PERIODICA POLYTECHNICA SER. CHEM. ENG. VOL. 48, NO. 1, PP. 41–52 (2004)

EFFECT OF THE UNCERTAINTY OF THE STABILITY DATA ON THE SHELF LIFE ESTIMATION OF PHARMACEUTICAL PRODUCTS

Kinga KOMKA and Sándor KEMÉNY

Department of Chemical Engineering
Budapest University of Technology and Economics
H–1521 Budapest, Hungary
Phone: 0036 1 4632209, Fax number: 0036 1 4633197
e-mail: kemeny@mail.bme.hu

Received: May 11, 2004

Abstract

The estimated shelf life of a drug product is highly influenced by the variability of the measured data. The fluctuation of the stability data is composed of the manufacturing process variation (batch-to-batch and within batch variability) and of the uncertainty of the analytical method (reproducibility and repeatability). The aim of the paper is to show a calculation method by which all of the variance components can be estimated before commencing the stability study. The effect of the uncertainty on the estimated shelf life is also considered: the expected variance of a single stability time point and the width of the 95% one-sided confidence limit after 2 storage years are calculated. For the computation the results of the content uniformity test and the validation (specifically the precision study) of the analytical method are used. The applied mathematical method is the analysis of variance. The advantage of the concept is that if the magnitude of the uncertainty is known in advance, one may consider whether the present manufacturing process and analytical method is suitable for the stability study.

Keywords: pharmaceutical stability data, method uncertainty, variance components, shelf life.

1. Introduction

In the pharmaceutical industry stability study is performed to estimate the shelf life of a drug product. The purpose of the stability study is to provide evidence on how the quality of a drug product varies with time under the influence of a variety of environmental factors such as temperature, humidity and light, and to establish the shelf life for the drug product and the recommended storage conditions [6]. A complete stability design includes accelerated and long term stability testing. Stability data obtained from the long term testing are used directly for the shelf life estimation, so this paper focuses only on the long term stability testing. The shelf life estimation is usually performed in the early stage of drug development, when limited number of stability data is available.

There are three steps in the stability analysis [8]. The first step is to collect the stability data at several time intervals for the samples stored under appropriate conditions. Stability data should cover all of those attributes of the drug product

that are susceptible to change during the storage and are likely to influence quality, safety and efficacy [6]. For the sake of brevity we concentrate here only on the active ingredient content analyses. The second step of the analysis is to choose the adequate model to describe the relationship between the measured data and the storage time. In this paper we assume that the assay content of the drug product decreases linearly with time, which is the commonly applied approach for describing the degradation of active ingredients [2]. The third step of the analysis is to estimate the shelf life based on the information from all batches assayed. According to the ICH Guideline [5] for an attribute known to decrease with time (e.g. the active ingredient content), the shelf life of the drug product is calculated as the time point where the lower one-sided 95% confidence limit for the mean degradation curve intersects the minimal required content. In most cases there is also an upper limit, which may cause troubles even if there is no real increase in assay content. This problem is not dealt with in this paper. The width of the confidence band depends on the number of experimental points, thus it would be advantageous to combine the data from the batches. Pooling the data should be supported by preliminary testing of batch similarity. The α value for these statistical tests is usually taken as 0.25. If the tests for equality of slopes and intercepts are accepted, the shelf life estimation is based on the pooled data. However, if the hypothesis of batch similarity is rejected, the shelf life is determined on the basis of the minimum of individual shelf lives obtained for each batch.

The method of the ICH Guideline has been criticized by several papers. Some papers contest the justness of the 0.25 significance level for the preliminary testing [10, 13]. CHOW and SHAO [4] state that if there is batch-to-batch variation, the shelf life estimated on the basis of one (the worst) batch does not represent the shelf life of all future batches manufactured under similar conditions. Thus the random effects model seems more adequate than the assumption of the fixed batch effect [3, 11]. In this paper the fixed effects model was used, however. There are several proposals to estimate model parameters of the stability study and consider adequately the error structure. In our former paper [7] a modified random error model was proposed for the shelf life estimation.

