Drug Information Journal, Vol. 31, pp. 1157–1166, 1997
Printed in the USA.
0092-8615/97

GOOD STATISTICS PRACTICE IN THE DRUG DEVELOPMENT AND REGULATORY APPROVAL PROCESS

SHEIN-CHUNG CHOW, PHD

Executive Director, Biostatistics and Data Management, Covance, Inc., Princeton, New Jersey

During the development and approval process of a new drug, the concept of good statistics practice (GSP) is necessarily implemented. GSP is a set of principles which assures the validity of the design and analysis of the intended studies conducted at various stages of the process of drug development and regulatory approval. GSP provides a fair assessment of the drug product with the desired accuracy, precision, and reliability. In essence, GSP not only concerns the validity of statistical inference regarding drug efficacy and safety, but also provides assurance of the proper identity, strength, quality, purity, and stability of the drug product. This paper describes regulatory requirements for statistics and the role of statistics in the drug development and regulatory process. The concept and importance of GSP are illustrated through some practical and/or regulatory statistical issues that commonly occur during the drug development and regulatory approval process.

Key Words: Code of Federal Regulations; United States Pharmacopeia and National Formulary; Good laboratory practice; Good clinical practice; Current good manufacturing practice; Good research practice; Good statistics practice

INTRODUCTION

THE DEVELOPMENT OF a pharmaceutical entity is a lengthy process involving drug discovery, formulation, laboratory development, toxicological studies, clinical development, and regulatory registration. The whole process is not only time consuming but also very costly. For example, for drug discovery, it may require the screening of a large number of compounds in order to obtain a few potential promising compounds. In practice, it is very likely that these potential promising compounds may never reach other stages of drug development such as animal studies or clinical evaluation due to the toxicity or intolerability in animals or humans. After a promising compound is obtained, the clinical development and regulatory approval process is also time consuming, which is necessary to assure the effectiveness and safety of the promising compound.

For the development of pharmaceutical products, different countries, such as the United States, the European Community (EC), and Japan have similar but slightly different sets of regulatory requirements. This paper focuses on the United States regulatory requirements for drug development. The United States regulations for drug development are developed based on the Federal Food, Drug and Cosmetic Act (FD&C Act) passed in 1938. The FD&C Act requires pharmaceutical companies to submit full reports of investigations on the safety of new drugs. In 1962, the significant Kefauver-Har-

Presented at the DIA "First International Taipei Symposium," August 29–30, 1996, Taipei, Taiwan.

Reprint address: Shein-Chung Chow, PhD, Biostatistics and Data Management, Covance, Inc., 210 Carnegie Center, Princeton, NJ 08540.

ris Drug Amendments to the FD&C Act were passed. The Kefauver-Harris Amendments not only strengthen the safety requirements for new drugs but also established for the first time an efficacy requirement for new drugs. In 1984, the United States Congress passed the Price Competition and Patent Term Restoration Act to provide increased patent protection to compensate for patent life lost during the approval process. Based on this act, the Food and Drug Administration (FDA) was authorized to approve generic drugs based on bioavailability and bioequivalence trials on healthy male subjects. The United States regulations for drug development are codified in the United States Code of Federal Regulations (CFR) which are to be carried out by the Food and Drug Administration.

Before the approval of a drug, the FDA requires that substantial evidence of the effectiveness and safety of the drug product be provided through the conduct of two well-controlled clinical studies. Before the drug can be tested in humans, however, it is also required that an appropriate analytical method be developed and a number of animal studies be conducted to assess the pharmacological characteristics, efficacy, and safety of the drug product in animals. To assist the sponsors in fulfilling the requirements, the FDA has issued a number of guidelines and guidances at various stages of drug development. The United States regulations codified in the CFR and FDA guidelines and/or guidances are then considered standards for good laboratory practice (GLP), good clinical (GCP), and good regulatory practice (GRP). After the drug is approved, the FDA also requires that the drug product be tested for its identity, strength, quality, purity, and stability before it can be released for use. For this purpose, the FDA also issued current good manufacturing practice (cGMP) guidelines to assure that the drug product possesses good drug characteristics such as proper identity, strength, quality, purity, and stability.

