of historical data; statistical thinking (design and analyze experiments)
	Evaluate excipient compatibility (if not performed previously)	Design and analyze experiments
	Develop Phase II dosage formulation	Design and analyze factorial and/or mixture experiments
	Develop Phase II manufacturing process	Design and analyze factorial and/or response surface experiments
	Stability studies	Regression analysis
Perform Phase III clinical studies	(If necessary, determine Phase III dosage type)	(Design and analyze experiments to investigate scalability and/or economic concerns with Phase II dosage type)
	Develop and/or scale Phase III dosage formulation	Design and analyze factorial and/or mixture experiments
	Develop and/or scale Phase III manufacturing process	Design and analyze factorial, mechanistic, and/or response surface experiments
	Develop PAT applications	Multivariate analysis
	Transfer technology to commercial manufacturing division	Write reports and consult
Submit new-drug application	Develop and/or scale commercial formulation and process	Design and analyze factorial, mechanistic, mixture, and/or response surface experiments
	Define design and knowledge spaces for DP formulation and process	Design and analyze product- and process-understanding experiments
	Conduct ICH campaign	Analyze ICH stability studies (set expiry)

TABLE 3. *Continued*

Milestones	Activities	Statistical support
Produce commercial product	Establish QA procedures	Assess process capability and establish quality systems to control the process (SPC, PAT, establish sampling plans, etc.)
	Monitor DP stability	Analyze data from annual stability lots
	Improve the process	Data mining, DoE, Six Sigma techniques, Lean techniques, JIT manufacturing, etc.

multifactor design of experiments. Another key area of statistical support is stability study design and manufacturing modeling. For some of the less stable vaccines, release specifications and manufacturing targets must be supported by strategic, and sometimes innovative, stability and manufacturing studies, in order to obtain reliable data to establish a release potency that guarantees safe and effective product through expiry and to establish a robust manufacturing process.

Another feature unique to large-molecule pharmaceuticals and vaccines is related to the complexity of their structure. These products can not usually be comprehensively characterized and, therefore, process or facility changes require carefully designed comparability studies. The industrial statistician may work with engineers to design an efficient and effective study.

Analytical Methods

Analytical methods are used by pharmaceutical manufacturers and regulatory authorities to help guide development and to characterize and control production and distribution of drugs, biologics, and vaccines to the public. Most methods are used to help demonstrate the quality of a product, its potency, and purity. Methods evolve over the course of development in correspondence to need and regulatory requirements. Statisticians are involved throughout the method lifecycle. This includes development of processing strategies for complex methods, such as bioassay and near infrared spectroscopy, employment of multifactor DoE to optimize the method, validation design and analysis to demonstrate the method is *fit-for-use*, and control strategies for maintaining satisfactory method performance. Methods range from measurement of physical characteristics of a

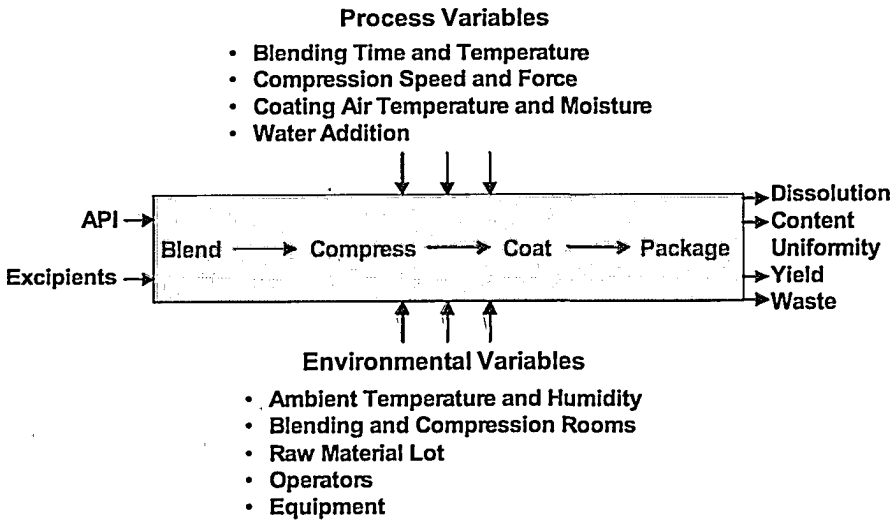


FIGURE 2. Tablet Manufacturing Process and Its Variables.

drug, to chemical composition, to biological activity in vivo.

Methods used to control pharmaceutical products are required to undergo validation in order to demonstrate that they are reliable measures of product quality. Validation standards are outlined in the *United States Pharmacopeia* (USP (2005)) and the ICH guidelines. These guidelines specify the parameters requiring validation, as well as validation methodology. Design considerations and other methods of analysis of validation studies are described in Schofield (2003) and Boulanger et al. (2007).

Pharmaceutical substances and pharmaceutical products are controlled and monitored through specifications and process capability or control limits. Specifications are defined in ICH (1999) as "a list of tests, references to analytical procedures, and appropriate acceptance criteria, which are numerical limits, ranges, or other criteria for the tests described." The limits may be specified in regional compendia such as the *United States Pharmacopeia* and the *British Pharmacopeia*, or developed by the manufacturer. Once defined, these are used to assure the quality of products or their intermediates.

The ideal pharmaceutical quality system is composed of specification limits, release limits, and process-capability limits. We denote by LSL, LRL, and LCL the lower specification limit, the lower release limit, and the lower capability limit, respectively. USL, URL, and UCL denote the repective upper limits. The specification limits (LSL, USL) reflect restrictions within which product is *fit for use*, and must conform throughout shelf-life. The release limits (LRL, URL) assure these limits are met at release and throughout product shelf-life and may be calculated per Apostol et al. (2008), while the control or process capability limits (LCL, UCL) describe both process and assay variability. See Figure 3 for an example.

Methods that have been developed and validated in the research laboratories must be transferred for implementation in the manufacturing quality-control (QC) laboratory. Various approaches are utilized to help assure that the method is performing properly. An equivalence approach correctly addresses the hypotheses of interest. The change in the average response between laboratories and the increase in variability are held to restrictions that must be met using methods similar to those used to establish bioequivalence (Schuirmann (1987)). Another approach using an β-expectation tolerance interval, similar to its use in method validation, has been advocated to demonstrate that measurements made in the QC laboratory conform to an acceptable shift from the distribution in the development lab (Dewé et al. (2007)).

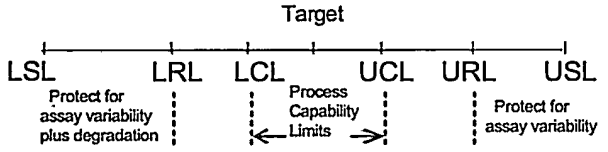


FIGURE 3. Assay Quality-Control Limits.

A key element of the quality assessment of product during development is the determination of product shelf-life. ICH Q1E provides a formulation for the design and analysis of development stability data (ICH (2003)). Studies are usually performed for each product image, which includes different dosages and container/closure profiles. A fixed-effects analysis is performed on stability measurements, which are designed over time to capture changes in important properties of the drug, such as potency and degradants. Rules for pooling the slopes and intercepts of the batches use hypothesis testing at an increased significance level, such as $\alpha = 0.25$, in order to improve the power of the test to detect meaningful differences among lots. After pooling has been assessed, shelf-life is estimated as the intersection of the one-sided 95% confidence interval on the mean and the product specification. Other approaches based on prediction interval (Carstensen and Nelson (1976)) and tolerance-interval concepts (Kiermeier et al. (2004)) have been proposed but do not appear to be much used in practice as yet.

Statisticians have contributed to development-stability studies by offering design strategies, such as bracketing and matrixing of stability time points, which greatly improve efficiency with little impact on the effectiveness of the study (Nordbrock (1992)). Bracketing is a strategy wherein extremes of a product image are tested rather than all images. Thus, if product is to be sold as 20-, 50-, and 100-mg doses, a bracketing study might study only the 20- and 100-mg doses. Matrixing involves testing only a subset of the batches at selected time points (a type of incomplete block design). The combination of bracketing and matrixing leads to a parsimonious evaluation of product stability. Mixed-effects modeling of "random batches" has been proposed by others (Chow and Shao (1991)). Efficient strategies for monitoring

stability of commercial product have been described in Fairweather et al. (2003).

Quality-control testing of manufactured product is carried out to help assure safe and effective product and to monitor the process for shifts or trends. Statisticians work with the QC labs to establish test strategies designed to obtain reliable estimates of product-quality attributes related to dose, potency, and purity. Replication may be utilized to increase the precision of the reportable value, and thereby reduce the risks to the customer of an out-of-specification (OOS) result. Some laboratories employ sequential-test plans in which product is tested using a variable number of assay measurements. The statistician works with the quality laboratory to establish acceptable quality criteria with minimum risk to the manufacturer and customer alike.

Regulatory guidelines prescribe a rigorous investigation of the QC laboratory and the manufacturing process when an OOS result is obtained (FDA (2006)). Statisticians may contribute to the process by providing the laboratory with retest and resampling strategies.

Over time, methods may change or a new method may be developed to measure a particular quality attribute of a product. When this is done, the laboratory performs a concordance analysis. Samples are tested across a range of responses, and a concordance correlation coefficient is determined as a measure of the agreement of the two methods (Lin (1989)). A related measure is the concordance slope (Schofield (2003)). This can be obtained from an eigenanalysis of the two assays, and it measures the slope of the relationship between the two methods.

Emerging Statistical Methods

Satisfying the FDA's encouragement of the pharmaceutical industry to achieve better understanding of their manufacturing processes and to quantify the risks associated with an OOS product requires a wider set of statistical tools than that commonly used in pharmaceutical manufacturing. This need is in large part driven by new technologies supported by the PAT initiative and by the need for better technology transfer from development to manufacturing plant.

Conventional pharmaceutical manufacturing is generally accomplished using batch processing with laboratory testing conducted on collected samples to evaluate quality. A key component of PAT is the desire to implement continuous, real-time quality assurance. As previously noted, the FDA considers PAT to be a system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality (FDA (2004a, b)). PAT is expected to produce gains in quality, safety, and efficiency by (1) reducing product-cycle times by using on-, in-, and/or at-line measurements and controls, (2) enabling real-time release, (3) increasing automation, and (4) facilitating continuous processing.

Many PAT tools tend to be multivariate in nature (in both the independent- and dependent-variable sense). Pharmaceutical products and processes involve complex, multifactorial systems. The understanding of these systems is achieved through the use of multivariate mathematical approaches, such as statistical design of experiments, response surface methodologies, process simulation, and pattern recognition tools, in conjunction with knowledge-management systems.

Such data is often collected by way of process analyzers. Such analyzers include those that take univariate process measurements (e.g., pH, temperature, pressure) and those that nondestructively measure biological, chemical, and physical attributes. These measurements may be taken at line, in line, or on line. Process analyzers typically generate large volumes of data. Multivariate statistical methods are often needed to extract critical process knowledge for real-time control and quality assurance (e.g., principal-components analysis, projection to latent structures, time series, batch modeling). Sensor-based measurements may provide a useful process signature.

One of the chief technologies associated with the PAT initiative is near-infrared (NIR) spectroscopy (Skibsted (2006)). This technology is a spectroscopic method utilizing the near-infrared region of the electromagnetic spectrum (from about 800 nm to 2500 nm). NIR spectroscopy produces a spectral trace that may characterize subtle physical and chemical properties of a substance (e.g., chemical intermediate or tablet excipient) (see Figure 4).

Such a trace is a high-dimensional vector response. As the number of experimental units (n) will be much smaller than the dimension of the vector (p), we have a "large p/small n" statistical inference situation. Such situations are typically dealt with by latent-

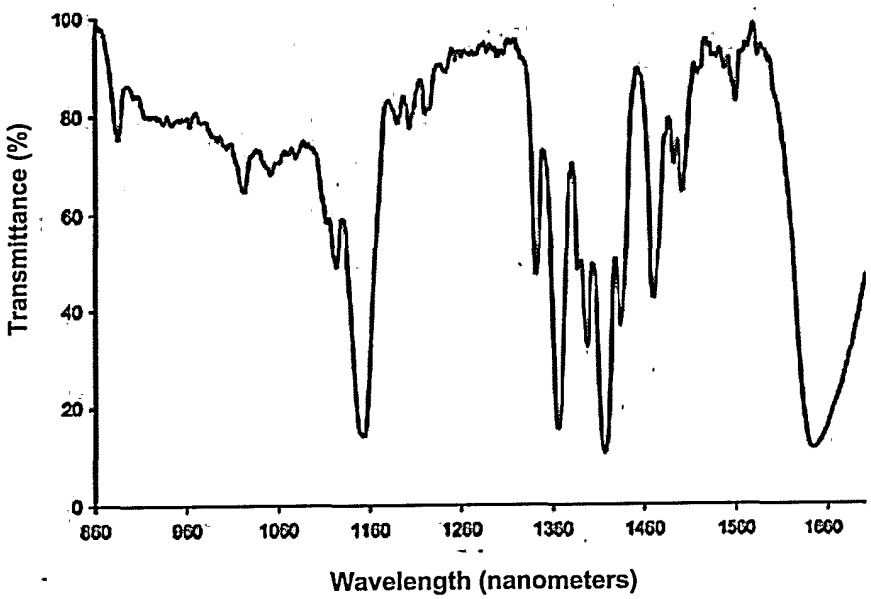


FIGURE 4. An NIR Trace.

variable methods. Such methods are popular in the field of chemometrics, which typically uses partial least squares (PLS) to analyze and reduce the dimensionality of such data.