This paper accepts the Guideline model in respect that how the uncertainty of the stability data influences the shelf life estimation. The shelf life is the point of intersection of the confidence limit and the acceptance criterion. Assuming linear function, the two-sided confidence limits for a given time point (x) are given as:

$$\hat{y} - t_{\alpha/2} s_y \sqrt{\frac{1}{n} + \frac{x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2}} < y < \hat{y} + t_{\alpha/2} s_y \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2}}, \qquad (1)$$

where $\hat{y}$ is the estimated value at the given time point, $t_{\alpha/2}$ is the Student's t coefficient at α (two-sided) risk and $n-2$ degrees of freedom, s_y is the standard deviation of the measured data, n is the number of the time points, x_j is the j^{th} time point and $\bar{x}$ is the mean of the time points. As it is seen from the equation above the width of the confidence band and thus the shelf life depends on the standard deviation of

the measured data (s_y). This means that highly variable stability data reduce the estimated shelf life.

Although the variability of the stability data is a common reality, the effect of the manufacturing and analytical method uncertainty on the shelf life has been studied only in a few articles. BAR [1] divided the reasons for change of measured data into two parts: the decrease of assay during the storage period and the random variation of the analytical method. He estimated the repeatability and reproducibility of an analytical method using the data from stability study of a stable drug. VAN DER VEEN et al. [12] discussed the uncertainty components of a stability study of certified reference materials in detail and showed how these components could be determined from the stability study itself. PAUWELS et al. [9] carried out a lot of simulations with different degradation rates, coefficients of variation and confidence levels to demonstrate the effect of these parameters on the shelf life estimation.

In the papers mentioned the variance components of the uncertainty have been estimated from the stability study (of a drug product or a certified reference material) itself. It would be, however, useful to obtain the expected uncertainty of the stability data before the start of the stability study, because if the magnitude of the uncertainty is known in advance, one may consider whether the present manufacturing process and analytical method is suitable for the stability study. Thus, the aim of this paper is to show a calculation method, by which all of the uncertainty components can be estimated before the start of the stability study and their effect on the shelf life estimation is considered in advance.

2. Method of Calculation

The fluctuation of the stability data is composed of the manufacturing process variation and the uncertainty of the analytical method. For the computation the following elementary variance components are taken into consideration:

- batch-to-batch variability (σ^2_{batch}),
- heterogeneity within each batch ($\sigma^2_{\text{heterogeneity}}$),
- period-to-period variability ($\sigma^2_{\text{circumstances}}$, for example σ^2_{day} or σ^2_{lab}),
- replicate-to-replicate variability (σ^2_{error}).

The batch-to-batch and the within batch variability belong to the manufacturing process variation. Consideration of the effect of these variance components is justified, because during a stability study several (usually three) batches are tested and at each time point or replicate different tablets are analysed.

The other two variance components join to the uncertainty of the analytical method. The replicate-to-replicate variability is characterized with the repeatability of the analytical method. The period-to-period variability is in principle reflected by the intra-laboratory reproducibility (intermediate precision), when the analytical method is replicated on different days. This analysis is generally performed on

days in succession. Opposite to this at the stability study several months are gone between the dates of analysis, so the circumstances of the measurements (e. g. analysts, equipments, reagents) may differ more from each other. We consider therefore that the scenario of performing analysis in so much different time points is often represented more adequately by the inter-laboratory reproducibility (when the analytical method is replicated under completely different conditions).

To estimate the contribution of these variance components to the total variance results of two kinds of analyses of the analytical developing procedure may be used: the content uniformity test and the validation of the analytical method. These analytical measurements are normally performed during the development of a drug product, thus no extra analysis is required to the calculation.