In drug development, almost all of the GLP, GCP, GRP, and cGMP guidelines require that:

1. Representative samples be drawn for testing, and
2. A valid design and appropriate statistical analysis be justified to ensure the accuracy and reliability of the test results.

As a result, the concept of good statistics practice (GSP) is the key to the success of GLP, GCP, GRP, and cGMP, and consequently, the success of the drug product. The objective of this paper is not only to introduce the concept of GSP but also to illustrate the importance and the implementation of GSP through some practical and/or regulatory issues that commonly occur during the development of a drug product.

In the next section, the United States regulatory requirements for drug development are briefly outlined. The role of statistics in the drug development and regulatory approval process is discussed in the section on "Good Statistics Practice." Also included in this section are the concept of GSP, some difficulties/concerns that commonly are encountered between statisticians and scientists, and suggestions for the implementation of GSP. In the section on "Practical Issues," the concept of GSP is illustrated by means of some practical and/or regulatory issues from the areas of nonclinical, preclinical, and clinical drug development. A brief conclusion is given in the last section.

REGULATORY REQUIREMENTS

For approval of drug products, the FDA has published regulations at various stages of drug development. These regulations include the investigational new drug application (IND) and new drug application (NDA) for new drugs, orphan drugs and over-the-counter human drugs and the abbreviated new drug application (ANDA) for generic drugs.

Table 1 summarizes CFR requirements at various stages of new drug and generic drug development. The cGMP as codified in 21 CFR 210 and 211 provides minimum requirements for the manufacture, processing, packing, and holding of a drug product. Its pur-

TABLE 1
The United States Regulations for Drug Development

CFR Number	Stage of Pharmaceutical Development
21 CFR 312	Investigational New Drug Application (IND)
21 CFR 314	New Drug Application (NDA)
21 CFR 314	Abbreviated New Drug Application (ANDA)
21 CFR 314	Supplements
21 CFR 10	Stability
21 CFR 50, 56, 312, 314	Good Clinical Practice (GCP)
21 CFR 210 & 211	Current Good Manufacturing Practice (cGMP)
21 CFR 320	Bioavailability/Bioequivalence

pose is to assure that the drug product meets the standard for the identity, strength, quality, and purity of the drug product. The standards are usually referred to as those specified in the United States Pharmacopeia and National Formulary (USP/NF) (1). The USP/NF consists of compendia of legally public standards for drug identity, strength, quality, purity, packaging, and labeling of drug products in the United States which are recognized as official compendia in the FD&C Act. The USP/NF describes not only the legally recognized standard testing methods and assay procedures for drug identity, strength, quality, and purity but also the number of samples to be tested and the acceptance criteria of various tests for release standards.

For new drug development, an IND permits the sponsors to gather the data on clinical safety and effectiveness that are needed for an NDA. An IND should contain information regarding chemistry, manufacturing, and controls (CMC) of the drug substance and drug product to ensure the identity, strength, quality, and purity of the investigational drug. In addition, the sponsors are required to provide adequate information about pharmacological studies for absorption, distribution, metabolism, and excretion (ADME) and acute, subacute, and chronic toxicological studies and reproductive tests in various animal species to support that the investigational drug is reasonably safe to be evaluated in clinical trials in humans. In addition, for approval of new drugs, the FDA requires at least two adequate well-controlled clinical studies be conducted in humans to demonstrate substantial evidence of the effectiveness and safety of the drug. As indicated in Section 314 of CFR Part 21, an NDA should include information on CMC, nonclinical pharmacology and toxicology, human pharmacology and bioavailability, clinical data, and statistics for review by various related divisions at the FDA before an approval decision can be reached.

Note that as indicated earlier, the development and approval of a new drug is a lengthy process. This lengthy process does not allow the access of promising drugs or therapies to patients with serious or life-threatening illness. Under the so-called *treatment IND* published in 1987, the FDA permits promising drugs or therapies currently under investigation to be available to patients with serious or life-threatening diseases. In 1992, the *Parallel Track Regulations* began allowing promising drugs or therapies for serious or life-threatening diseases to become available with considerably fewer data than required for approval. The FDA also published the *Accelerated Approval* based on surrogate endpoints to accelerate the approval process for promising drugs or therapies indicated for life-threatening diseases.

