Journal of Food and Drug Analysis
1995. 3(2) : 75-85

藥物食品分析
第三卷　第二期

Current Issues in Regulatory Requirements of Drug Stability

SHEIN-CHUNG CHOW AND ANNPEY PONG

Bristol-Myers Squibb Co. and Roberts Pharm. Corp., U.S.A.

ABSTRACT

For every commercial drug product, it is essential to indicate the expiration date to provide the consumer with some assurance that the drug product will retain its identity, strength, quality, and purity throughout the expiry period. For the determination of an expiration dating period, it is necessary to conduct stability studies to assess the intrinsic degradation of the drug product. In this paper, we provide an overview of regulatory requirements for stability including stability guidelines issued by the United States Food and Drug Administration and those established by the International Conference on Harmonization. In addition, we address current statistical design and analysis issues that often occur during the conduct of stability studies.

Key words : Expiration dating period, shelf-life, sampling times, batch similarity, matrixing design, bracketing design, international harmonization.

INTRODUCTION

For every drug product found on the market, government agencies require that an expiration dating period must be indicated on the immediate container label. The expiration date provides the consumer with a high confidence that the drug product will retain its identity, strength, quality, and purity throughout the expiry period. To provide such assurance, the sponsors (e.g., producers or manufacturers) find it necessary to conduct stability studies to collect, analyze, and interpret degradation data on their drug products throughout the expiry period. The purpose of a stability study is not only to assess the degradation of a drug product but also to establish an expiration dating period (shelf-life) applicable to all future batches of the same drug product.

In the early 1970's, although some drug products such as penicillin were known to be unstable, there were no regulatory requirements regarding drug stability. Since then, however, it has become a common concern that drug products may not be able to maintain their identity, strength, quality, and purity after stored over a period of time. This is especially in cases where the drug product is expected to degrade over time. To assure the integrity of commercial drugs, the United States Pharmacopeia (USP) included the first clause regarding drug expiration dating period in 1975. In 1984, the United States Food and Drug Administration (FDA)

Correspondence to: Shein-Chung Chow

Accepted for Publication: Mar. 31, 1995.

first required stability testing. However, specific requirements on statistical design and analysis of stability studies for human drugs and biologics were not available until the current FDA guideline issued in 1987. 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 European Community (EC), Japan, and the United States of America (USA)[1]. At the same time, the World Health Organization (WHO) also published draft guidelines for stability[2].

The purpose of this paper is not only to provide an overview of the requlatory requirements for stability, but also to discuss current statistical issues on design and analysis that often occur during the conduct of stability studies. In the next section, regulatory requirements for stability from the FDA and the ICH are outlined. Statistical considerations are described in Section 3. A brief discussion is given in Section 4.

REGULATORY REQUIREMENTS

I. *FDA Guideline*

In 1987, under 21 Code of Federal Regulations (CFR) 10.9, the FDA issued guidelines for the stability of human drugs and biologics[3]. The purpose of the guidelines is not only to provide recommendations for the design and analysis of stability studies for establishing an appropriate expiration dating period and product requirements, but also to provide recommendations for submission of stability information and data to the FDA for investigational drug (IND) application and new drug application (NDA) and biological product license application (PLA).

The FDA guideline indicates that a stability protocol must describe not only how the stability study is to be designed and carried out, but also the statistical methods to be used for the analysis of the data. The design of a stability study is intended to establish an expiration dating period applicable to all future batches of the drug product manufactured under similar circumstances. An appropriate design should be able to take into account the variabilities due to individual dosage units, containers within a batch, and batches. The purpose is to ensure that the resulting data for each batch are truly representative of the batch as a whole and to quantify the variability from batch to batch.