Content Uniformity Test

During the content uniformity test 10–10 tablets from usually 3 batches are analysed individually. By analysis of variance two variance components can be separated: σ_{batch}, which is the effect of batch-to-batch variability and σ^2_{tablet}, which includes both the effect of heterogeneity and the replicate-to-replicate variability. The effect of heterogeneity and the replicate-to-replicate variability can not be separated from one another, because the tablets are shattered during the analytical measurement, so the analysis may not be repeated with the same tablet. Since the tablets are analysed individually, the σ^2_{tablet} variance component is a simple combination of the effect of heterogeneity and the replicate-to-replicate variability:

$$\sigma^2_{\text{tablet}} = \sigma^2_{\text{heterogeneity}} + \sigma^2_{\text{error}}. \tag{2}$$

Validation of the Analytical Method

The main objective of the validation of an analytical method is to demonstrate that the procedure is suitable for its intended purpose. One of the examined validation characteristics is the precision, which is the measure of the random error. The determination of the precision may be subdivided into three steps: the examination of repeatability, intermediate precision and reproducibility.

Repeatability Study

Repeatability defines the ability to repeat the same analytical methodology by the same analyst, using the same equipment in the same laboratory, within the same day. Evaluating the data $\sigma^2_{\text{repeatability}}$ variance is given. This variance contains both the effect of heterogeneity and the replicate-to-replicate variability, because the determination of repeatability (similarly to the content uniformity test) may not be

repeated with the same tablets. However, opposite to the content uniformity test, during the testing of repeatability a mixture of several (usually 4–6) tablets are analysed within one replicate. In this case *Eq.* (2) is to be modified, because the effect of heterogeneity is averaged for replicates:

$$\sigma^2_{\text{repeatability}} = \sigma^2_{\text{heterogeneity}}/p + \sigma^2_{\text{error}}, \tag{3}$$

where p is the number of the tablets used at one replicate.

It should be noted, that the above equation is valid only in cases when the repeated measurements are always performed from new tablets. If the same homogenized pool of tablets is used for the repetition, the heterogeneity does not come into the picture, so the above relationship becomes simpler: $\sigma^2_{\text{repeatability}} = \sigma^2_{\text{error}}$. Sometimes even in case of homogenized sample the preparation method may have inherent bias, this will not be considered here.

To separate the effect of heterogeneity from the replicate-to-replicate variability the difference between the content uniformity test and the determination of repeatability can be applied. Using the *Eq.* (2) and (3) the effect of heterogeneity ($\sigma^2_{\text{heterogeneity}}$) and the replicate-to-replicate variability (σ^2_{error}) are separately given:

$$\sigma^2_{\text{heterogeneity}} = (\sigma^2_{\text{tablet}} - \sigma^2_{\text{repeatability}})/(1 - 1/p) \tag{4}$$

$$\sigma^2_{\text{error}} = \sigma^2_{\text{tablet}} - \sigma^2_{\text{heterogeneity}}. \tag{5}$$

Intermediate Precision Study

Intermediate precision is the ability to replicate the same analytical methodology in the same laboratory, but by different analysts or using different equipment or performing on different days. Analysis of variance leads to two variance components: $\sigma^2_{\text{circumstances}}$ refers to the different circumstances, while $\sigma^2_{\text{repeatability}}$ concerns the repeatability under the same conditions. The variance component of different circumstances shows either the effect of different analysts (σ^2_{person}) or equipments ($\sigma^2_{\text{equipmemt}}$) or days ($\sigma^2_{\text{day}}$).

Reproducibility Study

Reproducibility is the ability to repeat the same analytical methodology under completely different conditions: namely in different laboratories by obviously different analysts, using different equipments. Two variance components are given by analysis of variance: σ^2_{lab} belongs to the different laboratories and $\sigma^2_{\text{repeatability}}$ characterizes the repeatability under the same conditions.

It is apparent that from the precision study several (exactly three) estimates for $\sigma^2_{\text{repeatability}}$ variance component are given, namely from the examinations of repeatability, intermediate precision and reproducibility.

Table 1 summarises the two steps of the variance component estimation described above. The second and the third columns show the relationship between the estimated and the elementary variance components. It is worth mentioning that naturally only the estimates of these variances ($\hat{\sigma}^2$) are available, they may be calculated from the measured data.

