ELSEVIER

International Journal of Pharmaceutics 188 (1999) 97–104

international journal of pharmaceutics

www.elsevier.com/locate/ijpharm

An opportunistic stability strategy; simulation with real data

K. Dyrstad [a,*], C. Thomassen [b], K. Eivindvik [b]

[a] *Pharmaceutical Sciences, R&D, Nycomed Imaging AS, PO Box 4220 Torshov, N-0401 Oslo, Norway*
[b] *Stability Department, Nycomed Imaging AS, PO Box 4220 Torshov, N-0401 Oslo, Norway*

Received 7 April 1999; received in revised form 2 June 1999; accepted 9 June 1999

Abstract

A multivariate modelling procedure is proposed in order to identify factors influencing stability, to estimate shelf-life, to select new batches for further stability testing and to evaluate changes in new batches. A model developed by the proposed procedure predicts the degradation rate constant as a function of storage temperature, pH, concentration and volume. The predicted rate constants were compared with prospectively measured rate constants, primarily from new batches stored under stress conditions, which emphasised batch differences earlier than storage under normal conditions. Strong deviations from expected rate constants led to extended testing of the batches concerned. The new data were used to upgrade the multivariate model. The procedure proposed led to the formulation of an opportunistic stability strategy (OSSY). Of the 15 batches of injectable solutions, nine batches are proposed tested by using OSSY. This led to an approximately 75% reduction in analytical measurements. Hold samples are recommended for storage under several stability conditions for back up analysis. In general, a multivariate stability model should be based on scientific data obtained from early studies, such as preformulation and formulation studies to provide both a qualitative and quantitative understanding of the mechanisms involved.

Keywords: Stability design; Rate constant; Multivariate methods; Partial least squares

1. Introduction

The concepts of matrixing and bracketing for the estimation of shelf-life and batch to batch homogeneity at an acceptable cost have been thoroughly discussed (Ruberg and Stegeman, 1991; Helboe, 1992; Fairweather et al., 1995). As a result of the ICH process, guidelines have been made to suggest possible fractional stability designs which take account of the number and type of variables to be tested. Formulation variables such as drug concentration, container volume, packaging type, number of product batches and variation in raw material are typical main effects to be determined during a stability evaluation. However, reliable estimates of shelf-life also depend on additional factors such as Arrhenius estimations, experience in mathematic interpretation of scientific data, product homogeneity/uniformity, biological acceptability of product

* Corresponding author. Tel.: +47-23-18-57-72; fax: +47-23-18-60-00.
E-mail address: krd@nycomed.com (K. Dyrstad)

0378-5173/99/$ - see front matter
PII: S0378-5173(99)00209-4

variations and relevant climatic zones. Analytical precision expressed as repeatability, intermediate precision, reproducibility and robustness will also influence any relevant mathematical and statistical evaluation of the data. Information on all the factors above is usually of relevance to both regulatory authorities and manufacturers.

The time dependency of product change, can typically be examined by ordinary least square (OLS) regression analysis and variance analysis for testing the possibility of pooling batches (ANCOVA) to increase precision in shelf life estimates (Ruberg and Hsu, 1991; Ruberg and Stegeman, 1991; Fairweather et al., 1995). Ruberg and Hsu (1991) investigated the Tukey-Kramer method for multiple comparisons of slopes and multiple comparison of slopes with the worst slope as an alternative approach to ANCOVA. The Bayesian approach by Su et al. (1994) quantified the uncertainty in predicting shelf life from Arrhenius equation as a function of stability data consisting of various error distributions. Yoshika et al. (1996) used Monte Carlo simulation to test the power of ANOVA on matrixing data and single data to estimate shelf-life. The simulation showed that the large amount of data from a matrixing design gave more precise shelf-life estimates. The variation in Arrhenius plots as a function of pH was studied by Carstensen et al. (1992). Ertel and Carstensen (1990) improved the estimates in classical Arrhenius plotting by transforming data to their natural logarithm.

During the early stages of product development the knowledge about relevant future concentrations (doses), container volumes, packaging, etc., is only tentative. Therefore stability testing tends to be piecemeal and are undertaken throughout the products lifetime. Often, this results in testing one factor at a time with several batches at several conditions.

The primary objective in this paper is to make a scientific case for opportunistic stability strategy (OSSY). This strategy is based on the construction of a multivariate model to estimate the relationship between temperature, product concentration, container volume, etc., and the degradation rate constant. Predicted rate constants will be compared with rate constants gained from minor but highly informative stress studies of selected batches/formulations. The multivariate model should be built from as much relevant scientific data as possible to increase the power of prediction.