For determination of an expiration dating period, if the drug characteristic is expected to decrease (increase) with time, and there is no evidence of batch-to-batch variability, the FDA guideline suggests that the expiration dating period be estimated (determined) as the time period at which the 95% one-sided lower (upper) confidence bound for the mean degradation curve intersects the approved lower (upper) specification limit (see Figure 1). To ensure the accuracy and reliability of the estimated shelf-life, the FDA guideline also provides a number of requirements for conducting a stability study which are summarized below:

(I). *Batch Sampling Consideration*

The FDA guideline indicates that at least three batches and preferably more should be tested to allow for some estimate of batch-to-batch variability and to test the hypothesis that a si-

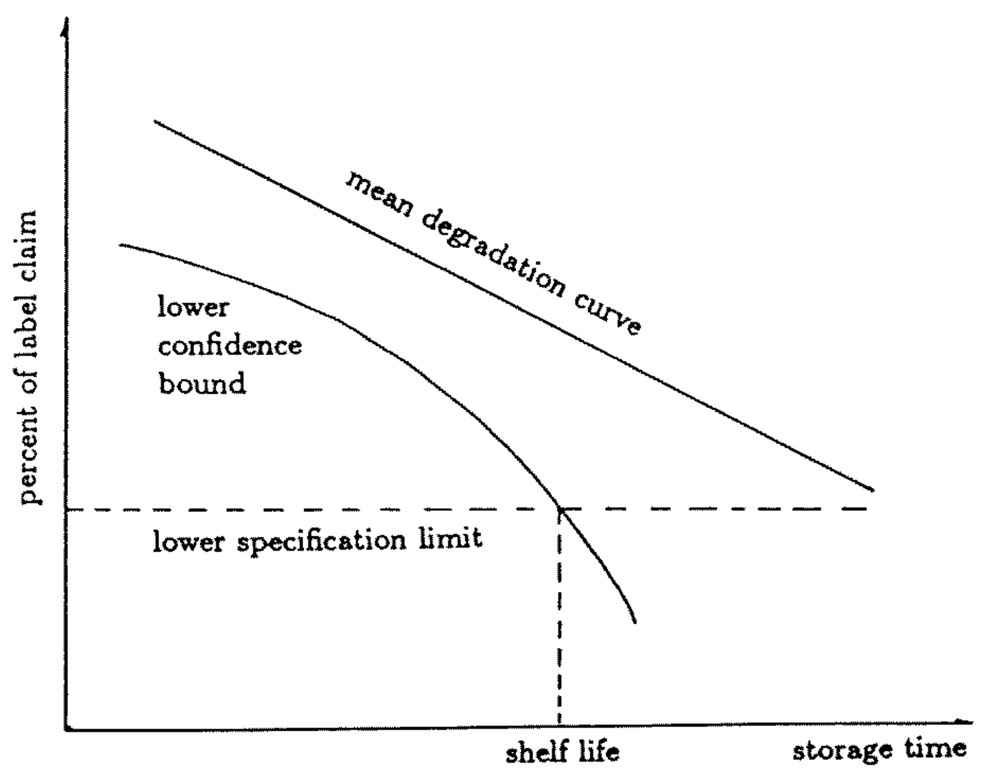

Figure 1. Determination of drug shelf-life

ngle expiration dating period for all batches is justifiable. This arises from concern that testing of a single batch may not permit assessment of batch-to-batch variability and that testing of two batches may not provide a reliable estimate. Therefore, the specification of at least three batches being tested has become the minimum requirement. In general, more precise estimates can be obtained from more batches.

(II). *Container-closure and Drug Product Sampling*

To ensure that the samples chosen for stability study can represent the batch as a whole, the FDA guideline suggests that selection of containers such as bottles, packages, and vials from the batches be included in the stability study. Therefore, it is recommended that at least as many containers be sampled as the number of sampling times in a stability study. In any case, sampling of at least two containers for each sampling time is recommended.

(III). *Sampling Time Considerations*

The FDA guideline suggests that stability testing be done at three month intervals during the first year, six month intervals during the second, and yearly thereafter. In other words, it is suggested that stability testing be performed at 0, 3, 6, 9, 12, 18, 24, 36, and 48 months for a four-year duration of a stability study. However, if the drug product is expected to degrade rapidly, then more frequent sampling is necessary.