Table 1. The estimation method of the variance components

Analytical measurements	The determined variance component	The elementary variance component
Content uniformity test	σ^2_{batch}	σ^2_{batch}
	σ^2_{tablet}	$\sigma^2_{\text{heterogeneity}}$ and σ_{error}
Validation Precision study	σ^2_{day}	σ^2_{day}
	σ^2_{lab}	σ^2_{lab}
	$\sigma^2_{\text{repeatability}}$	$\sigma^2_{\text{heterogeneity}}$ and σ^2_{error}

The Effect of the Uncertainty of the Stability Data on the Shelf Life Estimation

In this section it is outlined how the manufacturing and analytical method uncertainty influences the shelf life estimation. The first step of the computation is to determine the expected variance ($\hat{\sigma}^2$) of a single stability time point. The estimated elementary variance components mentioned in *Table 1* are used to the calculation:

$$\hat{\sigma}^2 = \frac{\hat{\sigma}^2_{\text{error}} + \dfrac{\hat{\sigma}^2_{\text{heterogeneity}}}{p}}{r} + \hat{\sigma}^2_{\text{batch}} + \hat{\sigma}^2_{\text{day}} + \hat{\sigma}^2_{\text{lab}}, \tag{6}$$

where r is the number of replicates at each time point. If the calculation is based on a single batch (batches are treated separately), σ^2_{batch} is obviously not considered.

Eq. (1) contains the standard deviation of stability data, usually obtained in the course of regression analysis. As our intent is to predict this standard deviation, but without having the stability data themselves, the $\hat{\sigma}^2$ value calculated by *Eq.* (6) is inserted in two ways into *Eq.* (1). In both cases stability study was modelled assuming linear degradation of the active ingredient content. The effect of the uncertainty was characterized by the width of the confidence limit.

In the first approach the square root of the estimated overall variance resulted by *Eq.* (6) was taken as s_y standard deviation. The assumed degrees of freedom is equal to the degrees of freedom of s_y as if it were estimated in the course of

regression analysis of stability data. The other parts of *Eq.* (1) are determined by the design of the stability study.

In the second approach the estimated elementary variance components (σ^2_{error}, $\sigma^2_{\text{heterogeneity}}$, σ^2_{batch}, σ^2_{day}, σ^2_{lab}) were taken as variances and a simulation study was performed to produce stability data (concentration values at different time points). The simulated points obtained this way reflect all the sources of fluctuation as if they were really measured. Then a straight line was fitted to the data and the confidence limit was calculated using the residual standard deviation.

3. Example

In this part concrete examples are shown to illustrate the process of the estimation of the variance components and to see the effect of the uncertainty of the stability data on the shelf life estimate.

These specific examples contain data on drug products where there is serious matrix effect in analysis. In addition the active ingredient content is low in the products, thus the analysis is a difficult task. The active ingredient content of the two drug products (A and B) was determined by an HPLC method. Details and source of the data are not given for confidentiality reason.

The design of the precision study was composed of the following steps. The determination of repeatability was performed with 6 repeated measurements under the same conditions. For the determination of intermediate precision 3 repeated measurements were carried out on 3 different days. During the determination of reproducibility the analytical method was repeated in 2 different laboratories performing 6–6 measurements in each laboratory on one day. At each replicate above 5–5 tablets were homogenised and analysed from a batch.

The determination of intermediate precision was also performed by individual tablets: 10–10 tablets were analysed individually on 3 different days. The results of this analysis were used to estimate the heterogeneity.

4. Results and Discussion

4.1. Variance Component Estimation

Estimation of the Repeatability Variance Component

The repeatability variance component can be estimated from all the three kinds of examinations of the precision study. The results performed by ANOVA are summarized in *Table 2*. The last row of the table contains the pooled mean square values from the three repeatability variance components.

Table 2. The estimated variance components of repeatability (assay %)

The source of the variance component estimation	$\hat{\sigma}^2_{\text{repeatability}}$ A	B
Repeatability study	1.67	1.13
Intermediate precision study	0.71	0.76
Reproducibility study	1.38	1.27
Pooled mean square	*1.26*	*1.09*

Estimation of the Effect of Days

The variance component regarding the days can be estimated from the examination of intermediate precision. The results are summarized in *Table 3*. As the intermediate precision study was performed both for individual tablets and samples containing more tablets, we have two estimates per drug products for the effect of days.