2. Materials and methods

Stability data from 15 batches of the diagnostic X-ray contrast agent, Visipaque (Aars and Eivindvik, 1995) filled in glass bottles and stored at elevated temperatures were used. The stability design is shown in Table 1. The 15 batches are referred to using letters from A to O and the corresponding numbers refer to the storage condition described in Table 1. The active ingredient and the finished product were produced at Lindesnes Plant and Oslo Plant, respectively, Norway, Nycomed Imaging AS.

The most important stability indicating parameter, inorganic iodide, was measured by potentiometric argentometric titration (Aars and Eivindvik, 1995). Change in inorganic iodide in g/ml per month was estimated by univariate least square regression and was expressed as k. The linearity is shown by the correlation coefficient (r^2) in Table 1. The measured k was transformed to logarithmic k, $\ln(k)$, to be used for multivariate mathematical modelling. The variation in $\ln(k)$ was determined as function of measured $[H^+]$, temperature (K^{-1}), product concentration (μg/ml) and container volume (ml) by using data from 25 to 50°C (Table 2).

The following multivariate linear relationship between $\ln(k)$ and the four variables were constructed;

$$\ln(k) = \beta_0 + \beta X_1 + \ldots + \beta_4 X_4 \quad (1)$$

where

$\ln(k)$ is the logarithmic rate constant
X_n is the value of variable
β_0 is a constant
β_n is a regression coefficient (effect of each variable)

Modeling was done by partial least squares (PLS) regression in Unscrambler 6.11, Camo AS, Trondheim, Norway.

Table 1
The batches and the conditions for stability testing of Visipaque and the calculated degradation rate k with the corresponding correlation coefficient r^2

Batch (letter)	Study start	Used for:	Temperature.(°C)	Concentration (mg/ml)	Volume (ml)	pH	k	r^2	In(k)
A11	01-91	Back up	4	150	50	7.50	0.02	0.976	−4.09
A12	01-91	Eq. (1)	25	150	50	7.50	0.14	0.956	−1.98
A13	01-91	Eq. (1)	40	150	50	7.40	0.84	0.983	−0.17
A14	01-91	Eq. (1)	50	150	50	7.50	5.29	0.987	1.67
B11	01-91	Back up	4	150	50	7.40	0.02	0.908	−3.83
B12	01-91	Eq. (1)	25	150	50	7.40	0.11	0.961	−2.17
B13	01-91	Eq. (1)	40	150	50	7.40	1.19	0.991	0.17
B14	01-91	Eq. (1)	50	150	50	7.40	4.86	0.987	1.58
G11	11-91	Back up	4	150	50	7.20	0.00	0.632	−5.70
G12	11-91	Backup	25	150	50	7.20	0.07	0.906	−2.61
G13	11-91	Eq. (2)	40	150	50	7.20	0.59	0.999	−0.54
G14	11-91	Back up	50	150	50	7.20	2.29	0.996	0.83
A21	01-91	Eq. (1)	40	150	200	7.50	1.24	0.940	0.22
B21	01-91	Eq. (1)	40	150	200	7.40	1.05	0.940	0.05
G21	11-91	Backup	40	150	200	7.20	0.67	0.994	−0.40
C11	04-91	Back up	4	270	20	7.50	0.02	0.933	-3.76
C12	04-91	Eq. (1)	25	270	20	7.50	0.27	1.000	−1.30
C13	04-91	Eq. (1)	40	270	20	7.50	2.09	1.000	0.73
D11	05-91	Back up	4	270	20	7.45	0.03	0.571	−3.56
D12	05-91	Back up	25	270	20	7.45	0.26	1.000	−1.34
D13	05-91	Back up	40	270	20	7.45	2.02	0.995	0.70
H11	02-92	Back up	4	270	20	7.50	0.02	0.996	−3.79
H12	02-92	Back up	25	270	20	7.50	0.26	0.999	−1.36
H13	02-92	Back up	40	270	20	7.50	2.32	0.999	0.84
C21	04-91	Back up	40	270	50	7.50	2.05	1.000	0.72
D21	05-91	Back up	40	270	50	7.45	1.95	0.998	0.67
H21	02-92	Back up	40	270	50	7.50	2.16	0.990	0.77
C31	04-91	Eq. (1)	40	270	100	7.50	2.04	0.997	0.71
D31	05-91	Back up	40	270	100	7.45	1.94	0.997	0.66
H31	02-92	Back up	40	270	100	7.50	2.24	0.996	0.81
K1	06-94	Back up	25	270	200	7.40	0.24	0.998	−1.43
K2	06-94	Back up	40	270	200	7.40	1.77	0.998	0.57
L1	06-94	Eq. (2)	25	270	200	7.20	0.06	0.870	−2.81
L2	06-94	Eq. (2)	40	270	200	7.20	0.96	0.998	−0.04
N1	03-95	Back up	25	270	500	7.30	0.17	1.000	−1.79
N2	03-95	Eq. (2)	40	270	500	7.30	1.42	0.995	0.35
E11	04-91	Back up	4	320	20	7.50	0.03	0.870	−3.38
E12	04-91	Back up	25	320	20	7.50	0.34	0.991	−1.08
E13	04-91	Eq. (2)	40	320	20	7.50	2.57	0.999	0.94
F11	05-91	Back up	4	320	20	7.40	0.02	0.465	−4.20