Spectroscopy, or other technologies that produce traces of output across space or time, produce what is known as functional data, that is, an observation viewed from the perspective of a 'profile' rather than a point in a one- or low-dimensional space. (See Ramsey and Silverman (2002) for examples of functional data and associated statistical-analysis techniques.) Because of the increase in functional data due to new technologies, the field of quality and industrial statistics is adjusting by starting to create new methods for statistical process control based on such data (Jeong et al. (2006)) and for process optimization (Nair et al. (2002)) involving functional data.

The use of data-mining statistical methods may help provide a better fundamental understanding of processes measured by functional response traces. However, practitioners need to understand that data mining is not magic and that the acquisition of new knowledge (or law-like relationships) among factors affecting a process requires the purposeful variation of factors according to well-designed experiments. However, the careful use of data-mining techniques to detect high-dimensional special causes may help here (Pamias (2005)).

Another technology being considered by some pharmaceutical companies is quasi- (or semi-) continuous manufacturing (Leuenberger (2001)). The idea here is to produce many small batches, similar in size to those produced for clinical-trial supplies. This obviates the need for scale-up and also produces rich batch-to-batch variation information, which invites the use of Bayes and empirical Bayes techniques. A recent overview of Bayesian approaches to process monitoring, control, and optimization can be found in Colosimo and del Castillo (2007).

Despite a predominance of batch manufacturing, a case can be made for utilization of continuous manufacturing, at least for some situations (Kossik (2002)). As such technology becomes integrated where needed in the pharmaceutical industry, it will generate a corresponding need for statistical feedback-control procedures (Box and Luceño (1997)) and for more sophisticated statistical process-monitoring methodologies (e.g., Alt (2007)).

It should be kept in mind that data-intensive methods may benefit from better data-gathering techniques (Steinberg et al. (2008)). Data mining as a practice is oriented toward analyzing a given collection of data. This alone may be acceptable for those industries for which it may be difficult to conduct designed experiments (e.g., credit-card companies). However, for an industry such as pharmaceuti-

cals, where most data is obtained from experiments, concepts of experimental design should be used for improving data gathering in the presence of data-intensive technologies.

In the various QbD regulatory documents, the word 'risk' turns up often, usually in reference to the fact that regulatory authorities see new quality initiatives for pharmaceutical manufacturing from a 'risk-based' perspective, with an eye to low risk to the patient. From a quality and industrial statistical viewpoint, this brings up two primary statistical methodologies, more directly those related to risk assessment and quantification (see ICH Q9 (2005)) and more indirectly those related to variation determination.

ICH Q9 stresses identification and assessment of associated risk likelihoods and consequences of risk events. Classical quality risk-analysis identification procedures, such as failure modes and effects analysis (FMEA) and fault tree analysis (FTA), are recommended. Probabilistic risk assessment is only mentioned briefly as associated with "supporting statistical tools." However, it does appear that some regulators recognize the important role of probability models for risk assessment. See, for example, Claycamp (2008). While sophisticated risk-assessment methodologies may be new to the pharmaceutical industry, the changing regulatory atmosphere appears to welcome any sound procedure for improving risk assessment. Probabilistic risk assessment and systems reliability are areas where quality professionals can make important contributions to pharmaceutical manufacturing.

Clearly, the determination of risk depends on variation. For example, a probabilistic reliability computation involving a normally distributed endpoint requires knowledge of variation as well as knowledge of the mean. Therefore, a proper reliability quantification may require a careful variance-components quantification. However, a more insidious issue is that both quality engineers and statisticians may simply follow statistical methods designed primarily for statistical inference on the means as a way to address assurance of quality. One example is the use of "overlapping mean response surfaces" as a way to find the "sweet spot" (Lind et al. (1960), Anderson and Whitcombe (1998)) of a multiple-response process with regard to meeting process-response specifications. Peterson (2004) shows that the probability of meeting specifications when operating well within such a sweet spot can be associated with a disappointingly small value. What is needed is a process-optimization approach tied directly to the probability of meeting specifications.

Design space, a key concept associated with risk in ICH Q8 (2005), actually is not directly related to experimental design. Rather, it is related to a region of process capability. The ICH Q8 guidance defines design space as the "multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality." The issue of risk is embodied in the phrase "assurance of quality." The guidance goes on to state that "working within the design space is not considered as a change." Of course this begs the question: "How much assurance?" Further, the ability to modify the production recipe without regulatory approval (if one stays within the design space) means that manufacturers could pursue continuous improvement with their manufacturing processes.

Currently, however, pharmaceutical manufacturers may have little business incentive for continuous improvement in their manufacturing processes after regulatory approval—for two reasons. First, typical manufacturing processes require costly and lengthy regulatory approval for any change in the specified manufacturing conditions. Second, there is the perceived risk of product batch failure due to changing the manufacturing conditions (i.e., process parameters and input variables).

However, a valid design space, approved by regulators, would provide pharmaceutical manufacturers with the ability to make small-to-moderate changes in their manufacturing conditions (within the design space) without the time-consuming process of regulatory approval. This would also allow manufacturers to safely experiment within the design space and thereby gain important information about their manufacturing process *after* approval by regulators.

The ICH Q8 Annex (2007) appears to suggest the classical approach of "overlapping mean response surfaces" as a way to construct a design space. But, as previously noted, such an approach does not quantify "how much assurance" of acceptable product, and the resulting "sweet spot" may possess factor combinations associated with poor process reliability.

A Bayesian approach to the ICH Q8 definition of design space has been given by Peterson (2008), although such a region for an HPLC assay had been

suggested earlier by Peterson (2004, Fig 6). Two case studies are also illustrated in Stockdale and Cheng (2009). Here, design space is defined as

$$\{x : \Pr(Y \in A \mid x, \text{ data}) \geq R\}, \qquad (1)$$

where $\Pr(Y \in A \mid x, \text{ data})$ is the posterior predictive probability that a new vector of relevant production responses, Y, will fall within the specification (i.e., acceptance) region, A, given the experimental data and a vector of controllable process factors, x. Here, R would be some prechosen level of reliability. (See Peterson (2008) for some comments on the choice of a value for R.) For the design space in (1), the input variables (e.g., raw material attributes) could be treated as noise variables.

On the other hand, a design space could be defined as

$$\{(x, z) : \Pr(Y \in A \mid x, z, \text{ data}) \geq R\}, \qquad (2)$$

where z is a vector of input-material measurements. The design space in (2) would be useful for feed-forward control procedures where movement within the design space would be motivated primarily by input-material measurements (MacGregor and Bruwer (2008)). A further discussion of other applications of the Bayesian approach can be found in the authors' technical report, Peterson et al. (2009).

The Impact of Statistical Software and Information Technology

For drug-discovery research, the pharmaceutical industry has already witnessed the rapid increase in software for number crunching and information storage and retrieval. Part of this has been due to the genomic revolution and the associated 'omic' technological platforms for generating such data. However, with the advent of the PAT initiative, technology will again be driving the need for more sophisticated software and information technology to store data, retrieve it, and manage it. Further, emerging statistical methods, as outlined in the previous section, will also require, to some degree, additional computing support or application development. Meanwhile, increased computing speed, storage, and interactive capability are also affecting quality and industrial statistics in the pharmaceutical industry. On one hand, need will, in some cases, drive software creation. On the other hand, software availability will guide what statisticians and quality engineers are willing to do with statistics and graphics.

The increased availability of easy-to-use statistical commercial software allows scientists, engineers, and various technicians to do their own analyses. Such software is useful not only for statisticians but for teaching and the enablement of (nonstatistical) engineers and scientists. It can also help to empower a trained statistical "champion."

The pharmaceutical industry is inundated with sophisticated assay equipment and measuring devices (e.g., for image or spectroscopic analyses). Some of this equipment has built-in proprietary algorithms that are statistical in nature. It can be difficult to ascertain how well these algorithms perform or what "fine print" assumptions are required for reliable statistical analyses. Some statistical packages have their algorithms well documented, while others do not.

Along with data mining, opportunities in the field of high-dimensional information visualization arise with PAT as well. Shop-floor technicians and chemical/biochemical engineers will have a strong desire to view information in their data whether or not it may be high dimensional. The visual package Spotfire® (Spotfire Inc.) can help here, but much more is possible in theory (see, e.g., Fayyad et al. (2002)).

Further opportunities for information technology to contribute lie in the field of database and knowledge management. A pharmaceutical-industry survey (Morris (2005)) has indicated that only 10% of database information has been leveraged to improve overall competitiveness and compliance. So good opportunities exist for database and knowledge-management professionals to impact the pharmaceutical industry from both competitiveness and regulatory perspectives. As mentioned above, the PAT initiative will only increase these needs due to data complexity and volume.

In addition to communication among repositories of collected data, there is a need for related databases of information that allow for easy storage and retrieval of data analyses (graphs, point estimates, confidence intervals, etc.) along with the language code and software that generated such entities from the data. Again, all of this is more pressing for the highly regulated pharmaceutical industry.

For all of the foregoing reasons, the pharmaceutical industry has become interested in "enterprise" statistical and information systems. Such systems allow users to store and track information across various groups within an organization. Examples of such systems are the *Statistica Enterprise wide SPC system* (StatSoft, Inc.) , the *SAS® Enterprise Guide® 4.1 for Statistical Analysis* (SAS Institute Inc.), and

the *Minitab Quality Companion* (Minitab Inc.). Such systems provide electronic venues for developing add-on modules for analyses particular to the pharmaceutical industry. But such systems also call for greater due diligence in software validation across a large group of company subunits that are in a regulated environment.

Despite the need for reliable software validation, increasing sophistication in technology is driving the need for increasing sophistication in statistical analyses and algorithms (Steinberg et al. (2008)). Many younger statisticians (particularly recent graduates) in both the pharmaceutical industry and associated regulatory agencies (e.g., the FDA) have learned their applied statistics using the R (open source) computing environment (Bell et al. (2006)). Consequently, these statisticians want to use R because it is familiar and often has relatively recent statistical methods that are not yet available in commercial packages. In fact, the Drug Information Association and the FDA cosponsor an "Open Toolbox Initiative Forum" to support vendor-neutral software products in an integrated environment.

Interestingly, there are no federal regulations prohibiting the use of open-source statistical software (Bell et al. (2006)). Instead, the FDA rightly supports good software validation that produces a "level of confidence" in the results. Further, the choice of statistical software should not alter the results (Bell et al. (2006)). Given the growing complexity of statistical software, this appears to imply that identical (or possibly similar) analyses using two or more different software applications may be needed to help provide such a level of confidence. Of course, such "level 1" validation exercises may not be considered sufficient for many statistical analyses subject to regulatory oversight. For a discussion of unmet software needs for industrial statistics, see the authors' technical report, Peterson et al. (2009).

The Changing Role of Statisticians

As the foregoing discussion makes clear, the utilization of statistical thinking and methods has made significant contributions to the success of the pharmaceutical industry. Undeniably, statistical methods now play an integral part in the industry—in discovery by chemists and biologists, in drug development and manufacturing by chemists, engineers, and pharmaceutical scientists, and in regulatory compliance and quality control.

But what of statisticians themselves? Certainly, they have played a major role in providing the infrastructure as well as in solving important problems and expanding the use of statistics in the pharmaceutical industry. However, as statistical thinking and methods become even more critical for success in the industry, statisticians will need to acquire new skills, particularly leadership skills, to move from their traditional role as passive advisor to a more dynamic role as creator of value. Statisticians in the pharmaceutical industry can provide leadership in three basic, and synergistic, areas: (i) within their own company, (ii) within the pharmaceutical industry (e.g., the Pharmaceutical Research and Manufacturers of America (PhRMA)), and (iii) in collaboration with statisticians at the FDA (e.g., the annual FDA/Industry Statistics Workshop).