II. *ICH Guideline*

In the interest of having an international harmonization of stability testing requirements for a registration application within the three areas of the EC, Japan, and the USA, a tripartite guideline for the stability testing of new drug substances and products was developed by the Expert Working Group (EWG) of the ICH [(1)]. The ICH guideline provides a general indication on the requirements for stability testing, but leaves sufficient flexibility to encompass the variety of different practical situations required for specific scientific situations and characteristics of the materials being evaluated. The ICH guideline establishes the principles that information on stability generated in any one of the three areas of the EC, Japan and the USA would be mutually acceptable in both of the other two areas provided that it meets the appropriate requirements of this guideline and the labeling is in accordance with national and regional requirements. It should be noted that the choice of test conditions defined in the ICH guideline is based on an analysis of the effects of climatic conditions in the three areas of the EC, Japan and the USA. Therefore, the mean kinetic temperature in any region of the world can be derived from climatic data.

Basically, the ICH guideline is similar to the current FDA 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 requires that the container to be used in the long-term, real time stability evaluation should be the same as or simulate the actural packaging used for storage and distribution. For the selection of batches, it requires that stability information from accelerated and long-term testing be provided on at least three batches and the long-term testing should cover a minimum of 12 months duration on at least three batches at the time of submission. For drug product, it is required that the three batches be of the same formulation and dosage form in the containers and closure proposed for marketing. Two of the three batches should be at least pilot scale. The third batch may be smaller (e.g., 25,000 to 50,000 tablets or capsules for solid oral doage forms). However, the ICH guideline also requires that the first three production batches of drug substance or drug product manufactured post approval, if not submitted in the original registration application, be placed on long-term

stability studies using the same stability protocol as in the appoved drug application. For storage conditions, the ICH guideline requires that accelerated testing be carried out at a temperature at least 15°C above the designated long-term storage temperature in conjunction with the appropriate relative humidity conditions for that temperature. The designated long-term testing conditions will be reflected in the labelling and re-test date. The re-test date is referred to the date when samples of the drug substance should be re-examined to ensure that the material is still suitable for use. The ICH guideline also indicates that where significant change occurs during six months storage under conditions of accelerated testing at 40 °C±2 °C/ 75%±5% relative humidity, additional testing at an intermediate condition (such as 30 °C±2 °C/ 60%±5% relative humidity should be conducted for drug substances to be used in the manufacture of dosage forms tested long-term at 25 °C/ 60% relative humidity. Note that a significant change at 40 °C/ 75% relative humidity or 30 °C/ 60% relative humidity is considered failure to meet the specification.

For the evaluation of stability data, the ICH guideline indicates that statistical methods should be employed to. test the goodness of fit of the data on all batches and combined batches (where appropriate) to the assumed degradation line or curve. If it is inappropriate to combine data from several batches, the overall re-test period may depend on the minimum time a batch may be expected to remain within acceptable and justified limits. A re-test period is defined as the period of time during which the drug substance or drug product can be considered to remain within the specification and therefore acceptable for use in the manufacture of a given drug product provided that it has been stored under the defined conditions.

STATISTICAL CONSIDERATIONS

For a single batch, as mentioned earlier, the shelf-life of drug characteristics (e.g., potency) can be determined as the time at which the 95% one-sided lower confidence limit (or 95% lower confidence bound) for the mean degradation curve intersects the acceptable lower specification limit assuming that the drug characteristics decrease with time. The degradation of the strength for a drug product over time can be described by the following simple linear regression model:

$$Yj=\alpha+\beta Xj+e_j,\ j=1,\cdots,\ n, \tag{3.1}$$

where Yj is the assay result (percent of label claim) at sampling time Xj, α is the batch effect and β is the degradation rate of that batch, and e_j are identically and independently distributed normal random variables with mean 0 and variance σ^2. The least squares estimates for α, β, and σ^2 are given by

$$b=\frac{S_{XY}}{S_{XX}},$$

$$a=\bar{Y}-b\bar{X}$$

$$s^2=\frac{S_{YY}-bS_{XY}}{n-2},$$

where

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\bar{Y}=\frac{1}{n}\sum_{i=1}^{n} Y_i,$$