For generic drug development, the sponsors are required to include the information regarding product information (eg, potency and stability), pharmacokinetic data and analysis, statistical analysis, analytical methodology and validation, and clinical data in an ANDA submission for review. A complete ANDA is then subject to a request for plant inspection, chemistry/micro review, labeling

review, and bioequivalence review. The ANDA would be approved if positive results are obtained from the plant inspection and all reviews. Note that a new policy recently issued by the Office of Generic Drugs of the FDA indicates that for drug product manufactured at a foreign facility and shipped to the United States for packaging, release and stability testing can now be conducted at any facility, United States or foreign.

To assist sponsors in fulfilling regulatory requirements, the FDA has published a number of guidelines and guidances at various stages of drug development and regulatory approval. For example, to establish an appropriate expiration dating period and product storage requirements, the FDA published a stability guideline under 21 CFR 10.90 (2). This guideline provides a means of meeting regulatory requirements for an IND (21 CFR 312.23), an NDA (21 CFR 314.50), and an ANDA (21 CFR 314.55). For clinical development of new drugs, the FDA issued a guideline which described general considerations for clinical evaluation of drug products (3). In addition, for the approval of new drugs, the FDA also developed a guideline to assist sponsors in the preparation of submissions (4). This guideline describes an acceptable format for submission of clinical and statistical information on the drug product. For the approval of generic drugs, a guidance on statistical procedures for bioequivalence studies was issued by the Division of Bioequivalence, Office of Generic Drugs, FDA (5). This guidance suggests that a crossover model with log-transformed data be used to assess average bioequivalence for approval of generic drugs.

GOOD STATISTICS PRACTICE

The Role of Statistics in Drug Development

When a new drug is discovered, it is important to design an appropriate dosage form for the drug so that it can be delivered efficiently to the site of action for the optimal therapeutic effect. The FDA requires that analytical methods for the active ingredients of the drug be developed and validated so that the assay results are in compliance with some established specifications according to the USP/NF requirements. Statistics are necessarily applied in order to meet certain standards of accuracy and reliability. At the same time, stability studies are usually conducted to establish the expiration dating period of the drug product according to the design and analysis as specified in the FDA stability guideline. The FDA requires that statistical inference of the estimated expiration dating period be included in the NDA submission of the drug.

Before the drug can be approved, the FDA requires that substantial evidence of the effectiveness and safety of the drug be provided in the *Technical Section of Statistics* of an NDA submission. Since the validity of statistical inference regarding the effectiveness and safety of the drug is always a concern, it is then suggested that a careful review be performed to ensure an accurate and reliable assessment of the drug product. In addition, to have a fair assessment, the FDA also established advisory committees, each consisting of clinical, pharmacological, and statistical experts and one advocate (not employed by the FDA) in designated drug classes and subspecialities, to provide a second but independent review of the submission. The responsibility of the statistical expert is not only to ensure that a valid design is used but also to evaluate whether statistical methods used are appropriate for addressing the scientific and medical questions regarding the effectiveness and safety of the drug.

After the drug is approved, the FDA also requires that the drug product be tested for its identity, strength, quality, purity, and stability before it can be released for use. For this purpose, the cGMP is necessarily implemented to:

1. Validate the manufacturing process,
2. Monitor the performance of the manufacturing process, and

3. Provide quality assurance of the final product.

At each stage of the manufacturing process, the USP/NF requires that sampling plans, acceptance criteria, and valid statistical analyses be performed for the intended tests such as potency, content uniformity, and dissolution. For each test, sampling plan, acceptance criteria, and valid statistical analysis are crucial for determining whether the drug product can pass the test based on the results from a representative sample.