Table 3. The estimated variance components of the days (assay %)

The source of the variance component estimation	$\hat{\sigma}^2_{\text{day}}$ A	B
Intermediate precision study	4.48	1.81
Intermediate precision study for individual tablets	3.76	0

If the estimates shown in *Table 3* are compared row by row, it can be proved that the variance components estimated from the intermediate precision study for individual tablets are lower. The reason for this is that comparing two designs of experiments for the observation of the same effect of a factor by one-way ANOVA is more difficult when the error (within) mean square is higher. At the examination of individual tablets the repeatability mean square is much higher due to the greater heterogeneity variance component, so the effect of the days can be observed with only lower probability.

It is also apparent from *Table 3*, that the days of analysis have a significant effect on the measured data. At the active ingredient A the variation caused by the days is particularly high.

Estimation of Reproducibility

The data from the determination of reproducibility were evaluated using one-way ANOVA. It was found that at this analytical method the laboratory as a factor has no effect on the measured data at 5% significance level.

Estimation of the Effect of Heterogeneity

The intermediate precision study for individual tablets is (similarly to the content uniformity test) also suitable for the estimation of the effect of heterogeneity. From the data the σ^2_{tablet} variance component can be given directly.

For the separation of the effect of heterogeneity from the replicate-to-replicate (error) variability *Eq.* (4) and (5) is used. The results summarized in *Table 4* show that at these drug products the effect of heterogeneity is much higher than the replicate-to-replicate variability.

Table 4. The estimated variance components belonging to the individual assay determination and the results of the separation of the effect of heterogeneity and the replicate-to-replicate variability (assay %)

	A	B
$\hat{\sigma}^2_{\text{tablet}}$	5.44	4.73
$\hat{\sigma}^2_{\text{heterogeneity}}$	5.23	4.55
$\hat{\sigma}^2_{\text{error}}$	0.21	0.18

The variance components given in *Tables 2–4* are only point estimates of the theoretical value. Because of the few replicates these point estimates have a great uncertainty. The effect of the uncertainty will be considered by the shelf life calculation of the drug product.

4.2. Use of Variance Components in Assessing a Virtual Design

Let us consider a general stability design, when 3 batches are tested and the testing frequency is every 3rd month over the first year. At each time point 3 repeated measurements are performed and at each replicate 5–5 tablets are homogenized and analyzed once.

The expected uncertainty of the stability data is characterized with the variance of a single time point and the width of the 95% one-sided confidence limit after 2 storage years. All calculations are performed for two cases:

a. the data of the three batches may not be pooled, so the shelf life estimation is based on the data of one (the worst) batch,
b. the data of the three batches may be pooled.

Table 5 summarizes the expected variance of a single time point calculated by *Eq.* (6) and the width of the confidence limit computed by the *first concept*, where s_y is the square root of the estimated overall variance. Since σ^2_{batch} is not available at these drug products, an average value (1.5) was used.

Table 5. The uncertainty of a single stability datum point and the calculated confidence limit after 2 years (assay %)

		A	B
The batches may not be pooled	$\hat{\sigma}^2$	4.90	2.17
	width of the 95% one-sided confidence limit	10.2	6.8
The batches may be pooled	$\hat{\sigma}^2$	6.40	3.67
	width of the 95% one-sided confidence limit	5.1	3.8

Table 6 shows the results using the *second approach* (simulation study) for the "A" active ingredient. Comparing the width of the confidence limit obtained by the two approaches it can be stated that the order of magnitude of the results is similar.