$$S_{XX}=\sum_{i=1}^{n}(X_i-\bar{X})^2,$$

$$S_{XY}=\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}),$$

$$S_{YY}=\sum_{i=1}^{n}(Y_i-\bar{Y})^2,$$

At a particular time point, t=x, the least squares estimate of mean degradation (i.e., $E(Y(x))=\alpha+\beta x$) is then given by

$$y(x)=a+bx,$$

with variance

$$Var(y(x))=\sigma^2\left[\frac{1}{n}+\frac{(x-\bar{X})^2}{S_{XX}}\right]$$

Thus, the least squares estimate of Var(y(x)) can be obtained as

$$\hat{Var}(y(x))=s^2\left[\frac{1}{n}+\frac{(x-\bar{X})^2}{S_{xx}}\right]$$

Consequently, the 95% lower confidence limit of the mean degradation line is given by

$$L(x)=a+bx-t(0.05,\ n\text{-}2)\ SE(x),$$

where

$$SE(x)=\left\{ s^2\left[\frac{1}{n}+\frac{(x-\bar{X})^2}{S_{xx}}\right]\right\}^{1/2}, \qquad (3.2)$$

and t(0.05, n-2) is the 5% upper quantile of a central t distribution with n-2 degrees of freedom. Thus, the points where (3.2) intersects the approved lower specification limit η, if exist, are the two roots, denoted by x_L and x_U, of the following quadratic equation

$$[\eta-(a+bx)]^2=$$

$$t^2(0.05,n-2)s^2\left[\frac{1}{n}+\frac{(x-\bar{X})^2}{S_{xx}}\right]$$

Kohberger (1988) indicated that if the slope is statistically significantly smaller than 0 and the intercept is statistically significantly larger than η, then $t=x_L$ is the estimated shelf-life[4].

Easterling (1969) also provided an interpretation based on hypothesis testing for estimation of an expiration dating period using the degradation curve rather than the mean degradation curve[5]. Carstensen and Nelson (1976) proposed that the shelf-life be determined based on the so-called prediction limit which is considered equally acceptable to the FDA[6]. More details on the estimation of an expiration dating period based on (3.1) for long-term stability studies can be found in Chow and Liu (1995)[7].

In this section, we will provide some insight to some statistical issues that often occurs in the design and analysis of stability studies.

I. *Design Issues*

Since the primary objective of a stability study is to characterize the degradation of a drug product and consequently to establish an expiration dating period for the drug product, a stability design should be chosen to achieve these objectives. First, an appropriate stability design should be able to adequately characterize the degradation of the drug product. For this purpose, the selection of sampling time intervals is important. If the drug product is expected to degrade linearly over time, then the sampling time intervals suggested in the FDA guideline should be used. However, if the drug product is expected to degrade in a quadratic fashion, then it is suggested that more sampling be taken at the time at which the curvature is expected to occur. Furthermore, since it is of interest to establish a single shelf-life for the drug product regardless the strength and package type, an appropriate design should be chosen to avoid any possible confounding effect and interaction effect. If there is an interaction between strength and package type, then the stability of the drug product should be evaluated at each combination of strength and package type. A significant interaction between strength and package type indicates that stability loss between strengths are not consistent across different package types. A typical stability design to achieve these objectives is a full factorial design.

However, in many cases, most pharmaceutical companies are unable to conduct a full factorial design with sampling time intervals suggested by the FDA due to limited resources. To achieve certain degree of precision without losing much information, a reduced stability design such as a matrixing or bracketing design with less sampling time points is usually considered as an alternative[1,7,9]. Nordbrock (1992) compared a number of commonly used stability designs and proposed a criterion for design selection based on their statistical powers for detection of a significant difference in slope[8]. Ju and Chow (1995), however, indicated that the criterion for selection of an appropriate design should be based on the precision of the estimated drug shelf-life obtained under the design [9].