Good Statistics Practice

As discussed above, statistics plays an important role during the development and regulatory approval process of new drug and generic drug products. A valid design and an appropriate use of statistical methods provide an accurate and reliable assessment of the drug under investigation. Therefore, good statistics practice is necessarily applied in the drug development and regulatory approval process. GSP is defined as a set of principles which assures the validity of the design and analysis of the intended studies conducted at various stages of the development and regulatory approval process of a drug product. A valid design is one which can reflect the study objectives and address scientific or medical questions regarding the drug product under investigation. Under the valid design, a representative sample is then selected from the target population at random for study. Statistical inference drawn based on appropriate statistical methods under a valid design can then provide a fair assessment with the desired accuracy, precision, and reliability. As a result, GSP is the key to the success of the development and regulatory approval process of a drug product.

The implementation of GSP, however, is a team project which requires mutual communication, confidence, respect, and cooperation among statisticians, pharmaceutical scientists in related areas, and regulatory agents. The implementation of GSP involves some key factors which have an impact on the success of GSP. These factors include:

1. Regulatory requirements for statistics,
2. The dissemination of the concept of statistics,
3. An appropriate use of statistics,
4. Communication and flexibility, and
5. Statistical training.

These factors are briefly described below.

In the drug development and approval process, regulatory requirements for statistics are the key to the implementation of GSP. They not only enforce the use of statistics but also establish standards for statistical evaluation of the drug product under investigation. Statistical evaluation provides useful information for pharmaceutical scientists and regulatory agents for determining whether the drug product under investigation has the claimed effectiveness and safety for the intended diseases, and whether the drug product possesses good drug characteristics such as the proper identity, strength, quality, purity, and stability.

In addition to regulatory requirements, it is always helpful to disseminate the concept of statistics whenever possible. GSP not only fulfills regulatory requirements but it also improves the accuracy and precision of the evaluation of the drug under investigation and the quality of the drug product manufactured after approval. It is then important for pharmaceutical scientists to understand the concept of statistics during the process of drug development and regulatory approval. It is also critical for pharmaceutical scientists and regulatory agents to recognize that:

1. A valid statistical inference is necessary to provide a fair assessment with certain assurance regarding the uncertainty of the drug under investigation,
2. A larger sample size is often required to increase such assurance, and
3. An invalid design and analysis may result in a misleading conclusion regarding the drug under investigation.

A commonly encountered problem in the drug development and regulatory approval process is the misuse of design and statistical methods in some studies. The misuse of design and statistical methods may result in either having the right question with the wrong answer, or having the right answer for the wrong question. As a result, it is not clear what the question is and what the answer is because the design may not be able to address the question or statistical analysis cannot reflect the design and hence cannot address the question. As a result, GSP recommends that statistical methods should be chosen to reflect the design which should be able to address the scientific or medical questions regarding the intended study objectives.

Communication and flexibility are important factors in the success of GSP. Inefficient communication between statisticians and pharmaceutical scientists or regulatory agents may result in a misunderstanding of the intended study objectives and consequently an invalid design and/or inappropriate statistical methods. Thus, efficient communications among statisticians, pharmaceutical scientists, and regulatory agents is essential for GSP. In addition, in many studies, the assumption of a statistical design or model may not be met due to the nature of the drug under study, experimental environment, and/or other causes related/unrelated to the studies. In this case, the traditional approach of doing everything by the book does not help. A concern from a pharmaceutical scientist or the regulatory agent may translate into a constraint for statistical design and analysis. In this case, GSP suggests that a successful statistician be flexible in the sense that a valid statistical design and analysis should be developed under the constraints.

Since regulatory requirements for the drug development and approval process may vary from drug to drug, various designs and/or statistical methods may be required to provide a valid assessment of a drug product. Therefore, GSP suggests that statistical training programs be routinely held for both statisticians and nonstatisticians, including pharmaceutical scientists and regulatory agents. The purpose of such training program is threefold:

1. It enhances communications within the statistical community. Statisticians can certainly benefit from such training programs by acquiring more practical experience from each other,
2. It provides the opportunity to exchange ideas and/or concepts regarding regulatory requirements between professional societies, and
3. Most importantly, it identifies critical practical and/or regulatory issues that are commonly encountered in the drug development and regulatory approval process. A panel discussion from different disciplines may result in some consensus to resolve the issues.