Table 1 (continued)

Batch (letter)	Study start	Used for:	Temperature.(°C)	Concentration (mg/ml)	Volume (ml)	pH	k	r^2	In(k)
F12	05-91	Back up	25	320	20	7.40	0.27	0.991	−1.33
F13	05-91	Back up	40	320	20	7.40	2.08	0.997	0.73
I11	02-92	Backup	4	320	20	7.50	0.03	0.917	−3.59
I12	02-92	Back up	25	320	20	7.50	0.30	0.999	−1.20
I13	02-92	Back up	40	320	20	7.50	2.56	0.983	0.94
E21	04-91	Back up	40	320	50	7.50	2.36	0.997	0.86
F21	05-91	Back up	40	320	50	7.40	1.81	0.995	0.59
I21	02-92	Back up	40	320	50	7.50	2.44	0.978	0.89
E31	04-91	Back up	40	320	100	7.50	2.55	0.972	0.94
F31	05-91	Eq. (2)	40	320	100	7.40	2.15	0.994	0.77
I31	02-92	Back up	40	320	100	7.50	2.52	0.977	0.92
J1	11-93	Back up	25	320	200	7.50	0.30	0.998	−1.21
J2	11-93	Eq. (2)	40	320	200	7.50	2.29	0.998	0.83
33	11-93	Back up	50	320	200	7.50	8.01	0.998	2.08
M1	06-94	Back up	25	320	200	7.50	0.40	0.996	−0.91
M2	06-94	Eq. (2)	40	320	200	7.50	2.61	0.998	0.96
M3	06-94	Back up	50	320	200	7.50	10.21	0.999	2.32
O1	03-95	Back up	25	320	500	7.40	0.25	1.000	−1.39
O2	03-95	Eq. (2)	40	320	500	7.40	2.60	0.967	0.96

3. Results and discussion

The data from the three earliest batches were used to develop the first model (calibration). To be able to determine the effect of [H^+], temperature, product concentration and sample volume on ln(k) the four main effects had to be linearly uncorrelated (unconfounded). By performing PLS on the stability data from the first three batches the effects of the four variables were determined (Table 2). The degradation rate expressed as ln(k) increased with an increase in product concentration, container volume, storage temperature and a decrease in [H^+]. However, container volume affected degradation only slightly (Table 2). No significant two-factor interactions were found. PLS was used as it dealt with covariant and noisy data better than OLS (Martens and Næs, 1989; Frank and Friedman, 1993). The small variation in measured [H^+] made it preferable to use PLS for the calibration. The accuracy and precision of the calibration are shown in Fig. 1. From the data obtained (Table 2), the expected worst case batches were chosen and tested. A typical worst case batch is a batch with low [H^+] and/or high product concentration. Measured ln(k) from real time data after storage at 40°C was compared with the predicted ln(k).

Two criteria were chosen for evaluating new batches of product. The first and most important criteria was to test whether the ln(k) at 40°C derived from a new batch was lower than the ln(k) predicted for the worst pH accepted (Table 3). The criterion is based on the assumption that the degradation rate at the most is still within the acceptable range for the product. Analytical measurements should be initialised on back up samples stored at 25°C if the difference between the measured ln(k) and the predicted ln(k) for the worst pH is negative. If the measured k at 25°C suggests a shorter shelf-life than the predicted value, to an extent which takes it outside product specification, action must be taken. Such an observation suggests that factors which have important effects on the stability of the product are inadequately modelled or absent in the model.