II. *Batch Similarity*

As indicated in the FDA guideline, a minimum requirement for a stability study is to test at least three batches. If batch-to-batch variability is small, it would be advantageous to combine the data into one overall estimate with high precision and a large degree of freedom for mean squared error. As indicated in the FDA guideline, combining the data should be supported by preliminary testing of batch similarity. The similarity of the degradation curves for each batch tested should be assessed by applying statistical tests of the *equality* of slopes and of zero time intercepts. Chow and Shao (1989) proposed several tests for batch-to-batch variability under normality assumption[10]. Note that the test suggested in the FDA guideline is the one for detecting the difference of slopes and intercepts among batches to be performed at the significance level of $\alpha=0.25$[11]. If tests for equality of slopes and for equality of intercepts do not result in rejection at the 25% level of significance, the data from the batches would be pooled. However, if tests resulted in p-values less than 0.25, a judgment would be made by the FDA reviewers as to whether pooling would be permitted.

It, however, should be noted that there are some criticisms regarding the use of the significance level of 0.25. Among these criticisms, the following are probably the most common:

(I) Acceptance or rejection of the null hypothesis of no difference in slopes among batches does not guarantee that the batches have similar degradation rates. This is because the problem of similarity is incorrectly formulated by the wrong hypothesis of difference;

(II) It is not a common practice to increase test power by increasing the level of significance. As an alternative, the concept of the interval hypotheses for bioequivalence problem can be applied to test batch similarity[12,13]. Under the assumption that the batch effect is fixed, the concept of the interval hypotheses is to claim similarity if the intercepts and slopes of the degradation curves for each batch are within an acceptable *equivalent* limits. If tests for equivalence of slopes and for equivalence of intercepts do not result in rejection at the 5% level of significance, the data from the batches would be pooled for determination of drug shelf-life.

The FDA guideline indicates that the design of a stability study is intended to establish, based on testing a limited number (at least three) batches of a drug product, an expiration dating period applicable to all *future* batches of the drug product manufactured under similar circumstances. In practice, there may be a slight or moderate variation among different batches and therefore the degradation curves have different intercepts and/or slopes from batch to batch. According to the FDA guideline, if there is a batch-to-batch variation (i.e., we reject the null hypothesis of batch similarity), the data cannot be pooled for determination of an expiration dating period. A method suggested in the guideline is to consider the minimum of individual shelf-lives where each shelf-life is obtained by fitting an ordinary linear regression within a batch. The minimum of individual shelf-lives is then used to reflect the shelf-life of all future batches of the drug product. The minimum approach, however, appears to be conservative and lacks statistical justification[14]. Note that if batch is assumed random, the shelf-life can be estimated without the preliminary test for batch similarity. However, estimation of drug shelf-life under the assumption of random batches requires more than three batches which are specified in the FDA guideline.

Recently, several methods for combining information from different batches have been proposed. Under assumption of the fixed batch effect, Ruberg and Hsu (1992) proposed the use of multiple comparison technique for pooling with the worst batch[15]. Several estimation procedures for shelf-life under assumption of a random batch effect have also been proposed[14,16,17]. Ho, Liu and Chow (1992) conducted a Monte Carlo simulation study to compare various methods including the approach proposed by the FDA under either fixed or random effects [18]. Morris (1992) also examined the con-

sequences in estimation of drug shelf-life when an incorrect linear model is used for the true exponential model[19].

III. *Sampling Time Considerations*

The FDA guideline indicates that sampling times should be chosen so that any degradation can be adequately characterized (i.e., at a sufficient frequency to determine with reasonable assurance the nature of the degradation curve). Usually, the relationship can adequately be represented by a linear, quadratic, or cubic function on an arithmtic or a logarithmic scale of the percent of label claim. Statistically, it is suggested that more frequent sampling be taken at which a curvature of the degradation curve is expected to occur in order to adequately characterize the degradation of the drug product. In addition, the FDA guideline also encourages to test an increased number of replicaties at the later sampling times, particularly the latest sampling time because this will increase the average sampling time toward the desired expiration dating period.