PRACTICAL ISSUES

This section illustrates the concept of GSP by means of some practical and/or regulatory issues commonly encountered during new drug and generic drug development. For illustration purposes, drug development is classified into the areas of nonclinical, preclinical, and clinical.

Nonclinical Applications

In the pharmaceutical industry, a scale-up experiment is usually employed to ensure that the results from a laboratory batch or small-scale production batch can be predictive of a regular scale production batch. The purpose of the scale-up experiment is not only to identify, evaluate, and optimize critical formulation and/or (manufacturing) process factors of the drug product, but also to maximize or minimize excipient ranges. If there are a large number of formulation and/or process factors, then the scale-up program could be very substantial. For example, suppose a pharmaceutical company is interested in studying the effects of eight process and ingredient factors on various properties of tablets of a drug product. If each factor has two levels, a full 2^8 factorial design would

require a total of 256 runs which it is almost impossible for the pharmaceutical company to conduct. As an alternative, a fractional factorial design is suggested. Suppose that it is believed that only three or four factors will have a major impact on the properties of the tablets of the drug product and that there are no interactions involving four or more factors based on prior information. Then, a 2^{8-4} fractional factorial design is useful. This design reduces the total of 256 runs to only 16 runs.

For another example, consider the issue of stability analysis for establishment of an expiration dating period of a drug. In 1993, the International Conference on Harmonization (ICH) issued guidelines on stability based on a strong industrial interest in international harmonization of requirements for marketing in the EC, Japan, and the United States (6). The ICH guideline is similar to the current FDA stability guideline. For example, the ICH guideline suggests that testing under the defined long-term conditions be normally done every three months over the first year, every six months over the second year, and then annually. It also requires that stability information from accelerated and long-term testing be provided on at least three batches. As a result, if there are three different strengths and three package types of a drug product, a total of 243 assays are required based on a full factorial design over four years. In practice, it is almost impossible to conduct a stability study of this size due to the limited resources available and budget constraints.

Alternatively, as suggested by the ICH guideline, a matrixing design or bracketing design may be used. A matrixing design is any subset of a full factorial design (7,8). A bracketing design is the design of a stability schedule such that at any time point only the samples on the extremes of container size and/or dosage strengths are tested (6). Nordbrock (9) examined a number of commonly used stability designs based on fractional factorial designs with partial sampling time points. Some of these designs are, in fact, special cases of matrixing designs. The use of a matrixing design or a bracketing design with partial sampling time points has become very popular because it reduces the number of assays tremendously. For example, if one only performs stability testing on two strengths per batch, the number of assays will be reduced by 33.3% (from 243 to 162).

The above two examples demonstrate that an appropriate use of statistics is not only very cost effective but can only assist in achieving the goal of the desired accuracy and reliability within the timeframe. It should be noted, however, that the choice of an appropriate fractional factorial design for a scale-up program or a reduced design such as a matrixing design or a bracketing design for stability studies should be statistically justified so that the results will have certain accuracy, precision, and reliability without losing much statistical power.

Preclinical Applications

In recent years, as more generic drug products became available, whether the quality and the therapeutic effect of the generic drugs were comparable to the innovator drug became a concern. Although the FDA indicates that an approved generic drug can be used as a substitute for the innovator drug, it is recognized that the current FDA regulation on average bioequivalence does not guarantee drug interchangeability between generic drugs and the innovator drug (10).

Drug interchangeability is usually classified as drug switchability and drug prescribability. Drug prescribability is usually referred to as the physician's choice for prescribing an appropriate drug product for his/her new patients between an innovator drug and a number of generic drugs which have been shown to be bioequivalent to the innovator drug. The underlying assumption of prescribability is that the innovator drug and its generic drugs can be used interchangeably in terms of the efficacy and safety of these drug products.