The second criterion is based on a comparison of an observed ln(k) for a new batch against the predicted ln(k) and its associated 95% confidence limit (CL) (Test 2 in Table 3). Here the 95% CL is set at twice the standard deviation on either side of the mean. Since no systematic residual bend was seen in the residuals (Fig. 1), such an interval is appropriate. The 95% CL is used to identify whether any significant change in the product occurred rather than to estimate shelf-life. The mechanisms behind any observed significant change may be of importance with respect to the clinical performance, process control or analytical methods. The appropriate action to be taken depends on the specific product concerned. Alternatively, to detect major deviations, outlier testing using a 99.7% limit equivalent to three times

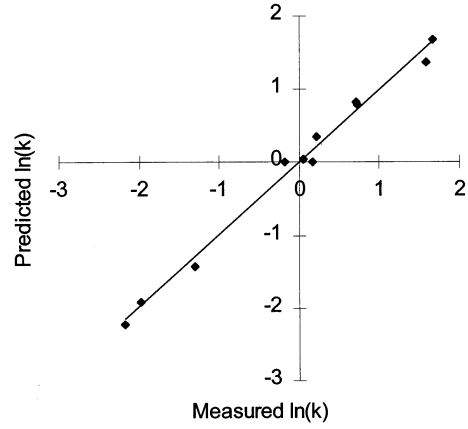

Fig. 1. The performance of Eq. (1) is illustrated. Standard deviation is 0.134 and r^2 is 0.995.

Table 2
The regression coefficients () for Eq. (1) and the upgraded Eq. (1), named Eq. (2), with weighed for ranking of the four variables

Variable	Regression coefficients (effects of variables)		
	Equation 1	Equation 2	Weighed (Eq. (2))
Concentration (g/ml)	0.00410	0.00312	0.289
Volume (ml)	0.00027	0.00041	0.047
H^+ (nM)	−0.0377	−0.337	−0.302
Temperature^{-1}, K^{-1}	−13.77	−14.29	−0.891

standard deviations may be used. When one of the criteria above is not met, the model needs to be updated and improved. Causes of the marked deviations have to be sought by for example exploring the possibility of changes in raw materials. Production process and analytical methods may require re-examination.

To evaluate product stability according to the procedure described, batches of product with different container volumes, product concentrations and pHs were chosen and evaluated. Two batches tested at 40°C, L (L2) and C (C31) fell outside the 95% CL (Test 2 in Table 3). Their corresponding measured ln(k)s at 25°C (L1 and C12) were within the predicted 95% CL. Since both batches also met the criterion 1 when using the worst case pH at the actual temperature (Test 1 in Table 3) the batches were found to have acceptable stability. The ln(k)s derived from various batches (Table 1) were used to validate and upgrade Eq. (1) resulting in a new equation (Eq. (2)) as given in Table 2. As expected, the regression coefficients in the equations were comparable. Eq. (2) was used to predict the ln(k) for 4, 25 and 40°C for all batches. Logarithmic rate constants were transformed to k and a plot of the predicted versus measured k is shown in Fig. 2a. The plot showed that no reaction rates deviated strongly from model predictions. No important factors influencing stability appeared to have been missed as shown by the lack of trends in the residuals from the regression lines (Fig. 1 and Fig. 2a). The slope (1.000) and the small intercept (Fig. 2a) indicated absence of systematic errors.

To further analyse the residuals shown in Fig. 2a, two relationships were evaluated. Firstly, the relationship between the size of k and the difference between measured and those predicted by Eq. (2) and expressed as % residual, was examined (Fig. 2b). Predictions based on low k values led to imprecise predictions.

Secondly, to assess whether the inaccuracy in prediction was related to analytical variation or to systematic error in the multivariate prediction model (Eq. (2)), the residuals were plotted against the correlation coefficients, r^2, from the original univariate linear regression (Table 1Fig. 2c). The figure shows a relationship between the inaccurate prediction of k and the inaccurate k calculated from the univariate regression. The correlation coefficient represents the ratio between sum of squares for the univariate regression and the corresponding sum of squares for the error. When k is low, this ratio and r^2 decrease, i.e. the random variation in the analytical results dominates over the real change in the product during time (decreased signal to noise ratio). The results shown in Fig. 2c suggest that the random variation in the

Table 3
Testing of ln(k) from new batches against ln(k) predicted by Eq. (1) against criterion 1 and 2