From Consultant to Collaborator to Leader

As statistics originally evolved with scientific management, statisticians became highly specialized functionaries in large organizations. They analyzed data that other people had created and passed along their analysis to engineers and business leaders who used the analysis in making decisions. In the pharmaceutical industry, as in other industries, statisticians first served in that consultative role, typically working one-on-one with internal "clients," designing studies, analyzing data, and providing training. They provided methods for data collection, strategies for design of experiments, and provided guidance on the most effective use of the concepts, methods, and tools of statistics. When statistical software emerged, they added that expertise to their portfolios.

A confluence of several powerful trends in recent years is requiring new roles for the statistician. First, the internet now makes many data sets accessible to anyone instantly, undermining the statistician's "ownership" of data. Second, the ubiquitous commercial statistical software described above enables nonstatisticians to perform many of the statistical operations that were formerly the province of specialists. Third, statistics classes are now commonplace in academia, including in business, engineering, economics, and social-science curricula; and the widespread use of statistics-intensive methodologies, like Six Sigma, has provided mass statistical training for nonstatisticians in many organizations.

But more positively, and perhaps most importantly, pharmaceutical industry scientists and engineers began to recognize the value that statisticians

TABLE 4. Expanded Role of Statisticians

Consultant (Old)	Leader (New Expanded)
Analyze data and design experiments	Determine the appropriate strategy and approach
Teach statistical tools	Design training systems, coach, mentor, and train
Work with technical people	Work with managers and technical personnel
Consult on other people's projects	Lead cross-functional projects
Narrow expertise and accountability	Broad expertise and accountability
Follow simple regulatory guidelines	Collaborate with regulatory agencies to influence new guidance
Reactive	Proactive

could provide. As statisticians were assigned to more and more project teams, their role changed from consultant to collaborator. They made unique contributions in their capacity as the primary interpreters of data, guiding interpretation and determining the most effective use of tools and methods. Today, in many industries, statisticians serve as team leaders, providing guidance and oversight of programs from beginning to end. While this role is less common in the pharmaceutical industry, it is a model that can work effectively and should be considered.

New Skills Needed

One index of the increasing value that the industry places on statistical methods may be found in the number of industry organizations devoted to addressing statistical issues. Those organizations include:

- PhRMA Chemistry, Manufacturing, and Control (CMC) Statistics Expert Team comprised of ~30 statisticians from ~20 PhRMA member companies;
- Informal Nonclinical Statistics Forum, a colloquium of nonclinical statistical managers who meet yearly to share experiences;
- Midwest Biopharmaceutical Statistics Workshop and the FDA/Industry Statistics Workshop, which has been including an increasing number of CMC-related statistics topics on their program list.

The role of statistics and statisticians will grow even more central as regulators move toward the risk-based approach to compliance embodied in PAT and QbD. With the door thus open for even broader and deeper use of statistical thinking and methods, industry statisticians and their employers will need not only to rethink their technical skills but also to learn to deploy PAT and QbD. As statisticians take on their expanded roles, they will need to expand their skills dramatically (See Table 4).

Many of these leadership skills are new to statisticians, and they will have to work to acquire them (Snee and Hoerl (2004)). Other new skills may play into some of the traditional strengths of statisticians. Among the most important leadership skills are the following:

- *Effective leaders not only know how to lead, they also have substantial business and regulatory knowledge.* In their new role, statisticians will need to understand how business works in general, how it works in the pharmaceutical industry, and how it works specifically for their companies. They will therefore need to understand strategic planning (the process for arriving at change objectives) and strategic deployment (the process of implementing strategy).
- *Leaders are comfortable with process and systems thinking.* In many ways, this is the easiest aspect of leadership for statisticians. After all, improving the way work gets done inevitably entails improving processes. Statisticians, steeped in analytic rigor, are uniquely positioned in that regard. Because leaders are also familiar with the proven, structured improvement methods, such as Six Sigma or Lean Sigma, that provide the practical means for improving processes, statisticians have a head start in these statistics-intensive methods. Statisticians are also most strongly positioned to develop sound statistical approaches and cogent quantifications relative to ideas put forth in new regulatory guidance documents, such as the "design space" concept. But it should be remembered that those methods also include such management and leadership skills

as project selection, project management, and results-tracking with clear metrics and milestones.

- *Leaders possess the so-called soft skills in abundance* (Snee (1998)). To be real leaders of critical projects, statisticians must know how to create stakeholders in a project, including those executive stakeholders and champions who are necessary for success. They must become adept at reviewing, coaching, and communicating. They must understand group dynamics, know how to lead teams, and know how to design and facilitate meetings. Projects, no less than grand strategy, require vision and direction for success. Statisticians must not only know how to set such direction but also communicate it concisely and clearly to their teams. In addition, they should be adept at removing barriers, like insufficient resources, lack of training, and inadequate time, that impede success. And they should remember some of those barriers lie in people's psyches—especially in their fear of change—and that the best way to overcome them is through coaching and counseling, not criticism.

With these leadership skills, coupled with technical expertise, statisticians should be ideally positioned to help deploy new risk-based compliance initiatives, to show their organizations where statistical methods can create value, and to create and sustain statistically driven continuous improvement with the aim of cutting costs, speeding time to market, and lightening regulatory burdens. The challenge is convincing others that statisticians are capable of being in these leadership positions. In a sense, current leaders are being asked to allow some newcomers into the game. This is not easy.

Pharmaceutical Case Studies

The four case studies discussed below provide a small sample of the types of problems encountered in the development and manufacturing of pharmaceuticals.

Case 1: Improving Yield for a Pharmaceutical Synthesis

As described in Aggarwal (2006), the yield from the synthesis of a small-molecule pharmaceutical (the API) was lower than desired (~40%). The laboratory decided to optimize the process using DoE. In a sequential experimental design, the initial design (a full factorial) was used to screen for significant factors. A follow-up experiment (a response surface study) was performed over the appropriate region to optimize the process, using factors suggested by the initial experiment. The factors studied in the screening experiment are illustrated in the following half normal plot of experimental results (Figure 5).

Solvent volume, the catalyst loading, and reaction temperature were clearly identified as having an impact on yield. Interaction plots (Figure 6) revealed some of the nature of the effects.

The labs used this information to generate a response surface design (a face-centered central com-

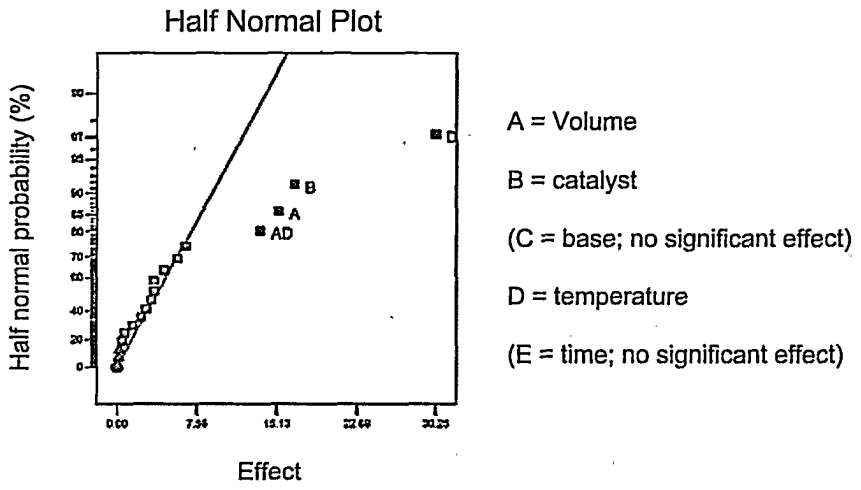
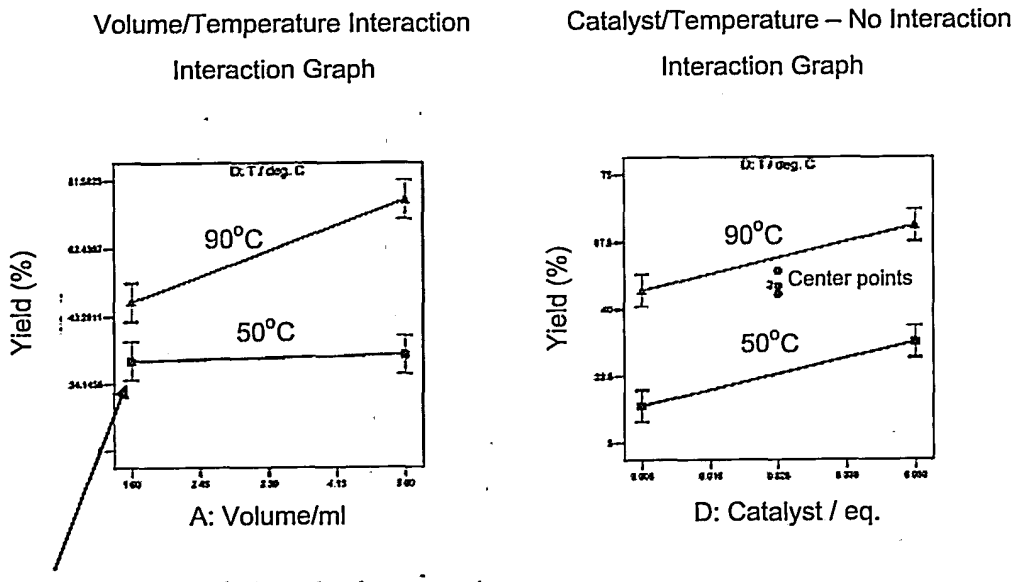


FIGURE 5. Half-Normal Plot of Effects from Pharmaceutical Synthesis Experiment.



Note: If error bars don't overlap the response is statistically significantly different.

catalyst = 0.5 eq.
base = 2.75 eq.
time = 6.5 h

volume = 3.30 ml
base = 2.75 eq.
time = 6.5 h

FIGURE 6. Plots for Volume–Temperature and Catalyst–Temperature Interactions.

posite design, to avoid setting levels in excess of the capability of the reaction), using the factors identified in the screening design. The labs also hoped to discover how to manage catalyst load using some of the other factors because the catalyst was expensive. The resulting response surface revealed that yield could be optimized with increases in temperature and solvent volume, while the amount of catalyst could be decreased due to the synergistic impact of these two factors. Final yields were confirmed to be in excess of 90% compared with average yields of approximately 40% that were observed prior to employing DoE. Thus, the lab doubled its capacity, while saving in cost by using less catalyst.

Case 2: Leadership for Design-Space Issue

As stated previously, the new regulatory concept of "design space" involves creating a multidimensional process capability region that has been demonstrated to provide "assurance of quality". Recently, in a large, multinational pharmaceutical company, various scientists (process chemists, pharmaceutical scientists, chemometricians, etc.) were meeting to formulate ideas about how to develop an approach to design-space construction within their company. Seeing this as an opportunity, some statisticians within the company decided to develop a probabilistic risk-based approach to design-space construction to address, in particular, the need to demonstrate "assurance of quality" as required by the ICH Q8 guidance.

The statisticians involved (in both the United States and Europe) worked together as an informal team and separately as individual consultants and members of matrix teams to formulate and build example Bayesian design spaces as defined in (1) in the Emerging Statistical Methods section. This approach was applied to problems in both (small-molecule) pharmaceutical and biopharmaceutical projects. Some prototype programs were written in R and SAS/IML, and then data from experiments were analyzed and presented to scientists and their management. Some external presentations and papers were generated. Eventually many in-roads were made to acceptance by the company scientists of this creative formulation of a design space. A simplified example of one of their reliability-based design spaces is shown in the probability surface plot in Figure 7.