Under current practice, the FDA only requires that evidence of bioequivalence in average bioequivalence of the generic drugs to the same innovator drug be provided. Bioe-

quivalence among generic drugs is not required. In practice, it is very likely that two generic drugs may not be bioequivalent when both drugs are bioequivalent to the same innovator drug. To overcome this disadvantage, it is suggested that population bioequivalence be assessed in addition to average bioequivalence. To ensure drug prescribability, Chow and Liu (11) proposed performing a meta analysis for bioequivalence review based on data from all submissions. The proposed method provides a useful tool for monitoring the performance of the generic drugs approved by the FDA based on current regulation on average bioequivalence.

Drug switchability is related to the switch from a drug (eg, an innovator) to an alternative drug (eg, a generic drug) within the same subject whose concentration of the drug has been titrated to a steady, efficacious, and safe level. For drug switchability, the FDA recommends that the assessment of individual bioequivalence be considered (12). Based on the concept of individual bioequivalence, several methods have been proposed. None of these methods, however, seems to serve the purpose of ensuring drug switchability due to the discrepancy from statistical and regulatory point of views in terms of the concept, definition, and criteria of individual bioequivalence (13,14,15).

Note that bioequivalence testing as a surrogate for clinical evaluation is based on the fundamental assumption that if the two drug products are bioequivalent (in terms of drug absorption), then they are therapeutically equivalent. Therefore, it is very important at least to assure that the two drugs have similar drug absorption profiles and consequently have a similar therapeutic effect. Thus, alternative regulatory criteria for bioequivalence should be developed in such a way that they can address both drug prescribability and drug switchability. This task, however, provides a challenge not only to regulatory agents but also to biostatisticians.

Clinical Applications

In clinical trials, it is common to observe an inconsistent result between a significant statistical difference and a clinically meaningful difference in the assessment of the effectiveness and safety of a drug product. This inconsistent result often creates confusion and/or an argument among clinicians and statisticians. A significant difference may be referred to as a statistically significant difference or a clinically significant difference. A difference which is unlikely to occur by chance alone is considered a statistically significant difference. A statistically significant difference may be relatively small as compared to the treatment mean. A large difference may not be of statistical significance if the sample size is too small.

A clinically significant difference is a difference that is considered clinically meaningful and important to the investigators. Basically, there are four different outcomes for an observed difference in a clinical trial. The observed difference may show that:

1. The difference is both statistically and clinically significant,
2. There is a statistically significant difference yet the difference is not clinically significant,
3. The difference is of clinical significance yet it is not statistically significant, and
4. The difference is neither statistically significant nor clinically significant.

If the difference is both clinically and statistically significant or it is neither clinically nor statistically significant, then there is no confusion and a consistent conclusion can be drawn. In many cases, however, a statistically significant difference does not agree with the clinically significant difference. For example, a statistical test may reveal that there is a statistically significant difference. The difference is too small (which may be due to a unusual small variability or a relatively large sample size), however, to be of any clinical importance and hence it is not clinically significant. In this case, a small p-value may lead one to conclude the effectiveness of the treatment. On the other hand, the result may indicate that there is a clinically significant difference but the sample size

may be too small (or variability is too large) to claim a statistically significant difference. In this case, the evidence of effectiveness is not substantial due to a large p-value.

For another example, consider the use of active control trials in clinical development. For approval of a new drug, conducting a clinical trial comparing the new drug with a control is required. Section 314.126 in Part 21 of the CFR indicates that a control could be a placebo control, no treatment control, positive control, or historical control. In practice, since it may not be ethical to conduct a placebo control study with very ill patients with severe or life-threatening diseases to establish efficacy of a new drug, a positive control trials is often considered as an alternative to evaluate the effectiveness and safety of the new drug by comparing it with a positive control which has been shown to be effective and safe for the intended disease. A positive control trial is also known as an active control trial.

As indicated by Pleager and Hall (16), the primary objective of an active control trial could be to:

1. Establish the efficacy of the test drug,
2. Show that the test drug is equivalent to an active control, or
3. Demonstrate that the test drug is superior to the active control.