Assuming that the drug characteristic is expected to decrease with time, for long-term stability studies under ambient conditions such as NDA stability studies, the FDA guideline suggests that stability testing be done at 3-month intervals during the first year, 6-month intervals during the second year, and yearly thereafter. However, for drug products predicted to degrade rapidly, more frequent sampling is necessary. For marketing stability studies, less frequent sampling is usually considered on more batches.The purpose of a stability study is to characterize the degradation of the drug product and consequently to establish drug shelf-life. For this purpose, the following are a list of statistical issues which are of concern:

(I) How to select the number of time points and the allocation of the selected time points in a manner such that the degradation of ingredients of a drug product is adequately characterized?

(II) How frequent sampling is necessary in order to have a desired degree of accuracy and precision for the estimated shelf-life?

(III) Is it reliable to predict durg shelf-life beyond the time interval under study?

(IV) Is it necessary to have replicates at each sampling time point?

(V) How to efficiently allocate the number of assays at each sampling time point if a fixed number of assays are to be done?

Mathematically, if the degradation curve is linear, it can be uniquely determined by two time points. One may consider to have stability testing at the initial (i.e., the time at which the batch is manufactured) and at the latest sampling time point. It, however, should be noted that, statistically, there is no degrees of freedom for the error term if only two sampling time points are considered. The pharmaceutical companies are usually interested in acquiring stability information regarding the drug product within a short period of time after the drug product is manufactured. However, if only two time points are employed in a long-term stability study, no information about degradation can be obtained between these two time points. In addition, if the latest sampling time point is too close to the initial sampling time point, the fitted degradation line may not be reliable for establishing an expiration dating period beyond the time interval under study.

It is therefore of interest to study the impact of the frequency of sampling on the characterization of the degradation curve and the determination of drug shelf-life. Moreover, the 95% confidence interval for the mean degradation at time point such as the initial and last time points, which is further away from the middle of the range of time points, could be very wide. Consequently, the estimated shelf-life may not be reliable. The study of the reliability of an estimated drug shelf-life beyond the time interval under study is then an interesting and important topic in stability analysis.

In the current guideline, the FDA does not require that stability testing be repeatedly done

at each time point. Replicates at each time point certainly enable us to estimate the degradation curve more precisely in terms of the width of the confidence intervals about the estimated curve around the average of the sampling times included in the study. As a result, replicates improve the accuracy and precision of the estimated shelf-life. In addition, the information of replicates can be used to perform goodness-of-fit test for the fitted degradation curve. The above discussion may provide some justification as to why it is required that each individual test be repeated three times in Japan. When the pharmacuetical company can only perform a fixed number of assays due to limited resources available, it is important to efficiently allocate the number of assays at each sampling time point included in the study. As indicated earlier, we may consider either to place more assays at the latest sampling time point and the time point at which the curvature is expected to occur or to place equal number of assays at each sampling time point. However, there are little or no literature available for selection of sampling time points and/or the allocation of assay numbers at each time point. The problem is then worth of investigation.

IV. *Drug Characteristics*

Generally there are different criteria for acceptable levels of stability with respect to chemical, physical, microbiological, therapeutic, and toxicological characteristics of drug products[20]. The requirements of stability on these five characteristics are also different from dosage form to dosage form. Table 1 gives a list of drug characteristics for different dosage forms which should be evaluated in a stability study. As indicated earlier, the objective of stability studies is to characterize the degradation of drug products in terms of some essential drug characteristics and consequently to establish an expiration dating period. The approach suggested in the FDA guideline for determination of drug shelf-life is primarily based on single drug characteristic such as strength. The strength of a drug product is defined as either (I) the concentration of the