Further progress toward company-wide accep-

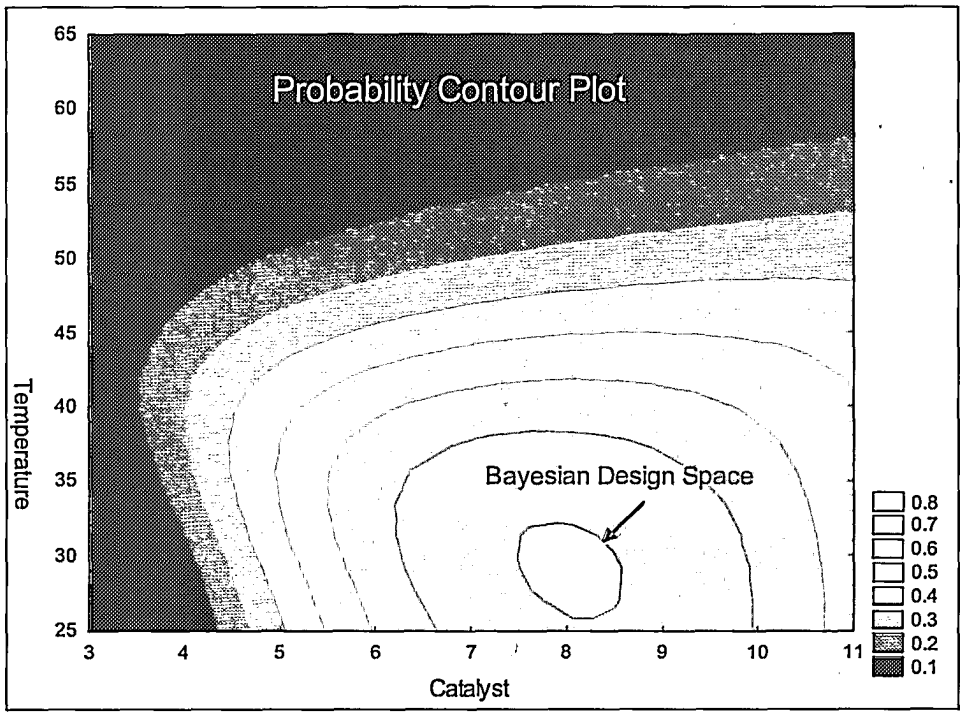


FIGURE 7. Reliability Contour Plot (See Level Scale on Right). The elliptical region in white is a prototype design space. This region has higher relative probability of simultaneously meeting all product-release specifications as a function of the temperature and catalyst factor levels. (Plot obtained from Stockdale and Cheng (2009). Used with permission from the Association for Quantitative Management.)

tance is still ongoing. Some future hurdles remain (e.g., addressing functional data). A definitive measure of success would be for such a design space to be submitted by the company, approved by regulatory authorities, and implemented in the manufacturing plant.

Case 3: Validation of a Complex Tedious Biological Assay

Validation of biological assays, particularly assays using large numbers of animals, is a tedious and expensive prospect. It's particularly important, as well, to make sure that a minimum number of animals are utilized for these purposes. An animal potency assay was developed for a vaccine product that uses 120 animals per run and takes 6 weeks to perform. The challenge was to obtain the maximum amount of information with minimum use of mice and time. The laboratory chose to include an assessment of the product distribution together with the assay validation and thereby gain valuable information about both assay and product variability. Ten lots of product were strategically identified, which were manufactured across a range of processing conditions, in order to best simulate the product distribution. Each lot was assayed twice in the bioassay, in a manner that would allow for independent estimates of inter-run variability throughout the experiment using the design shown in Figure 8.

A reference (Ref) was included in each run to calibrate bioassay results, while a clinically relevant lot (Clin) was included in order to set a specification for commercial materials tested in the bioassay. The reference was tested in duplicate in order to derive a criterion for "parallelism" in the bioassay according to Hauck et al. (2005). The validation results are illustrated in Figure 9.

Excellent precision was observed between replicate runs of the bioassay for all lots, while commercial materials performed as well as or better than lots tested in clinical trials (POC Lots 1–3 in Figure 9). The precision estimated from these data was used together with a minimum threshold to derive a release potency limit for lots of commercial material according to Apostol (2008).

Run									
1	2	3	4	5	6	7	8	9	10
Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Lot 1	Lot 1	Lot 2	Lot 4	Lot 4	Lot 5	Lot 7	Lot 7	Lot 8	Lot 9
Lot 2	Lot 3	Lot 3	Lot 5	Lot 6	Lot 6	Lot 8	Lot 9	Lot 10	Lot 10
Clin	Ref	Clin	Ref	Clin	Ref	Clin	Ref	Clin	Ref

FIGURE 8. Design for Biological Assay Validation.

Case 4: Regulatory Action and Use of Statistical Methods

A brief conceptualized version of a recent incident illustrates the leverage exercised by regulatory authority. A pharmaceutical company manufactures a tablet with two active ingredients. One of the ingredients is present in a very small absolute amount by weight. Small absolute deviations in the actual weight of this ingredient produce large relative (percent) deviations from the targeted dosage. As with most drug products, the specifications for this active ingredient are expressed as an allowable range of percentage of target dose. These ranges are typically 90 to 110% of target dose.

More important for this discussion, the manufacturer is required to report all OOS events to the FDA. These events are monitored and the agency has a series of escalating remedies to enact. Early among these is the right to send inspectors to the manufacturing facility at any time, announced or unannounced. The inspectors must have a stated purpose, but once an inspection begins, their audit trail may lead in many directions, particularly if irregularities are encountered along the way. In the example here, a series of OOS lots of product occurred over the course of a number of months—not every lot failed, but there were enough failures to cause notice. The facility was inspected; the OOS reports were audited, and the conclusion was reached by the FDA that the manufacturer did not understand the root cause of the failures and was therefore presenting an unmitigated risk to the public. Note that this is not a case of direct contamination or of an immediately harmful level of active ingredient. The regulator was principally concerned about the lack of apparent understanding demonstrated by the manufacturer.

The manufacturer's responses to the regulatory concerns were deemed unsatisfactory and the FDA determined to use its ultimate stricture—seizure of product at the manufacturing site. Federal marshals entered the manufacturing site and seized all lots of the particular product. Manufacturing of the product was halted. The FDA appointed a third party "monitor" to oversee the investigation of root cause, the implementation of process changes, and the assessment of new manufacture postimprovement. The third-party monitor was hired at the expense of the product manufacturer. Additionally, the manufacturer was required to post a substantial bond that would be forfeited in the event of insufficient cooperation and progress.

Where were the statisticians? The ultimate questions in this incident regarded patient safety. One slightly underdosed tablet may not be harmful, but when a medicine is taken daily to treat a chronic ailment (e.g., asthma or diabetes), the potential long-term effects are of chief concern. Therefore, the statisticians were asked to calculate probabilities such as, "How many defective tablets were likely to occur in a monthly supply of 30?" or "How many defective tablets might a patient obtain in a year's supply?" These probabilities depend on the failure rate in the manufacturing process. The underlying statistical questions then become "the best estimate of the tablet failure rate" and "the upper bound on the failure rate." The data suggest that the tablet failure rate, π, varies from batch to batch. Binomial mixture models are suggested, where π is a random effect associated with variations in the set-up, inputs, and conditions of each manufacturing run.

In terms of uncovering root causes for the failures, a primary question is "When did the problem begin?" This means extensive examination of time-series manufacturing data. An element of statistical interest is that the batch sizes vary and range from hundreds of thousands to millions of tablets,

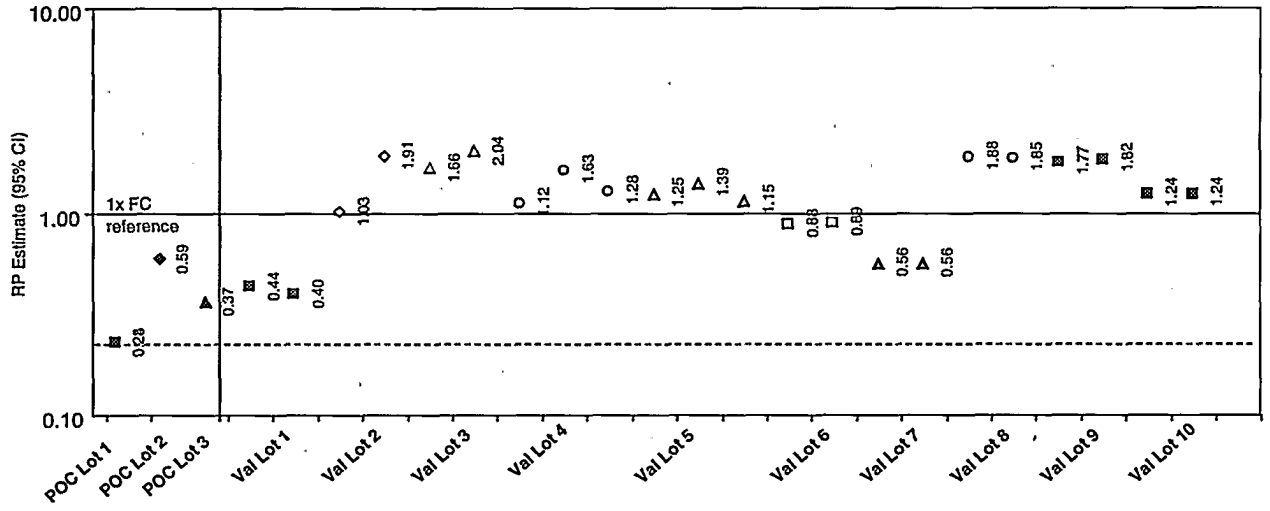


FIGURE 9. Results for Biological Assay Validation Study. (Note: run 4 of the design was repeated due to a performance issue related to the duplicate of the reference; thus, three replicates of lots 4 and 5 appear in the figure.)

and the number of tablets sampled from batches also varies—ranging from 20 to a few hundred. Control-chart techniques were heavily utilized. In addition, a variety of graphical techniques, including box plots and cusums, were helpful in explaining statistical conclusions to both regulators and company senior managers. Identification of a starting point allowed the manufacturer to identify process changes or incidents that occurred in the specified time frame. This generated hypotheses regarding ultimate causes. The statisticians were again involved in helping design experiments that would test those hypotheses. Resolution and prevention were largely engineering considerations. However, an extensively revised monitoring system was put in place—again with heavy emphasis on the use of control-charting techniques, including appropriate training and interpretation. Ultimately, the manufacturer was successful in reinstating their product in the marketplace.

Summary

Throughout most of the 20th century, pharmaceutical regulation was largely a matter of reactive legislation, but this century has seen the rise of science-based regulatory initiatives, including ICH guidelines on pharmaceutical development, pharmaceutical quality, the pharmaceutical quality system, and the FDA's guidance on Good Manufacturing Practice (GMP), Process Analytical Technology (PAT), and Quality by Design (QbD). These initiatives have brought a new emphasis on process understanding and quality and the use of statistical methods for both. These regulatory trends, as well as the inherent complexity of pharmaceutical manufacturing, economic pressures, and the increased need for technology transfer, will continue to accelerate the industry's need for sophisticated, statistics-driven approaches to quality and process understanding as well as the statisticians to apply them.

Today, pharmaceutical development, following the discovery of an active pharmaceutical ingredient (API), typically proceeds along several parallel paths, each requiring particular statistical techniques. These tools and statistical methods include Design of Experiments (DoE), screening experiments, optimization studies, regression modeling, process optimization, and robustness studies. Analytical-method development employs such tools as analysis of variance (ANOVA), variance-component studies, method ruggedness studies, and basic statistical techniques, including graphics. In manufacturing, statistical process control (SPC) is used extensively to monitor and improve processes. Statistical techniques found in Six Sigma, Lean Manufacturing, PAT, Design for Six Sigma, and QbD are also increasingly being used to improve processes.

The FDA's encouragement of the pharmaceutical industry to achieve better understanding of manufacturing processes and to quantify the risks associated with out-of-specification product has widened the set of statistical tools to include the use of multivari-

ate mathematical approaches, such as response surface methodologies, process simulation, and pattern recognition tools, in conjunction with knowledge-management systems. These emerging statistical methods are also driving the development of more powerful commercial statistical software, which is enabling nonstatisticians to do their own statistical analyses. Information technology is also making contributions to data mining, database management, and knowledge management, increasing the industry's interest in "enterprise" statistical software that can be used across an organization.

As statistical thinking and methods become even more critical for success in the industry, statisticians have begun to move from their traditional role as passive advisor to a more dynamic role as creator of value. Their role is becoming even more central as regulators move toward the risk-based approach to compliance embodied in PAT and QbD, requiring industry statisticians to rethink their technical skills, learn to deploy PAT and QbD, and acquire new leadership skills commensurate with the increased importance of their discipline within the industry. With those new skills and methods, they can become catalysts for both process-development understanding and product-quality improvement.

Acknowledgments

First, we would like to thank the Discussants for their valuable input and review of our paper. We would also like to thank Professor Enrique del Castillo (*JQT* Editor for 2006–2008) for inviting us to develop such an expository article. Finally, we thank Professor Joe Pigeon and Dr. Darryl Downing for their review of the (expanded) technical report version of this paper.

Appendix

Key Abbreviations and Acronyms

Abbreviations/ acronyms	Meaning
API	Active pharmaceutical ingredient
cGMP	Current Good Manufacturing Practice
CMC	Chemistry, manufacturing, and control
DCU	Dose-content uniformity
EMEA	European Medicines Agency
FDA	Food and Drug Administration
GMP	Good Manufacturing Practice
ICH	International Conference on Harmonisation
LC	Label claim
LOA	Limit of agreement
NDA	New drug application
OOS	Out of specification
PAT	Process analytical technology
PhRMA	Pharmaceutical Research and Manufacturers Association
QbD	Quality by design
QC	Quality control
RSD	Relative standard dSeviation
USP	*United States Pharmacopeia*

References

AGGARWAL, V. K.; STAUBITZ, A. C.; and OWEN, M. (2006). "Optimization of the Mizoroki-Heck Reaction Using Design of Experiment (DoE)". *Organic Process Research and Development* 10, pp. 64–69.