Batch	Predicted ln(k)	Measured ln(k)	Criterion 1		Criterion 2	
			WC[a]	test 1[b]	95% CL[c]	test 2[d]
C31	−0.88	−0.54	1.07	1.60	−1.15,-0.61	−0.07
L2	−0.34	−0.04	1.09	1.13	−0.38,–0.07	−0.03
L1	−2.56	−2.81	−1.12	1.69	−2.83,–2.29	+0.02
N2	0.23	0.35	1.17	0.82	−0.04,0.50	+0.15
E13	1.00	0.94	1.25	0.30	0.73,1.27	+0.21
F31	0.71	0.77	1.27	0.50	0.44,0.98	+0.21
J2	1.05	0.83	1.30	0.47	0.78,1.32	+0.05
M2	1.05	0.96	1.30	0.34	0.78,1.32	+0.18
O2	0.82	0.96	1.42	0.46	0.55,1.09	+0.13

[a] Prediction of ln(k) for the worst case pH (7.6).
[b] Difference between measured ln(k) and ln(k) predicted for the worst case pH.
[c] 95% CL defined as predicted ln(k) ± twice standard deviation (0.27) gained from Eq. (1).
[d] Negative values is equal to significant deviation of measured ln(k) from predicted ln(k).

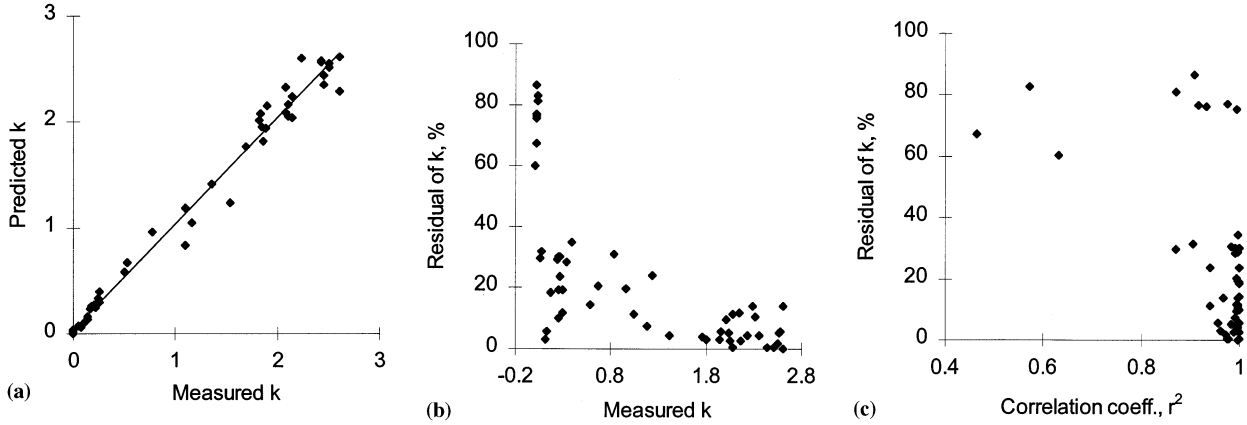

Fig. 2. (a) The correlation of predicted and measured k by Eq. (2) shows that no critical failure estimations have been performed. The correlation coefficient, r^2, the slope and the intercept is 0.985, 1.000 and 0.039, respectively. (b). The % residual derived from measured k and k predicted by multivariate model versus measured k illustrates that low k result in uncertain model prediction. (c). The % residual of k derived from the multivariate model is plotted against the correlation coefficient, r^2, from original linear regression of each reaction rate k. The plot shows increased probability of inaccurate model predictions of k for low r^2.

analytical results was the primary reason for the observed deviations. The results suggest that stress conditions will emphasise batch differences more accurately than normal storage conditions. Therefore, the stability data from 4°C (Table 1) were not used for calibration, validation and batch to batch testing. Instead data from higher temperature studies are preferable. However, such should be thoroughly evaluated before use in predicting differences between batches at room temperature. Higher temperatures may initiate irrelevant chemical/physical reactions not observed at room temperature.

3.1. The proposed OST for evaluating a solution filled in glass bottles

3.1.1. Step 1. Establish model

A. Calculate the degradation rate constants, k, for each stability indicating parameter using three batches product according to ICH guidelines and covering minimum three temperatures typically within the range of 25–50°C. Evaluate the estimated k using the correlation coefficient and the significance testing from each of the univariate linear regressions. Evaluate residual structure for systematic errors.