The equivalence and superiority to the active control, however, do not guarantee that the test drug is effective. As an example, consider the case where a test drug (denoted by T) is superior to an active control (denoted by A), denoted by T > A (ie, a statistically or clinically significant difference between T and A is observed). In this case, it is possible that the actual outcome is T > A > P (ie, both T and A are effective), T > P > A (ie, only T is effective), or P > T > A (ie, both A and B are ineffective), where P denotes the placebo. For another example, when T and A are equivalent (denoted by TA), it may fall in one of the following possible outcomes: TA > P (ie, both T and A are equally effective), TAP, T > AP (ie, there is no significant decreasing improvement among T, A, and P; however, T is superior to P and hence is effective), or P > TA (ie, both T and A are ineffective).

As a result, showing equivalence or superiority may imply that T and A are both equally effective or equally ineffective. This statement is especially true for some drug products such as antianxiety agents, antidepressants, antianginal agents, or appetite suppressants which may not necessarily beat the placebo. For these drug products, the FDA prefers placebo control trials, while the EC resists placebo control trials. It is then suggested that trials including the test drug, an active control, and a placebo be used with a different ratio of patients assigned whenever possible.

The above two examples indicate that the communication of statistical concept between clinicians and statisticians is essential for the implementation of GCP in clinical development for establishment of drug efficacy and safety.

CONCLUSIONS

During the development and regulatory approval process, GLP, GCP, GRP, and cGMP are necessarily applied to:

1. Assure the effectiveness and safety of the drug under investigation before approval, and
2. Ensure that the drug product possesses good drug characteristics such as proper identity, strength, quality, purity, and stability in compliance with the standards as specified in the USP/NF after regulatory approval.

In essence, GSP is the foundation of GLP, GCP, GRP, and cGMP. The implementation of GSP is a team project which involves statisticians, pharmaceutical scientists, and regulatory agents as well. The success of GSP depends upon mutual communication, confidence, respect, and cooperation among statisticians, pharmaceutical scientists, and regulatory agents.

REFERENCES

1. USP/NF. The United States Pharmacopeia XXIII and the National Formulary XVIII. Rockville, MD: United States Pharmacopeidal Convention, Inc.; 1995.
2. FDA. Guideline for Submitting Documentation for the Stability of Human Drugs and Biologics. Rockville, Maryland: Center for Drugs and Biologics, Food and Drug Administration; 1987.
3. FDA. Guideline for General Considerations for Clinical Evaluation of Drugs. Rockville, Maryland: Center for Drug Evaluation and Research, Food and Drug Administration; 1977.
4. FDA. Guideline for the Format and Content of the Clinical and Statistical Sections of an Application. Rockville, Maryland: Center for Drug Evaluation and Research, Food and Drug Administration; 1988.
5. FDA. Guidance on Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design. Rockville, Maryland: Division of Bioequivalence, Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration; 1992.
6. ICH. Stability Testing of New Drug Substances and Products. Tripartite International Conference on Harmonization Guideline. International Conference on Harmonization; 1993.
7. Chow SC. Statistical design and analysis of stability studies. Presented at the 48th Conference on Applied Statistics, Atlantic City, New Jersey: 1992.
8. Chow SC, Liu JP. *Statistical Design and Analysis in Pharmaceutical Science*. New York, New York: Marcel Dekker, Inc.; 1995.
9. Nordbrock E. Statistical comparison of stability study designs. *J Biopharm Stat*. 1992;2:91–113.
10. Chow SC, Liu JP. Current issues in bioequivalence trials. *Drug Inf J*. 1995;29:795–804.
11. Chow SC, Liu JP. Meta-analysis for bioequivalence review. *J Biopharm Stat*. 1997;7(1):97–111.
12. Chen ML. Individual bioequivalence—a regulatory update. *J Biopharm Stat*. 1997;7(1):5–11.
13. Wellek S. A comment on so-called individual criteria of bioequivalence. *J Biopharm Stat*. 1997;7:17–21.
14. Endrenyi L. Some issues for the consideration of individual bioequivalence. *J Biopharm Stat*. 1997; 7:35–39.
15. Liu JP, Chow SC. Some thoughts on individual bioequivalence. *J Biopharm Stat*. 1997;7:41–48.
16. Pledger GW, Hall D. Active control trials: do they address the efficacy issue? Proceedings of the Biopharmaceutical Section of the American Statistical Association; 1986:1–7.