ALT, F. B. (2007). "A Bayesian Approach to Monitoring the Mean of a Multivariate Normal Process". In *Bayesian Process Monitoring and Optimization*, Colosimo, B. M. and del Castillo, E. (eds.), pp. 139–166. Boca Raton, FL: Chapman and Hall.

ANDERSON, M. J. and WHITCOMB, P. J. (1998). "Find the Most Favorable Formulations". *Chemical Engineering Progress* April, pp. 63–67.

APOSTOL, I., I.; SCHOFIELD, T. L.; KOELLER, G.; POWERS, S.; STAWICKI, M.; and WOLFE, R. A. (2008). "A Rational Approach to Setting and Maintaining Specifications for Biological and Biotechnology-Derived Products." *Biopharm International* 21(6), 21(7), 27(8).

BELL, S. B.; MORRISH, K.; HARRISON, F.; PETULLO, D.; THOMPSON, L. and GRAY, G. (2006). "Times 'R' a Changing: FDA Perspectives on 'Open Source' ". Presented at the *Joint Statistical Meetings*, August 6th, Seattle, WA.

BOULANGER, B.; DEWÉ, W.; GILBERT, A.; GOVAERTS, B.; and MAUMY, M. (2007). "Risk Management for Analytical Methods Based on the Total Error Concept: Conciliating the Objectives of the Pre-Study and In-Study Validation Phases". *Chemometrics and Intelligent Laboratory Systems* 86, pp. 198–207.

BOX, G. E. P. and LUCEÑO, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment.* New York, NY: John Wiley & Sons.

BUNCHER, C. R. and TSAY, J.-Y. (2005). *Statistics in the Pharmaceutical Industry*, 3rd edition. Boca Raton, FL: CRC Press.

CARSTENSEN, J. T. and NELSON, E. (1976). "Terminology Regarding Labeled and Contained Amounts in Dosage Forms". *Journal of Pharmaceutical Sciences* 65, pp. 311–312.

CHOW, S. C. and SHAO, J. (1991). "Estimating Drug Shelf-Life with Random Batches." *Biometrics* 47, pp. 1071–1079.

CLAYCAMP, H. G. (2008). "Room for Probability in ICH Q9: Quality Risk Management". Presented at the *Pharmaceutical Statistics 2008 Confronting Controversy* conference organized by the Institute of Validation Technology, March 18–19, Arlington, VA.

CLEOPHAS, T. J.; ZWINDERMAN, A. H.; and CLEOPHAS, T. F. (2006). *Statistics Applied to Clinical Trials*, 3rd edition. New York, NY: Springer.

COLOSIMO, B. M. and DEL CASTILLO, E. (2007). *Bayesian Process Monitoring and Optimization.* Boca Raton, FL: Chapman and Hall/CRC.

DEWÉ, W.; GOVAERTS, B.; BOULANGER, B.; ROZET, E., CHIAP, P.; and HUBERT, PH. (2007). "Using Total Error as Decision Criterion in Analytical Method Transfer". *Chemometrics and Intelligent Laboratory Systems* 85, pp. 262–268.

FAIRWEATHER, W. R.; MOGG, R.; BENNETT, P. S.; ZHONG, J.; MORRISEY, C.; and SCHOFIELD, T. L. (2003). "Monitoring the Stability of Human Vaccines". *Journal of Biopharmaceutical Statistics* 13(3), pp. 395–413.

FAYYAD, U.; GRINSTEIN, G. G.; and WIERSE, A. (2002). *Information Visualization and Knowledge Discovery.* London: Academic Press.

FDA. (2004). "Guidance for Industry: PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance".

FDA. (2004). "PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance".

FDA. (2006) "Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production".

GRIFFIN, J. P. and O'GRADY, J. (2006). *Textbook of Pharmaceutical Medicine*, 5th edition. Malden, MA: Blackwell Publishing.

HAUCK, W. W.; CAPON, R. C.; CALLAHAN, J. D.; MUTH, J. E. D.; HSU, H.; LANSKY, D.; SAJJADI, N. C.; SEAVER, S. S.; SINGER, R. R.; and WEISMAN, D. (2005). "Assessing Parallelism Prior to Determining Relative Potency". *Journal of Pharmaceutical Science and Technology* 59(2), pp. 127–137.

ICH. (1999). "ICH Harmonised Tripartite Guideline Q6A, Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances".

ICH. (2003). "ICH Harmonised Tripartite Guideline Q1E Evaluation of Stability Data".

ICH. (2005). "ICH Harmonized Tripartite Guideline: Pharmaceutical Development, Q8".

ICH. (2007). "ICH Harmonized Tripartite Guideline: Pharmaceutical Development: Annex to Q8".

ICH. (2005). "ICH Harmonized Tripartite Guideline: Quality Risk Management, Q9".

JEONG, M. K.; LU, J.-C.; and Wang, N. (2006). "Wavelet-Based SPC Procedure for Complicated Functional Data". *International Journal of Production Research* 44, pp. 729–744.

KIERMEIER, A.; JARRETT, R. G.; and VERBYLA, A. P. (2004). "A New Approach to Estimating Shelf Life". *Pharmaceutical Statistics* 3, pp. 3–11.

KOSSIK, J. (2002). "Think Small: Pharmaceutical Facility Could Boost Capacity and Slash Costs by Trading in Certain Batch Operations for Continuous Versions". Pharmamag.com, article ID/DDAS-SEX 52B/ http://www.pharmamanufacturing.com.

LEUENBERGER, H. (2001). "New Trends in the Production of Pharmaceutical Granules: Batch Versus Continuous Processing". *European Journal of Pharmaceutics and Biopharmaceutics* 52, pp. 289–296.

LIN, L. I. (1989). "A Concordance Correlation Coefficient to Evaluate Reproducibility". *Biometrics* 45(1), pp. 255–268.

LIND, E. E.; GOLDIN, J.; and HICKMAN, J. B. (1960). "Fitting Yield and Cost Response Surfaces". *Chemical Engineering Progress* 56, p. 62.

MACGREGOR, J. and BRUWER, M.-J. (2008). "A Framework for the Development of Design and Control Spaces". *Journal of Pharmaceutical Innovation* 3, pp. 15–22.

MILLARD, S. P. and KRAUSE, A. (2001). *Applied Statistics in the Pharmaceutical Industry.* New York, NY: Springer.

MINITAB, INC., 1829 Pine Hall Rd, State College, PA.

MORRIS, K. (2005). "Making the Most of Drug Development Data". *Pharmaceutical Manufacturing* April, p. 73.

NAIR, V.; TAAM, W.; and YE, Q. (2002). "Analysis of Functional Responses from Robust Design Experiments". *Journal of Quality Technology* 34, pp. 355–370.

NORDBROCK, E. (1992). "Statistical Comparison of Stability Study Designs". *Journal of Biopharmaceutical Statistics* 2, pp. 91–113.

PAMIAS, J. (2005). "PAT Success and Speed Hinge on Data Mining". PharmaManufacturing.com, at http://www.pharmamanufacturing.com/articles/2005/276.html.

PETERSON, J. J. (2004). "A Posterior Predictive Approach to Multiple Response Surface Optimization". *Journal of Quality Technology* 36, pp. 139–153.

PETERSON, J. J. (2008). "A Bayesian Approach to the ICH Q8 Definition of Design Space". *Journal of Biopharmaceutical Statistics* 18, pp. 959–975.

PETERSON, J. J.; SNEE, R. D.; MCALLISTER, P. R.; SCHOFIELD, T. L.; and CARELLA, A. J. (2009). "Statistics in Pharmaceutical Development and Manufacturing". DDS Technical Report, at www.biometrics.com (Statistical and Quantitative Sciences).

RAMSEY, J. O. and SILVERMAN, B. W. (2002). *Applied Functional Data Analysis.* New York, NY: Spriger-Verlag.

SAS INSTITUTE INC., Cary, NC.

SCHOFIELD, T. L. (2003). "Contributions on 'Assay Development' and 'Assay Validation' " in *Encyclopedia of Biopharmaceutical Statistics*, S.-C. Chow (ed.). New York, NY: Marcel Dekker, Inc.

SCHUIRMANN, D. J. (1987). "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability". *Journal of Pharmacokinetics and Pharmacodynamics* 15, pp. 657–680.

SKIBSTED, E. (2006). "Near Infrared Spectroscopy: The Workhorse in the PAT Toolbox". At www.spectroscopy.com, vol. 18, no. 5.

SNEE, R. D. (1998). "Non-Statistical Skills that Can Help Statisticians Become More Effective". *Total Quality Management Journal* 9(8), pp. 711–722.

SNEE, R. D. and HOERL, R. W. (2004). "Statistical Leadership—As Traditional Workplace Roles Change, Learn to Transition from Consultant to Leader". *Quality Progress* October, pp. 83–85.

SNEE, R. D. and HOERL, R. W. (2005). *Six Sigma Beyond the Factory Floor—Deployment Strategies for Financial Services, Health Care and the Rest of the Real Economy.* Upper Saddle River, NJ: Pearson Prentice Hall.

SPOTFIRE, INC. 212 Elm Street, Somerville, MA.

STATSOFT, INC., 2300 East 14th Street, Tulsa, OK.

STEINBERG, D. M.; BISGAARD, S.; DOGANAKSOY, N.; FISHER, N.; GUNTER, B.; HAHN, G.; KELLER-MCNULTY, S.; KETTENRING, J.; MEEKER, W. Q.; MONTGOMERY, D. C.; and WU, C. F. J. (2008). "The Future of Industrial Statistics—A Panel Discussion". *Technometrics* 50(2), pp. 103–127.

STOCKDALE, G. W. and CHENG, A. (2009). "Finding Design Space and a Reliable Operating Region Using a Multivariate Bayesian Approach with Experimental Design". *Quality Technology and Quantitative Management*, to appear.

UNITED STATES PHARMACOPEIA 28 (2005). "Validation of Compendial Methods," chapter 1225.

Discussion

ROGER W. HOERL

GE Global Research, One Research Circle, Niskayuna, NY 12309

I WOULD first like to thank the authors for writing this article and the editor for inviting me to comment on it. The authors have provided a window into the world of statisticians working in the pharmaceutical industry, beyond the scope of the clinical trials that we typically think of in pharma. They have made a clear case that statistical thinking and methods are just as integral outside of clinical trials as they are in them, and for this I would like to thank them.

From reading their paper, it is clear that this is a time of significant change in pharma, especially as it relates to statistics and statisticians. We therefore have a unique opportunity to expand our roles and become even more value adding to this industry. For example, the authors imply that the FDA is driving much of the change that is occurring today, especially that change related to better understanding of manufacturing processes. This is a good thing from the FDA, but begs the question of why the statistical community is following the FDA rather than leading it. Of course, in a regulated industry statisticians cannot dictate to the FDA, but I would hope that, in the future we could play a larger role in proactively influencing their thinking toward more modern, scientific approaches that have proven to be effective outside of pharma. As noted by the authors, new skills will be needed to assume this proactive role, as opposed to the more passive consulting role we have traditionally sought.

The authors also noted that the commercial statistical software widely available today allows scientists and others to perform many of their own analyses. This is certainly true and I think provides us with another opportunity. We can no longer "own" statistical methods, and that is a good thing, I think. For example, I have found the massive statistical training at GE that has occurred through its Six Sigma initiative to be liberating to professional statisticians. We no longer have to spend time doing routine analyses or explaining to our peers why you divide by $n-1$ instead of n. Rather, with people able to do the basics for themselves, we are freed to tackle the bigger and more challenging unsolved problems where our advanced expertise is truly needed. This allows us to have greater impact on the organization and provide greater value to the bottom line.

There are several indications given by the authors that we need to move quickly to establish more proactive, leadership roles for statisticians working in pharma. This is because, based on this paper at least, it appears the industry is lagging other industries in the United States in the areas of modern quality management and also in expanding the application and influence of statistical methods. For example, in their discussion of developments in the 21st century, the authors note: "Process development and manufacturing issues have only come to the forefront with recent FDA initiatives." Such topics have been the focus areas of the statistics profession in electronics, chemicals, and many other applications for years.