Table 6. The results of the simulation study (assay %)

'A' active ingredient	The width of the 95% one-sided confidence limit after 2 years	
	Mean*	Standard deviation*
The batches may not be pooled	8.98	3.84
The batches may be pooled	4.06	1.24

(*calculated from the 1000 cases)

The specification limit for the active ingredient content is usually 90–105% at the end of shelf life. Let us compare this specification limit with the width of confidence limit given in *Table 5* or *6*. If the data of the examined batches may not be pooled and the effect of analysis dates is high (for example at the 'A'

active ingredient), the width of confidence limit may be as high as 10%. This means that even in the favourable case when the initial active ingredient content is approximately 100% and no degradation occurs, the lower limit of the confidence bound intersects the specification limit already at the end of the 2nd year. If the data of the examined batches may be pooled, the width of confidence limit becomes narrower (3.8–5%). However, if the initial active ingredient content decreases by 5% over 2 years, the 5% width of the confidence limit means that the estimated shelf life will be also only about 2 years.

It should be also noted that at this analytical method the effect of inter-laboratory reproducibility was not significant. If $\hat{\sigma}^2_{\text{lab}} > 0$, the expected variance of a single time point will be even higher and the width of the confidence-limit becomes wider, because the additional fluctuation caused by different laboratories (which means completely different conditions of the analytical measurements) may be also taken into consideration.

5. Conclusions

In the paper a calculation method is shown by which the fluctuation of the stability data is forecasted. First all of the uncertainty components influencing the variability of stability data are determined using the results of two analyses (the content uniformity test and the validation study) of the developing procedure of a drug product. At the second step the effect of this uncertainty on the shelf life is estimated.

We stated that the highly variable stability data significantly reduce the estimated shelf life. The results of our examples may not be general, but not unreal. Thus, it is worth considering the source of the uncertainty at the planning of any stability and validation studies.

Acknowledgement

This work has been supported by the Hungarian National Science Foundation (OTKA) under grants no. T 033005.

References

[1] BAR, R., Statistical Evaluation of Stability Data for Assessing Variability of an Analytical Method, *Accredit. Qual. Assur.*, **4/6** (1999), pp. 235–239.
[2] CHOW, S. C. – LIU, J. P., *Statistical Design and Analysis in Pharmaceutical Science*, Marcel Dekker, New York, 1995.
[3] CHOW, S. C. – SHAO, J., Estimating Drug Shelf-Life with Random Batches, *Biometrics*, **47** (1991), pp. 1071–1079.
[4] CHOW, S. C. – SHAO, J., Test for Batch-to-Batch Variation in Stability Analysis, *Stat. Med.*, **8** (1989), pp. 883–890.
[5] ICH Harmonised Tripartite Guideline, Evaluation for Stability Data, 2003.

[6] ICH Harmonised Tripartite Guideline, Stability Testing of New Drug Substances and Products, 2003.
[7] KOMKA, K. – KEMÉNY, S., A Modified Error Model for the Assessment of Stability of Pharmaceutical Products, *Chemometr. Intell. Lab.*, in Press.
[8] LIN, K. K. – LIN, T. Y. D. – KELLY, R. E., in BUNCHER, C. R. – TSAY, J. Y., (Eds.): *Statistics in the Pharmaceutical Industry*, Marcel Dekker, New York, 1994, pp. 419–444.
[9] PAYUWELS, J. – LAMBERTY, A. – SCHIMMEL, A., Quantification of the Expected Shelf-Life of Certified Reference Materials, *Fresen. J. Anal. Chem.*, **361** (1998), pp. 395–399.
[10] RUBERG, S. J. – STEGEMAN, J. W., Pooling Data for Stability Studies: Testing the Equality of Batch Degradation Slopes, *Biometrics*, **47** (1991), pp. 1059–1069.
[11] SHAO, J. – CHOW, S. C., Statistical Inference in Stability Analysis, *Biometrics*, **50** (1994), pp. 753–763.
[12] VAN DER VEEN, A. M. H. – LINSINGER, T. P. J. – LAMBERTY, A. – PAUWELS, J., Uncertainty Calculations in the Certification of Reference Materials 3. Stability Study, *Accredit. Qual. Assur.*, **6/6** (2001), pp. 257–263.
[13] WESSELS, P. – HOLZ, M. – ERNI, F. – KRUMMEN, K. – OGORKA, J., Statistical Evaluation of Stability Data of Pharmaceutical Products for Specification Setting, *Drug. Dev. Ind. Pharm.*, **23/5 (**1997), pp. 427–439.