B. Determine the effect of temperature and product characteristics, such as pH, on the degradation rate constant by constructing a multivariate model by primarily using PLS regression or secondary OLS regression. Use other scientific information/data gained from degradation, prestability and formulation studies to increase the robustness and reliability of the model. Evaluate the performance of the multivariate regression by the correlation coefficient, intercept and slope. A leverage test may also be adequate in order to check if some individual results influence the regression coefficients strongly.

3.1.2. Step 2. Testing a new variable

A. When the effect of a new variable such as volume is to be tested, store three batches under normal and stressed stability conditions according to the ICH guideline.

B. Analyse the two expected worst case batches, one stored at 25 and 40°C and the other at 40°C after 3 or 6 months. Test whether the stability of the product using the new volume is different to that with the old volume according to:

Criterion 1. Whether the measured reaction rate constant exceeds the predicted worst case reaction.

Criterion 2. Whether the measured reaction constant exceeds the 95% confidence limit for the predicted degradation constant.

If there is no difference based on both criteria one may exclude testing of volume effects or only undertake volume reduced testing under stress conditions.

3.1.3. Step 3. Upgrade model

A. If the volume is different, upgrade the first model with results at 25 and 40°C for the first batch. Then test whether the second batch meets criteria 1 and 2.

B. If the second batch meets criteria 1 and 2 then only test the new volumes at stress temperatures in future. If the second batch yields different, then initiate testing of all three batches at all normal storage conditions. Look for new factors that influence the product stability of these batches and take appropriate action

3.1.4. Step 4. Further evaluation of new variables by repeating step 2 and 3

Repeat step 2 and 3 for each new variable such as minor formulation changes, raw materials, product concentration, packaging, etc. Generally, more extensive testing has to be initiated for major formulation changes.

The criteria and stress conditions to be chosen for assessing degradation rates for new batches by OST should be based on the intended market for the pharmaceutical product. Typically, different climatic zones will result in various needs depending on the type of packaging (glass, plastic, other) and type of product (solid, solutions, emulsions, suspensions, etc.). Products, whose stability change markedly with small changes in storage conditions, may not be suitable for OST. For systems that may brake down (emulsions, suspensions, etc.) the storage time to brake down may be the response of interest instead of the reaction rate.

4. Conclusion

The classical statistical approach for batch testing is often nonoptimal although the predicted shelf-life for new batches is usually satisfactory. This may be more related to the improved knowledge about the product during its lifetime, than to statistical prediction made on the basis of the data from three formal batches of product.

The proposed opportunistic stability testing scheme suggested in this paper may be an improvement on current approaches. Of the 15 batches tested extensively at all formal conditions during a 5 year period, nine batches were tested by the OST method. Since the chemical analyses required for testing at 25°C is more extensive than at 40°C, a significant reduction in analytical measurements can be achieved using OST.

Acknowledgements

The authors are grateful for support and advise from Kari Dyvik and Anne Matthiasen of the pharmaceutical sciences department and Janne Veggeland of the stability Department.

References

Aars, E.V., Eivindvik, K., 1995. Formulation, stability and compatibility of iodixanol. Acta Radiol. 36 (Suppl. 399), 50–58.

Carstensen, J.T., Franchini, M., Ertel, K., 1992. Statistical approach to stability protocol design. J. Pharm. Sci. 81, 303–308.

Ertel, K.D., Carstensen, J.T., 1990. Examination of a modified Arrhenius relationship for pharmaceutical stability prediction. Int. J. Pharm. 61, 9–12.

Fairweather, W.R., Lin, T.D., Kelly, R., 1995. Regulatory, design, and analysis aspects of complex stability studies. J. Pharm. Sci. 84, 1322–1326.

Frank, I.E., Friedman, J.H., 1993. A statistical view of some chemometrics regression tools. Technometrics 35, 109–135.

Helboe, P., 1992. New design for stability testing programs: matrix or factorial designs. Authorities viewpoint on the predictive value such studies. Drug Inf. J. 26, 629–634.

Martens, H., Næs, T., 1989. Multivariate Calibration. Wiley and Sons, New York.

Ruberg, S.J., Hsu, J.C., 1991. Multiple comparison procedures for pooling batches in stability studies. Technometrics 34, 465–472.

Su, X.Y., Po, A.L.W., Yoshioka, S., 1994. A bayesian approach to Arrhenius prediction of shelf life. Pharm. Res. 11, 1462–1466.

Yoshika, S., Aso, Y., Kojima, K., 1996. Statistical evaluation of shelf life of pharmaceutical products estimated matrixing. Drug Stab. 1, 147–151.