As noted previously, this begs the question of why we waited for the FDA to assume leadership and mandate us to develop our processes? Why didn't we, the statistics community, see the need for this and make a convincing case to our own management and the FDA for the overall benefit of process development and manufacturing excellence, relative to the investment required? Government regulations have certainly helped our profession in important ways, such as mandating statistically-based clinical trials and analysis. However, I am concerned that they have also slowed our progress, at least to some degree, in that we may have focused too much on meeting existing regulations rather than being proactive and pushing for improvement.

The regulatory action case study would seem to add credence to this point. Insufficient efforts had been made to understand the process, resulting in the FDA seizing product at the manufacturing site. As noted by the authors, "The regulator was principally

Dr. Hoerl is Manager of the Applied Statistics Laboratory at GE Global Research. He is an ASQ Fellow. His email address is hoerl@crd.ge.com.

concerned about the lack of apparent understanding demonstrated by the manufacturer."

As another example, the role of statisticians in Table 3 is listed as "statistical support". The authors note in the section on the Changing Role of Statisticians that we need to be moving from passive "support" roles to being equal collaborators, and eventually to being organizational leaders. I wholeheartedly agree. However, in the same case study, we read: "Resolution and prevention were largely engineering considerations." Engineers should certainly have a role in resolution and prevention of these manufacturing problems, perhaps the largest role. I would hope, however, that as equal collaborators statisticians would also have an important role. The engineers bring more subject-matter knowledge, but we bring better understanding of the scope and limitations of data analyses, as well as how to verify that prevention has actually been achieved.

In terms of quality management, the systems outlined in the Quality Control section seem to be primarily inspection based. Moving from product inspection to process control and process design was one of Deming's main points going back to the 1950s and 60s. Much of American industry has begun this journey, and while no one is so good at design and control that they don't need to think about checking the product, it would be fair to say that most industries have a strong focus on process design and control. Continuing to expand the focus in pharma to include process design, process understanding, and process control, as noted by the authors, needs to happen rapidly, to enhance public safety and also to bring down costs.

It is particularly disappointing to hear from the authors that the current system in pharma provides little incentive for continuous improvement and, in fact, some disincentive. The needs for regulatory approval for process changes and the risks associated with changes that do not work are real and must be considered. However, in other industries, even in regulated ones such as the food industry, statisticians have often been able to create a culture of continuous improvement, in conjunction with enlightened management (which doesn't have to be an oxymoron). While we can't create such an environment on our own, we shouldn't use this fact as an excuse and wait for someone else to initiate changes. Leadership is a choice.

As we think about our path forward as a profession in pharma, I would like to highlight the three arenas mentioned by the authors where we can and must show more leadership going forward. These are within our own companies, within the pharmaceutical industry, and in collaboration with statisticians at the FDA. I completely agree with the authors that within our own companies we need to take the initiative to exhibit more leadership behaviors, waiting for someone to stop us, rather than waiting for someone to ask us to lead. If we wait until we are asked to lead, we will be waiting a long time, I'm afraid! However, if we begin to proactively exhibit leadership by promoting modern approaches to process development and quality management, we will be helping our organizations be more competitive and may even catch management's attention along the way. We simply can't afford to wait until the FDA mandates continuous improvement.

Because driving organizational change can be so difficult and because there are industry-wide regulations to consider, it is also important for us to show leadership within the pharmaceutical industry as a whole, so that it can speak with one voice and drive broader change across the industry. The Pharmaceutical Research and Manufacturers of America (PhRMA) is one example of an industry-level organization within which statisticians can provide greater leadership, particularly in the area of promoting modern approaches to process development and control.

Eventually, we need to be working with FDA statisticians to determine the direction for change that will enhance both the cost and safety of pharmaceuticals. Rather than waiting for the FDA to determine what steps are needed, we should be taking the initiative to reach out to them and partner to develop the statistical approaches and methods that will be needed in the future. I would suggest that this partnership focus not only on statistical techniques per se, but also look at the broader question of how existing statistical technology can be best utilized for broad, organizational improvement. That is, I suggest we also research how to best integrate the available tools in our tool kit into overall strategies for improvement in pharma. I strongly believe that this broader, more holistic approach will provide significant value to both our organizations and society as a whole.

Discussion

JEFF HOFER

Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana 46285

I WOULD like to thank the authors for an excellent overview of the use of statistics in the pharmaceutical industry in the area of chemistry, manufacturing, and controls (often referred to as CM&C). The stated goal of the article was to provide an overview of the use of statistical thinking and methods in the R&D and manufacturing functions of the pharmaceutical industry, and the authors clearly achieved this objective. The authors covered a great deal of ground in their paper and my discussion will focus primarily on the future opportunities and challenges for CM&C statisticians, some of which were mentioned by the authors and a few others that were not discussed.

Themes for the Future

I would like to focus my discussion on themes for the future for CM&C statisticians. Many factors have and will continue to change the role of the CM&C statistician in the decade ahead. In reviewing the list, many, if not all, of these factors are likely similar challenges faced by many industrial statisticians, with some unique aspects linked to pharmaceuticals. A few of the most important factors I see facing pharmaceutical CM&C statisticians in the coming years include:

- the increasing use of PAT technology, resulting in the generation of large amounts of data yearning to be deciphered and converted into essential learning to reduce the cost of development for similar molecules or platforms and to improve the quality of manufactured pharmaceuticals,
- the increasing use of cross-functional teams of formulators, chemists, engineers, chemometricians, and statisticians collaborating to design, aggregate, and analyze data to identify fundamental relationships that can be leveraged across multiple compounds to speed the development process and improve processes. These fundamental learnings are often referred to as prior knowledge or institutional learning and will become a key focus area of interest at many companies in the coming years, if it is not already occurring,
- the increasing prevalence of outsourcing aspects of development and manufacturing,
- improvements in the availability and capabilities of user-friendly software to facilitate sophisticated analyses,
- the challenge of defining requirements for CM&C quality attributes that ensure acceptable product performance,
- developing and obtaining clear agreement among all parties involved on the interpretation of specification acceptance criteria for analytical properties,
- the development of more definition around the target level of assurance of quality within a design space, and
- the need for statisticians to continue to increase their leadership both internally and externally on key issues.

A brief discussion of each of these factors is provided.

As highlighted by the authors, the use of process analytical technology will increase the amount of data and also change the type of data that are encountered by industry statisticians. The data will be multivariate in nature and may also be time series in nature. One of the key goals of looking at data of this type will be data compression and the ability to develop models to identify and quantify the underlying relationships that exist between the responses of interest and the latent variables or fundamental material properties that are identified as the variables for reliable prediction. Statisticians should look to collaborate with formulators, analytical chemists, engineers, and chemometricians as spectroscopic calibration models are developed and validated allowing for rapid assessment of one or more analytical properties

Mr. Hofer is a Research Advisor. He is a member of ASA and ASQ. His email address is jhofer@lilly.com.

on large quantities of individual dosage units. The experimental design outlining what types of batches to include in the model-building process, the evaluation of the partial least-squares model-building options, and the development of the product-control strategy for the implementation of this technology are key areas where statisticians can contribute to the project team and ensure efficient model development and successful registration.

Another area where statisticians can contribute to key initiatives is in the collaboration toward the development of predictive models based on material attributes combining data across many compounds that use the same manufacturing platform. Similar to spectroscopic model building, the ability to work constructively on multidisciplinary teams will become an essential skill set for this endeavor. The goal of this effort is to generate prior knowledge that can be used to reduce the development time for new compounds. For successful completion of this activity, the ability to define and manage large data sets and combine and collectively model data from compounds using the same manufacturing platform will be necessary. The efficient use of experimental design techniques and multivariate data analysis will be the statistical tools of choice. Once again, the statisticians will benefit greatly by collaborating with chemometricians and engineers on the interpretation of the analysis results. If it is possible to build reliable models, development timelines may be reduced as optimal process conditions may be rapidly predicted with a few trials by using the prior knowledge of those projects that preceded it. Both technical and non-technical skills (e.g., influence) will be necessary to most effectively contribute to these large efforts.

As more and more companies outsource various aspects of development and manufacturing, the ability to work with third parties will increase in importance. The ability to define the essential data that must be generated and the data-format requirements will be important aspects for statisticians to rapidly support projects and collaborate with teams to efficiently identify the best formulation and process conditions. In manufacturing, the ability to create tools to obtain, analyze, and summarize third-party results as efficiently as possible will be key productivity enhancements that can minimize the impact of third-party manufacture and allow for rapid decision making.

In all situations, a key area for CM&C statisticians to focus on is the ability to concisely and clearly convey the results of their analysis to a wide variety of scientists and levels of management. The data-analysis methods will likely become more complex but it will be important to explain the findings in a manner that all parties can understand the benefits and what limitations, if any, exist for the models. An emphasis on good presentation skills will be an important aspect for success for statisticians and should not be overlooked as the nature of the data analysis becomes more complex.

The need for user-friendly software (in particular, multivariate and that related to the concept of design space) will be very important. The ability to easily manage the data and the capability to produce clear visualizations of the analysis results and sensitivities will be important considerations as these tools become more commonly used. Given the highly regulated nature of the pharmaceutical industry, one of the key aspects of software use is the ability for commercially available software to have validated built-in algorithms (such as those discussed in the design-space discussion) so that its use can be rapidly employed. The ability for vendors to effectively implement some of the design-space methodologies outlined in the article and other analysis methodologies (e.g., mixture designs, partial least squares) will be important to enabling statisticians to effectively collaborate with scientists on formulation, process, or analytical studies.

The authors mentioned that specifications are ideally established based on fitness for use. Currently, we do not explicitly accomplish this goal. In general, most acceptance criteria are established more based on the perceived process capability, based on the data provided at the time of submission. By doing so, it is an attempt to ensure fitness for use by striving for similarity to past data. An area of opportunity for statisticians in both development and clinical use is to develop approaches to determining specification-acceptance criteria that are more closely linked to true patient fitness for use. This is not an easy endeavor but is very much in line with the FDA goals. The ability to partner with clinical and pharmacokinetics experts in this area is an opportunity waiting to be tapped.

Another area ripe for resolution is the fundamental interpretation of acceptance criteria. This interpretation is currently not consistent among various regulatory guidances (reference USP, FDA OOS Guidance, ICH Q1E, Analysis of Stability Data) and is an area where logical alignment would facilitate

rational development of solutions. The bottom-line question of this issue is "To what do acceptance criteria refer?", or, put another way, "To what am I drawing an inference?"

As the authors noted, design-space development is an area where there is a great deal of opportunity. Some excellent work has been done in this area, as pointed out by the authors, and the development or enhancement of software to facilitate the implementation of these analysis methods will be necessary to bring these methods into common practice. The development of a clearer definition of the target level of assurance of the design space will assist in the ability to more consistently define design spaces from project to project.

The final theme of the future is that of leadership. As the authors correctly pointed out, the role of the statistician is changing in a positive way. CM&C statisticians simultaneously provide support across a wide variety of projects. This allows them to learn and see trends across compounds and/or platforms. This luxury is also a responsibility, as it is necessary for statisticians to become leaders in sharing best practices and driving changes across teams. This multiproject support and broad overview is one of the reasons that more is being expected of statisticians beyond just the design and analysis of data sets. "How can study designs be standardized across projects? How can study results be generalized? Can data from multiple projects be combined to find fundamental relationships that can be leveraged for future projects? Is it cost effective to invest in a new technology to improve the measurement system for a key analytical property? What is the most effective way to present a control strategy in a submission?" These are important questions that statisticians are well positioned to answer and we must be willing to step up and take on these challenges. In some cases, there may not be a clear solution but a sound methodical investigation into the possibilities is necessary and many of the skill sets we possess will be called on in the evaluation.

Concluding Thoughts

Once again, I would like to thank the authors for an excellent summary of the general drug-development process and the breadth of areas and types of data analyses supported by CM&C statisticians. The document highlights the exciting times that lie ahead for those statisticians willing to flex with the changing tide and learn new skills. The need for leadership by CM&C statisticians is paramount. Moving forward, more and more data will become available for analysis and the ability of the statistician to collaborate with scientists and leaders in deciphering and applying the relationships uncovered will be essential to improving the efficiency of both the development and manufacturing processes. The decade ahead will be full of opportunities for CM&C statisticians—it will be incumbent on us to be proactive and seek out unique ways to have a positive impact on improving the quality of information generated and reducing the time to develop such information. For those graduate statistics students looking for a challenging and rewarding career, the area of industrial statistics in the pharmaceutical industry should be one that is considered, as many opportunities exist for talented and dynamic statisticians.

Acknowledgments

I would like to thank my colleagues Kristi Griffiths and Gary Sullivan for their input and discussion on the future direction and opportunities for CM&C statisticians.