Table 1. Drug characteristics for different dosage forms

Dosage Form	Drug Characteristics
Tablets	Appearance, Friability, Hardness, Color, Odor, Moisture, Strength, Dissolution
Capsules	Strength, Moisture, Color, Appearance, Shape, Brittleness, Dissolution
Emulsions	Appearance, Color, Odor, pH, Viscosity, Strength
Oral Solution & Suspensions	Appearnace, Strength, pH, Color, Odor, Redispersibility, Dissolution, Clarity
Oral Powder	Appearance, pH, Dispersibility, Strength
Metered Dose Inhaler Aerosols	Strength, Delivered dose per actuation, Number of metered doses, Color, Clarity, Particle size, Loss of Propellant, Pressure, Valve corrosion, Spray pattern
Topical and Ophthalmic Preparations	Appearance, Clarity, Color, Homogeneity, Odor, pH, Resuspendibility, Consistency, Particle size distribution, Strength, Weight loss
Small-volume Parenterals	Strength, Appearance, Color, Particulate matter, pH, Starility, Pyrogenicity
Large-volume Parenterals	Strength, Appearance, Color, Clarity, Particulate matter, pH, Volume, Extractables, Starility, Pyrogenicity
Suppositories	Strength, Softening range, Appearance, Dissolution

drug substance or (II) the potency, i.e., the therapeutic activity of the drug product which can be determined by an appropriate laboratory test or by an adequately developed and controlled clinical data. However, in the FDA guideline, the strength of a drug product is interpreted as a quatitative measure of the active ingredient of a drug product, as well as other ingredients requiring quantitation, such as alcohol and preservatives. For the analysis of stability data, the FDA requires that percent of label claim, not percent of initial average value, be used as the primary variable for strength. Furthermore, the stability of the characteristics of a drug product ofr a particular dosage form may be influenced by the storage conditions such as temperature, humidity, light, or air and by the package types such as high-density polyethylene (HDPE).

For any drug product that is intended for use as an additive to another drug product, the possibility of incompatabilities may exist. In such cases, the FDA guideline requires that the drug product labeled to be administered by addition to another drug product (e.g., parenterals or aerosols) be studied for stability and compatability in a mixture with the other drug product. A suggested stability protocol should provide for tests to be conducted at 0, 6 to 8, and 24 hour intervals, or as appropriate over the intended use period. These thests should include assay of the drug product and additive, pH (especially for unbuffered large-volume parenterals), color, clarity, particulate matter, and interaction with the container.

As indicated in Table 1, for a given dosage form, the FDA guideline requires a number of drug characteristics be evaluated for determination of drug shelf-life. However, in most stability studies, the determination of drug shelf-life is usually based on the primary drug characteristic of interest such as strength (or potency) of the drug product rather than all of the drug characteristics. On the other hand, a drug product may have more than one active ingredients. In practice, to fulfill with the FDA requirements, we may determine shelf-lives for each drug characteristic of each active ingredient and consider the minimum shelf-life if different drug characteristics among different active ingredients have the shelf -life of the drug product.

DISCUSSION

Although an expiration dating period x_L can be determined from (3.2), it is of interest to study statistical properties of x_L. Since x_L is a point estimate, it may overestimate or underestimate the true shelf-life. If the labeled shelf-life underestimates the true shelf-life, then the drug product will still remain its identity, strength, quality, and purity beyond the expiry period. On the other hand, if the labeled shelf-life overestimates the true shelf-life, then the drug product may not be safe even prior to its expiration date.

In a recent EWG meeting held in Washington D.C., the EC, the Ministry of Health and Welfare (MHW) of Japan and the FDA representatives had reached a consensus on the implementation of the ICH guideline. They recommended that a three-year phase in period be allowed to implement once the ICH guideline is adopted by the three authorities. Thus, the requirements of the harmonized guideline would have to be met at the time of submission of applications by January of 1997.

REFERENCES

1. ICH 1993. Stability Testing of New Drug Substances and Products. Tripartitie International Conference on Harmonization Guideline.
2. WHO 1993. World Health Organization Guidelines on Stability Testing of Pharmaceutical Products Containing Well-established Drug Substances in Conventional Dosage Forms. Reference WHO/PHARM/92, 158/Rev.4.
3. FDA 1987. Guideline for Submitting

Documentation for the Stability of Human Drugs and Biologics. Center for Drugs and Biologics, Office of Drug Research and Review, United States Food and Drug Administration, Rockville, Maryland.