Discussion

LYNN TORBECK

Torbeck and Assoc., Inc., 2000 Dempster Plaza, Evanston, Il 60202

THE AUTHORS are to be commended for an excellent paper that will serve as an introduction and overview of current statistical quality control in the pharmaceutical industry. It will be a valuable reference for statisticians new to the industry and for scientists, engineers, and managers taking on statistical leadership. The following discussion supports the article and presents more specific issues and topics for further consideration.

The quality of most consumer products is readily apparent upon inspection by the buyer. Thus, the voice of the customer is available for the company to use for improvement. The quality of a pharmaceutical product is more difficult to determine. Who is the customer? The patient cannot, for the most part, determine if the pill taken is of high quality. Given that the placebo effect is almost one third for some situations, getting better is not proof of efficacy. This has lead to increased counterfeiting of drugs. The fact is that few people know if the drug product is of high quality. The patient doesn't know, the pharmacist doesn't know, and the doctor doesn't know. They all assume it is. Only the manager of the quality-control analytical laboratory really has the information at hand to determine if the product is fit for use. Coupled with the ever needed and thankfully present regulations and agency inspections, quality control of pharmaceutical products is clearly not the same as making toaster ovens. While activities like process validation are done at nonpharmaceutical companies, it is required for our industry. As is corrective action /preventive action (CAPA) for out-of-specifications (OOS) results. Other examples abound. Our activities and statistics should reflect the differences of our industry.

While the industry is moving forward with new and advanced statistical tools, as noted in the article, we need to assure that the simple tools are still being used correctly. Hundreds of small and midsized pharmaceutical companies do not have statisticians or statistically trained engineers and are in need of industry and agency assistance in doing simple routine statistics correctly. Regulatory agencies assist by enforcing the requirements for training. For example, many quality-control staff members routinely use attribute- and variables-sampling plans for incoming, in-process, and outgoing inspection. But they have never been trained in the theory of sampling, design, or the correct implementation of the plans. They are expected to learn it on the job, often by just reading a standard operating procedure. One can only guess the number of lots that have been rejected that were acceptable or the number of lots that were unacceptable but shipped. Even the act of data collection needs to be carefully monitored. Once a sample has been taken from a batch, there is no way to know if it is truly representative. Only by watching the sample being taken can assurance be made. Detailed standard operating procedures are needed to support this.

According to available information, OOS results still are a major source of regulatory citations and recalls. Industry statisticians need to join forces with the agencies to codify scientific and statistically valid approaches to determining specification criteria that reflect the actual medical and manufacturing situations. Wishful thinking by any party doesn't serve the patient or change the process. As an example, Bergum's method (1990) can give real insight to meeting criteria for as many as 14 USP tests that have multiple criteria.

ICH Q9 on risk management has as its core concept the risk to the patient. It does a good job of discussing risk management but surprisingly doesn't discuss the risk to the patient for type II or beta errors. It also doesn't discuss the "consumer" risk for the limiting quality (LQ) for sampling plans. This should be part of the discussion on setting specification criteria and determining sample sizes for testing.

The authors are to be commended for raising the

Mr. Torbeck is Principal Statistician. He is a member of ASA and a Senior Member of ASQ. His email is Lynn@Torbeck.org.

topic of designed experiments in general and mechanistic models specifically. The industry should look more closely to teaching and implementing mechanistic model building. Many scientists and engineers are put off by the use of generic empirical models, as they aren't "scientific" enough. They want to work from first principles. More journal articles and successful case studies are needed.

Statisticians, including this one, said for years that the square root of n plus one was not a valid sampling plan because we could not find a statistical authority to reference. Then we discovered that it does indeed meet the definition of a sampling plan as given in ANSI/ASQ Z1.4 (2003), section 9.1. "A sampling plan indicates the number of units of product from each lot or batch which are to be inspected (sample size or series of sample sizes) and the criteria for determining the acceptability of the lot or batch (acceptance and rejection numbers)." Further, calculation of the sample sizes finds them to be very close to the sample sizes for Z1.4, General Inspection Level I. Because the OC curves can be calculated as well, it is a valid sampling plan and can be used with the same care and caution as any other sampling plan. Statisticians need to assist in implementing this correctly.

Statisticians outside of biology, chemistry, and the pharmaceutical industry are surprised to find the wide-spread use and abuse of the percent relative standard deviation, %RSD, better known to statisticians as the coefficient of variation (CV). A statistic that statisticians love to hate, it has a number of properties that make it less than desirable. Yet, over the last couple of decades, it has become even more popular and widely abused. Laboratory analysts routinely add, subtract, and average %RSD's with abandon. It is almost universally used as a replacement for the standard deviation; often times for calculating confidence intervals and t-tests. It is also widely misinterpreted. For example, a constant RSD over different levels of concentration is perceived as not a change in the variability. One can only wonder at the number of incorrect decisions as a result. Given that this abuse is nearly 100%, it is not clear how industry statisticians can correct it in this generation.

Another statistical tool subject to wide-spread abuse is the sample average plus and minus three times the sample standard deviation without consideration of the sample size. It is used as a confidence interval, specification setting criteria, looking for statistical significance, identifying outliers and other statistical questions. The worst offense is when non-U.S. regulatory agencies require companies to set specification criteria using it with sample sizes as small as five. The probability of rejecting good lots in the future can be high. Cpk, with its multiplier of three, is a related example. This writer has seen companies accept and reject lots of product based on Cpk values calculated from small samples.

While applied statistics must be pragmatic, it cannot be incorrect. Good science and good manufacturing practice regulations demand that the correct tools be used correctly for a given problem. Using %RSD and $\overline{X} \pm 3S$, as in the above discussion, is incorrect statistically, particularly so for small sample sizes. Given the wide-spread abuse, the topics beg to be clarified. While advanced statistics bring benefits, the misuse of simple statistics endangers patients.

Conclusion

To conclude, the authors have presented a well-written article that is on target for its intended audience. With that as a foundation, we, as an industry, also need to address some of the areas discussed above.

Reference

BERGUM, J. S. (1990). "Constructing Acceptance Limits for Multiple Stage Tests". *Drug Development and Industrial Pharmacy* 16(14), pp. 2153–2166.

Discussion

YI TSONG, ROSWITHA KELLY, MEIYU SHEN, and JINGLIN ZHONG

Division of Biometrics VI, Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, FDA, Silver Spring, MD 20993

WE COMMEND the authors' efforts to document the historical development of pharmaceutical quality and manufacturing controls as well as the statisticians' past roles and future opportunities in this area. We recognize that advances in quality and manufacturing controls are continuously evolving and, hence, it is difficult to cover all of them. However, there are three statistical concepts that play important roles in modern approaches to quality assessment and control that are relevant to this paper. The first concept is the shift from accepting a null hypothesis that cannot be rejected to proper hypothesis testing. This concept is demonstrated when "proving quality (parallelism, equivalence)" by rejecting the proper null hypothesis instead of "assuming quality (parallelism, equivalence) because one is not able to prove otherwise". The second concept deals with defining quality in terms of "a high proportion of the product being within desired specifications" or "a low proportion of the product being outside the desired specifications" instead of describing quality by "its attribute and variable measurements in a non-cohesive manner", which is currently the approach taken by the *U.S. Pharmacopeia*. The third concept relates to achieving quality control by "understanding the relationship between the finished product and the influencing factors" instead of only "understanding how the quality assurance tests relate to the quality of the product". A more detailed treatment of recent improvements in statistical approaches to these three areas may enhance the article. We provide an example and additional recent references for each concept.

Recent work in the analysis of stability data and shelf-life estimation provides an example of the advancement to "proving quality" instead of the historical approach of "accepting quality if the null hypothesis is not rejected". As early as 1992, Ruberg and Stegeman (1991) and Ruberg and Hsu (1992) questioned the efficiency of the conventional poolability test based on "accepting the null hypothesis" when there was insufficient evidence to show a difference. Ruberg and Hsu (1992) proposed to pool batches after demonstrating "equivalence of slopes". Although their proposed "equivalence of slopes" was not considered feasible at that time (Lin and Tsong (1991)), the research interest of "proving poolability" continued. Yoshioka et al. (1997) suggested a "shelf-life equivalence" method, which Tsong et al. (2003) revised. The latter authors also developed a batch-pooling method by establishing that all batches have equivalent content at the targeted shelf life. Liu et al. (2006, 2007) expanded on the content-equivalence approach. Djira et al. provided, in 2008, a good survey and review of equivalence approaches to batch pooling.

The second concept deals with defining quality in terms of "a high proportion of the product being within desired specifications" (or: "a low proportion of the product being outside the desired specifications"). instead of describing quality by "its attribute and variable measurements in a noncohesive manner". Over the past 10 years, such methods have been seen in the area of specification and sampling acceptance testing as well as in method-transfer studies. In specification and sampling acceptance testing, the conventional approaches were often based on two criteria, one for variance and one for zero tolerance. For example, the *U.S. Pharmacopeia* XXIV sampling acceptance plan is a two-stage sampling plan. In the

Dr. Tsong is a Deputy Division Director and Acting CMC Team Leader. He is a member of ASA and ICSA. His e-mail address is yi.tsong@fda.hhs.gov

Ms. Kelly is a senior mathematical statistician. She is a member of ASA. Her email address is Roswitha.Kelly@fda.hhs.gov.

Dr. Shen is a senior mathematical statistician. She is member of ICSA. Her email address is meiyu.shen@fda.hhs.gov

Dr. Zhong is a mathematical statistician. She is a member of ASA. Her email address is jinglin.zhong@fda.hhs.gov.

first stage, a sample of 10 tablets is assayed. The lot complies with the dose-content uniformity (DCU) requirement if dose content of each of the 10 tablets is within 85%–115% label claim (%LC) and the relative sample standard deviation (RSD) is $\leq$6%. It fails to comply if more than one tablet has dose content outside (85%, 115%) LC or if at least one tablet has dose content outside (75%, 125%) LC. If the lot neither passes nor fails at this point, one moves to the second stage. In the second stage, an additional 20 tablets are randomly sampled and assayed. The lot complies with the DCU requirement if the dose content of each of the 30 tablets is within (75%, 125%) LC, if no more than one tablet has dose content outside (85%, 115%) LC, and if the RSD $\leq$7.8%. Otherwise, the lot fails the DCU test. This counting rule is for the fixed sample sizes at the two stages. Hauck and Shaikh (2001) suggested that the determination of specification and sampling acceptance procedures be based on the tolerance-interval approach. Tsong et al. (2004) also used a tolerance-interval approach for a multitiered sampling acceptance procedure for the dissolution test. Further, the notion of an alpha spending function associated with group-sequential designs was first mentioned in connection with multitiered sampling acceptance procedures. Tsong and Shen (2007b) proposed a two-tiered sampling acceptance procedure based on the tolerance interval where the probabilities of the overfill end and the underfill end were controlled separately by using 2 one-sided tests (Schuirmann (1987)). An FDA working group of statisticians and chemists also adopted the 2 one-sided tests in the acceptance sampling procedure (Tsong et al. (2008)) for delivery dose uniformity of inhalation sprays and intranasal products.

Traditional assessment of method equivalence has relied on concordance analysis. In recent years, it has been recognized that the concordance statistic is not a good measurement agreement between new and initial methods. More suitable methodology focusing on individual measurement agreement has been developed. First Bland and Altman (1986) proposed the limit of agreement (LOA) method to describe the individual agreement between two clinical measurements. In 1999, Bland and Altman extended the LOA method to more complicated situations. Lin et al. summarized the tools for measurement agreement in 2002. Zhong and Shao (2003) and Shao and Zhong (2004) proposed an approach that combined the mean and variances of individual differences into one statistic. In 2008, Zhong et al. proposed to also adopt the 2 one-sided tolerance limits approach (Tsong and Shen (2007a)) to assess the exchangeability of two test methods.

References

BLAND, J. M. and ALTMAN, D. G. (1986). "Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement". *Lancet* 8, pp. 307–310.

BLAND, J. M. and ALTMAN, D. G. (1999). "Measuring Agreement in Method Comparison Studies". *Statistical Methods in Medical Research* 8, pp. 137–160.

DJIRA, G. D.; HOTHORN, L. A.; and TSONG, Y. (2008). "Equivalence Tests for Shelf Life and Average Drug Content in Stability Studies". *Journal of Biopharmaceutical Statistics* 18(5), pp. 985–995.

HAUCK, W. W. and SHAIKH, R. (2001). "Sample Size for Batch Acceptance from Single and Multitier Designs Using Two-Sided Normal Tolerance Interval with Specification". *Journal of Biopharmaceutical Statistics* 11, pp. 335–346.

LIN, L.; HEDAYAT, S.; SINHA, X.; and YANG, X. (2002). "Statistical Methods in Assessing Agreement: Models, Issues, and Tools". *Journal of the American Statistical Association* 97, pp. 257–270

LIN, T.-Y. D. and TSONG, Y. (1991). "Determination of Significance Level for Pooling Data in Stability Studies". *Proceedings of Biopharmaceutical Section of Joint Statistical Meetings, American Statistical Association*, pp. 195–201.

LIU, J.; TUNG, S.; and PONG, Y. (2006). "An Alternative Approach to Evaluation of Poolability for Stability Studies". *Journal of Biopharmaceutical Statistics* 16, pp. 1–14.

LIU, W.; JAMSHIDIAN, M.; YHANG, Y.; BRETZ, F.; and HAN, X. L. (2007). "Pooling Batches in Drug Stability Study by Using Constant-Width Simultaneous Confidence Bands". *Statistics in Medicine* 26, pp. 2759–2771.

RUBERG, S. J. and HSU, J. C. (1992). "Multiple Comparison Procedures for Pooling Batches in Stability Studies". *Technometrics* 34, pp. 465–472.

RUBERG, S. J. and STEGEMAN, J. W. (1991). "Pooling Data for Stability Studies: Testing the Equality of Batch Degradation Slopes". *Biometrics* 47, pp. 1059–1069.

SCHUIRMANN, D. J. (1987). "The Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability". *Journal of Pharmacokinetics and Biopharmaceutics* 15, pp. 657–680.

SHAO, J. and ZHONG, B. (2004). "Assessing the Agreement Between Two Quantitative Assays with Repeated Measurements". *Journal of Biopharmaceutical Statistics* 14(1), pp. 201–212.

TSONG, Y.; CHEN, W.; and LIN, T. D. (2003). "Shelf Life Determination Based on Equivalence Assessment". *Journal of Biopharmaceutical Statistics* 13, pp. 431–449.

TSONG, Y. and SHEN, M. (2007a). "An Alternative Approach to Assess Exchangeability of a Test Treatment and the Standard Treatment with Normally Distributed Response". *Journal of Biopharmaceutical Statistics* 17, pp. 329–338.

TSONG, Y. and SHEN, M. (2007b). "Parametric Two-Tier Sequential Assurance Test of Dose Content Uniformity". *Journal of Biopharmaceutical Statistics* 17, pp. 143–157.

TSONG, Y.; SHEN, M.; LOSTRITTO, R. T.; and POOCHIKIAN, G. K. (2008). "Parametric Two-Tier Sequential Quality As-

surance Test of Delivery Dose Uniformity of Multiple-Dose Inhaler and Dry Powder Inhaler Drug Products". *Journal of Biopharmaceutical Statistics* 18(5), pp. 976–984.

TSONG, Y.; SHEN, M.; and SHAH, V. P. (2004). "Three-Stage Sequential Statistical Dissolution Testing Rules". *Journal of Biopharmaceutical Statistics* 14(3), pp. 1–32.

YOSHIOKA, S.; ASO, Y.; and KOJIMA, S. (1997). "Assessment of Shelf-Life Equivalence of Pharmaceutical Products". *Chemical & Pharmaceutical Bulletin* 45(9), pp. 1482–1484.

ZHONG, J.; LEE, K.; and TSONG, Y. (2008). "Statistical Assessment of Analytical Method Transfer". *Journal of Biopharmaceutical Statistics* 18(5), pp. 1005–1012.

Rejoinder

JOHN J. PETERSON, RONALD D. SNEE, PAUL R. McALLISTER,
TIMOTHY L. SCHOFIELD, and ANTHONY J. CARELLA

THE AUTHORS wish to thank the discussants for their time and contribution to this article. They have reinforced many of our views and raised some important issues. Through our response, we hope to summarize the broad views expressed throughout the article and among the discussants.

Dr. Hoerl states that "... the authors imply that the FDA is driving much of the change that is occurring today, especially that change related to better understanding manufacturing processes. This is a good thing from the FDA, but begs the question of why the statistical community is following the FDA, rather than leading it". This is an important question that requires additional comment and context.

The transformation going on within the industry is indeed highlighted by the call from regulators to build quality into the manufacturing process rather than to test quality into the product. There is also a call from regulators to change the manner in which pharmaceutical-development and manufacturing practices are communicated to authorities, through documentation of risk analyses, development experiments, and process-maintenance strategies in regulatory filings. In a highly regulated environment, however, there is a great deal of organizational momentum to continue doing whatever has been approved in the past. When an application is submitted to the FDA for approval, a vast amount of resources have already been invested by the company. Because the company's market exclusivity is limited by patent expiry, each day of delay to market costs tremendously in lost opportunity. This creates pressure to use the techniques, approaches, and statistical arguments that have been used previously. The chemical, manufacturing, and control (CMC) sections of a new drug application are primarily reviewed by relevant scientists/engineers and only infrequently reviewed by FDA statisticians.

Thus, a confluence of effects unite to perpetuate traditions that, in some cases, may benefit from improved statistical thought. We believe that these realities partially underlie the difficulty pharmaceutical CMC statisticians have in providing breakthrough leadership in the industry. This environment produces pressures that challenge change.

It is hoped that this article will inspire industrial statisticians and industrial engineers to collaborate with pharmaceutical scientists, chemical engineers, and government regulators to apply creative strategies to provide continuous improvement for pharmaceutical development and manufacturing in the face of these issues.

These pressures notwithstanding, CMC statisticians have historically collaborated within their companies, across institutions, and together with their regulatory counterparts to define the process and to measure the properties of manufactured materials. The tools espoused in Six Sigma have long been used by pharmaceutical CMC statisticians. The call to quality by design clearly elevates these to the standing of other scientific tools that have been used historically to develop safe and effective drugs and vaccines.

The discussants have aptly viewed statisticians as having a leadership role in the pharmaceutical industry. While this has long been the case in many areas of nonclinical development and control of pharmaceuticals, the FDA call for better process understanding and control, as well as the advent of information-rich technologies, has reinforced the need for statistical leadership. As discussed in this paper, pharmaceutical CMC statisticians have shown leadership through the PhRMA CMC Statistics Expert Team and the annual FDA/Industry Statistics Workshop. In addition, statisticians can show, and indeed have shown, leadership through presentations to the FDA and other regulatory bodies on specific drug-application issues. Some pharmaceutical industry statisticians publish improved statistical methods in relevant biopharmaceutical statistics and other scientific journals. We believe that pharmaceutical-industry statisticians and FDA statisticians have communicated well with each other in the past decade. However, perhaps a more impor-

tant question is how well have pharmaceutical industry statisticians and FDA statisticians been able to influence their own (nonstatistician) scientific collaborators at their respective institutions. Such influence is less visible directly but will show itself over the years as new industry guidance on CMC-related issues is developed and implemented. Much like the need for statistical leadership in the application of sound clinical study design and data interpretation in the 20th century, the need for leadership and support from CMC statisticians has been reinforced through regulatory opportunities, in this case, through the advent of quality by design in the 21st century.

Dr. Hoerl's comment that "It is particularly disappointing to hear from the authors that the current system in pharma provides little incentive for continuous improvement, and in fact some disincentive" was likely drawn from a statement that currently pharmaceutical manufacturers have little business incentive for continuous improvement after regulatory approval. This is precisely one of the reasons for establishing a pharmaceutical process design space, so that continuous improvement activities can take place within this defined region without time-consuming regulatory approval after a process change. Industry and regulatory statisticians have been communicating regularly on this issue.

Application of statistical tools related to the definition of a pharmaceutical process design space has become the focus of many activities undertaken by the PhRMA CMC Statistics Expert Team, including the introduction of special sessions devoted to the topic at the ASA-sponsored FDA/Industry Statistics Workshop. Furthermore, the recent articles by Peterson (2008) and Stockdale and Cheng (2009) make specific proposals for design space development. Mr. Hofer challenges the statistical community to develop "more definition around the target level of assurance of quality within a design space." We agree and believe that the referenced articles will help to provide some quantitative clarification around this issue.

We need to keep in mind that, while changes to the registered manufacturing process are subject to regulatory oversight, operational changes may be, and should be, implemented in the manufacturing environment in order to improve operations and reduce costs. Thus, processes such as inventory management, document flow, and cycle time may be studied and improved using Lean and Lean Six-Sigma methods.

Mr. Torbeck acknowledges the utility of designed experiments in deriving process understanding and defining design space, but states "The industry should look more closely to teaching and implementing mechanistic model building. Many scientists and engineers are put off by the use of generic empirical models, as they aren't 'scientific' enough." Clearly, more work needs to be done with regard to mechanistic models and models for functional data, as well as software for implementation. The basic theory, methods, and software exist for fitting (mechanistic) models that are nonlinear (Gallant (1987), Bates and Watts (1988), Seber and Wild (2003)). Optimal design theory for estimating parameters of nonlinear mechanistic models also exists (Seber and Wild (2003)), but easy-to-use software is just starting to become available (e.g., SAS/JMP®). As mentioned previously in the section of our paper on The Impact of Statistical Software and Information Technology, "... software availability will guide what statisticians and quality engineers are willing to do with statistics ...". As such, mechanistic models provide an important area for further statistical-software development, particularly in the area of optimal design.

The authors agree with Mr. Torbeck that "Industry statisticians need to join forces with the agencies to codify scientific and statistically valid approaches to determining specification criteria that reflect the actual medical and manufacturing situations," and with Mr. Hofer that "A few of the most important factors I see facing pharmaceutical CM&C statisticians in the coming years include ... the challenge of defining requirements for CM&C quality attributes that ensure acceptable product performance." As Mr. Hofer points out, our paper challenges statisticians to work together with their development and regulatory colleagues to define specifications on the basis of fitness-for-use (for the patient) rather than indirect process capability limits (e.g., dissolution in a vessel).

In this regard, an expanded view must be taken of Mr. Torbeck's statement that "Only the manager of the quality-control analytical laboratory really has the information at hand to determine if the product is fit for use." While it's true that the analytical laboratory possesses information regarding the properties of their methods and the manufacturing process, quality is defined by the customer. The CMC statistician has the role of collaborating with the analytical laboratory, clinical researcher, and regulatory authorities to help provide assurances to the patient that a drug or vaccine is safe and effective.

Mr. Hofer states that we must achieve "clear agreement among all parties involved in the interpretation of specification acceptance criteria for analytical properties." Parameters such as quality attributes of clinical materials and shelf-life determination are directed toward the true mean of development materials, while some regulatory guidances urge manufacturers to regard individual measurements against specification acceptance criteria. This creates confusion and hinders the statistician's ability to implement "scientific and statistically valid" approaches for maintaining product quality.

Dr. Tsong et al. have made some very important points, worth reinforcing here. Nonstatisticians, and even some statisticians, make the mistake of not heeding the dictate that "absence of evidence is not evidence of absence." Dr. Tsong highlights contributions to stability assessment; indeed, there are many other opportunities to reformulate the null and alternative hypotheses and apply an equivalence approach. Another application is method validation, which seeks to establish conformance of assay properties to acceptance criteria, while method transfer looks to establish equivalence of performance between laboratories. In bioassay, the difference between, or ratio of, slopes of the test and reference materials needs to conform to an acceptable range in order to help assure the accuracy of the potency measurement. These and many other applications require an equivalence approach to hypothesis testing in order to appropriately address the research objective.

Regarding Tsong et al.'s second point, we note that the approaches to design space described in Peterson (2008) and Stockdale and Cheng (2009) do indeed quantify quality in terms of "a high proportion of the product being within specifications." Various industrial problems may require different approaches to defining the experimental units, but we agree that there is a need for good probabilistic-based risk assessment.

It is our hope that this article and the contributions of the discussants have clearly communicated that the success of pharmaceutical CMC statisticians, in improving the efficiency and effectiveness of pharmaceutical manufacturing, depends on a combination of factors. In addition to technical skills, industry statisticians must hone their leadership and collaborative skills. Decision science must join the other sciences routinely employed by scientists and engineers, as well as by statisticians.

We fully support Mr. Hofer's statement that the pharmaceutical industry now provides an excellent opportunity for graduate students interested in a career in applying industrial statistical methods. We believe that, for those statisticians in industrial engineering and applied statistics departments in academia, the pharmaceutical industry, now more than ever, presents many interesting and fruitful opportunities for research and consulting in the area of quality and efficiency improvement.

References

BATES, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications.* Hoboken, NJ: John Wiley and Sons, Inc.

GALLANT, R. A. (1987). *Nonlinear Statistical Models.* Hoboken, NJ: John Wiley and Sons, Inc.

SEBER, G. A. F. and WILD, C. J. (2003). *Nonlinear Regression.* Hoboken, NJ: John Wiley and Sons, Inc.