4. Kohberger, R.C. 1988. Manufacturing and Quality Control. Chapter 14 in Biopharmaceutical Statistics for Drug Development. Ed. K. Peace. pp.605-629, Marcel Dekker, Inc. New York, U.S.A.
5. Easterling, R. G. 1969. Discrimination Intervals for Percentilies in Regression. J. Amer. Stat. Assoc. 64: 1031-1041.
6. Carstensen, J.T. and Nelson, E. 1976. Terminology Regarding Labeled and Contained Amounts in Dosage Forms. J. Pharm. Sci. 65: 311-312.
7. Chow, S.C. and Liu, J.P. 1995. Statistical Design and Analysis in Pharmaceutical Science. pp. 329-353, Marcel Dekker, Inc. New York, U.S.A.
8. Nordbrock, E. 1992. Statistical Comparison of Stability Study Designs. J. Biopharm. Stat. 2: 91-113.
9. Ju, H.L. and Chow, S.C. 1995. On Statistical Design of Stability Studies. J. Biopharm. Stat. 5: 201-214.
10. Chow, S.C. and Shao, J. 1989. Test for Batch-to-batch Variation in Stability Analysis. Stat. Med. 8: 883-890.
11. Bancroft, T.A. 1964. Analysis and Inference for Incompletely Specified Models Involving the Use of Preliminary Test (s) of Significance. Biometrics. 20: 427-442.
12. Schuirmann, D.J. 1987. A Comparison of the Two One-sided Tests Procedure and PowerApproach for Assessing the Equivalence of Average Bioavailability. J. Pharm. Biopharm. 15: 657-680.
13. Chow, S.C. and Liu, J.P. 1992. Design and Analysis of Bioavailability and Bioequvalence Studies. pp. 88-89, Marcel Dekker, Inc., New York, U.S.A.
14. Chow, S.C. and Shao, J. 1991. Estimating Drug Shelf-life with Random Batches. Biometrics 47: 1071-1079.
15. Ruberg, S. and Hsu, J. 1992. Multiple Comparison Procedures for Pooling Batches in Stability Studies. Technometrics. 34: 465-472.
16. Murphy, J.R. and Weisman, D. 1990. Using Random Slopes for Estimating Shelf-life. Proceedings of the Biopharmaceutical Section of the American Statistical Association. pp. 196-203, American Statistical Association, Alexandria, Va. U.S.A.
17. Shao, J. and Chow, S.C. 1994. Statistical Inference in Stability Analysis. Biometrics 50: 753-763.
18. Ho, C.H., Liu, J.P. and Chow, S.C. 1992. On Analysis of Stability Data. Proceedings of the Biopharmaceutical Section of the American Statistical Association. pp. 198-203, American Statistical Association, Alexandria, Va. U.S.A.
19. Morris, J.W. 1992. A Comparison of Linear and Exponential Models for Drug Expiry Estimation. J. Biopharm. Stat. 2: 83-90.
20. USP/NF 1990. The United States Pharmacopeia XXII and the National Formulary XVII. United States Pharmcopeial Convention, Inc., Rockville, Md. U.S.A.

Journal of Food and Drug Analysis. 1995. 3(2)

藥物安定性規範之研究分析

周賢忠　彭安佩

美國必治妥—施貴寶藥廠　美國羅勃藥廠

摘　要

對於市場上的任何一種藥物，政府規定其包裝標籤上必須註明該藥物的有效期限(expiration dating period)，藥物的有效期限不僅提供消費者對該藥物的信心且保證在該藥物有效期限內仍然保持該藥物應具有的良藥特性，諸如成份(identity)，效力(strength)，品質(quality)以及純度(purity)。一般而言，藥廠經由藥物安定性的測試研究來描述藥物退化曲線進而估計藥物的有效期限。在這篇文章中，我們對現行美國安定性規範及國際標準化規範做了深入的探討與評估。同時，我們也列舉了一些經由評估所引發在藥物安定性測試中，實驗設計與統計分析的一些未來問題的研究方向。