Journal of Thermal Analysis, Vol. 41 (1994) 713-724

General

In celebration of the 60th birthday of Dr. Andrew K. Galwey

PREDICTING SHELF-LIVES OF PHARMACEUTICAL PRODUCTS
Monte Carlo simulation using the simulation package SIMAN

X. Y. Su[1], *A. Li Wan Po*[1]* *and S. Yoshioka*[2]

[1]Drug Delivery Research Group, The School of Pharmacy,
The Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland
[2]National Institute of Hygienic Sciences, 1-18-1, Kamiyoga Setagaya, Tokyo 158, Japan

Abstract

The computer simulation package SIMAN® is used to carry out Monte Carlo simulations with a view to estimating and predicting the shelf-lives of pharmaceutical products. The input data take account of assay variance and low extents of decomposition typical of data sets submitted by pharmaceutical companies in support of their product licence applications. It is shown that Monte Carlo methods provide estimates of the predicted shelf-life with a narrower and more symmetrical distribution than obtainable with integral methods. The results indicate that the median may be a more reliable estimate of shelf-life than the mean particularly if assay variance is high. Despite its usefulness SIMAN is a difficult package to use and is not generally recommended.

Keywords: pharmaceutical products, SIMAN program

Introduction

Pharmaceutical manufacturers submitting applications for products to be licensed are required to provide data in support of product stability during the projected shelf-lives under the intended storage conditions. Typically, in tem-

* To whom correspondence should be addressed.

John Wiley & Sons, Limited, Chichester
Akadémiai Kiadó, Budapest

perate climates, the stability profile is such that a shelf-life of three years at 25°C is sought.

At the time of submission of their applications, manufacturers usually would have six months stability data including some generated under accelerated conditions. Both isothermal and non-isothermal [1] methods may be used. On the basis of those data, estimated shelf-lives for the products, stored under ambient storage conditions, would be proposed.

There are several difficulties with making such predictions. If the product being investigated is to be stable for the projected 3 year shelf-life, then, at the time of submission of the application, only sparse data is available for analysis. This sparsity of data leads to two further difficulties. Firstly, it is not possible to determine the order of reaction reliably and hence the model required for making the shelf-life predictions. Secondly, the low extent of decomposition means that the accuracy of the estimates for drug content is limited by the precision of the assay. The maximum extent of decomposition is, at that stage, often of the same order of magnitude as the assay coefficient of variation.

Those difficulties have led several authors to investigate the best method for handling the sparse data using Monte Carlo simulations [2, 3]. Such simulations are made relatively simple with the availability of statistical packages which enable sampling from various probability distributions.

Theoretical background

If drug decomposition follows zero-order kinetics, the residual amount of drug remaining (D) in the dosage form, as a function of time (t), can be represented as

$$D = D_0 - kt \tag{1}$$

where D_0 is the initial drug content and k is the zero-order rate constant. Hydrolytic reactions in suspension systems often follow such kinetics.

Assuming that the Arrhenius equation holds, the rate constant k can be written as

$$k = A \exp [-E_a/(RT)] \tag{2}$$

where A is the pre-exponential factor, E_a the activation energy, R the gas constant and T the storage temperature (K).

By definition, the shelf-life is the time (t_{90}) taken for the original drug content to drop to 0.9 D_0, unless toxic decomposition products are formed.

A can be expressed in terms of the t_{90} at room temperature (25°C) as

$$A = 0.1\ D_o \exp\left[E_a/(298R)\right]/t_{90(25)} \tag{3}$$

Substitution into Eqs (1) and (2) leads to the following relationship

$$t_{90(25)} = 0.1\ D_o\ t_{D(40)} \exp\left[(E_a/R)\ (1/298 - 1/T)\right]/(D_o - D) \tag{4}$$

$t_{D(40)}$ is the sampling time at which the drug content D is observed in a sample stored at 40°C.

If one assumes that the assay used in estimating the residual drug concentration has a variance of σ^2, then the probability density function (p.d.f.) for the concentration D, with mean μ, can be described by

$$p(D) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(D-\mu)^2}{2\sigma^2}\right] \tag{5}$$

Now, typically in stability testing, at the time of product licence application, data at elevated temperatures (e.g. 40°C) are used to estimate the shelf-life at room temperature (25°C).

If one assumes zero-order kinetics at both temperatures, then use of the Arrhenius equation enables the probability density function for D to be expressed in terms of t_{90} at 25°C ($t_{90(25)}$).

$$p(D) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(100 - B_o/t_{90(25)} - D)^2}{2\sigma^2}\right] \tag{6}$$

where

$$B_o = 10 t_{D(40)}\left[\exp\left[-\frac{E_a}{R}\left(\frac{1}{313} - \frac{1}{298}\right)\right]\right] \tag{7}$$

The corresponding p.d.f. for $t_{90(25)}$ can be written, as shown by Yoshioka *et al.* [4], as

$$p(t_{90(25)}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(100 - B_o/t_{90(25)} - D_i)^2}{2\sigma^2}\right] \frac{B_o}{(t_{90(25)})^2} \tag{8}$$

where D_i is now a specific observation of a random variable D representing the drug content.

The corresponding p.d.f. for different reaction orders are as follows

First-order

$$p(t_{90(25)}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{[100 \exp(-B_1/t_{90(25)}) - D_i]^2}{2\sigma^2}\right] \cdot 100 \exp(-B_1/t_{90(25)}) \frac{B_1}{(t_{90(25)})^2} \quad (9)$$

where $B_1 = 0.1054 \exp[(E_a/R)(1/298 - 1/313)t_{D(40)}$

$$D = D_o \exp\left[t_{D(40)}\left(\frac{0.1054}{t_{90(25)}}\right)\exp\left[\frac{E_a}{R}\left(\frac{1}{298} - \frac{1}{313}\right)\right]\right] \quad (10)$$

Second-order

$$p(t_{90(25)}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{[100/(1 + B_2/t_{90(25)}) - D_i]^2}{2\sigma^2}\right] \cdot 100\, B_2/(t_{90(25)} + B_2)^2 \quad (11)$$

where $B_2 = 0.1111 \exp[(E_a/R)(1/298 - 1/313)]t_{D(40)}$

Those distributions can therefore be used to assess the effect of assay variance on the reliability of the shelf-life estimates. We did this using Monte Carlo methods to generate data which were then fitted by both linear and non-linear regression. We assumed assay standard deviations ranging from 1 to 5% and extents of decomposition at the time of assay in samples stored at 40°C ranging from 75 to 90% of the original content.

SIMAN - Simulation language

In a SIMAN program (Systems Modelling Corporation, Sewickley, PA, USA, 1989), the user is required to define a MODEL frame, an EXPERIMENT frame, a LAYOUT frame and a STATE subroutine.

The MODEL frame describes the components of the system and their interactions. In our study, the MODEL frame generates the error terms drawn from a normal distribution with a mean of zero and a standard deviation estimated for our assay procedure (range 1 to 5%). This error is then added to the errorless residual drug content (range 75 to 90%) to provide random observations of the latter.

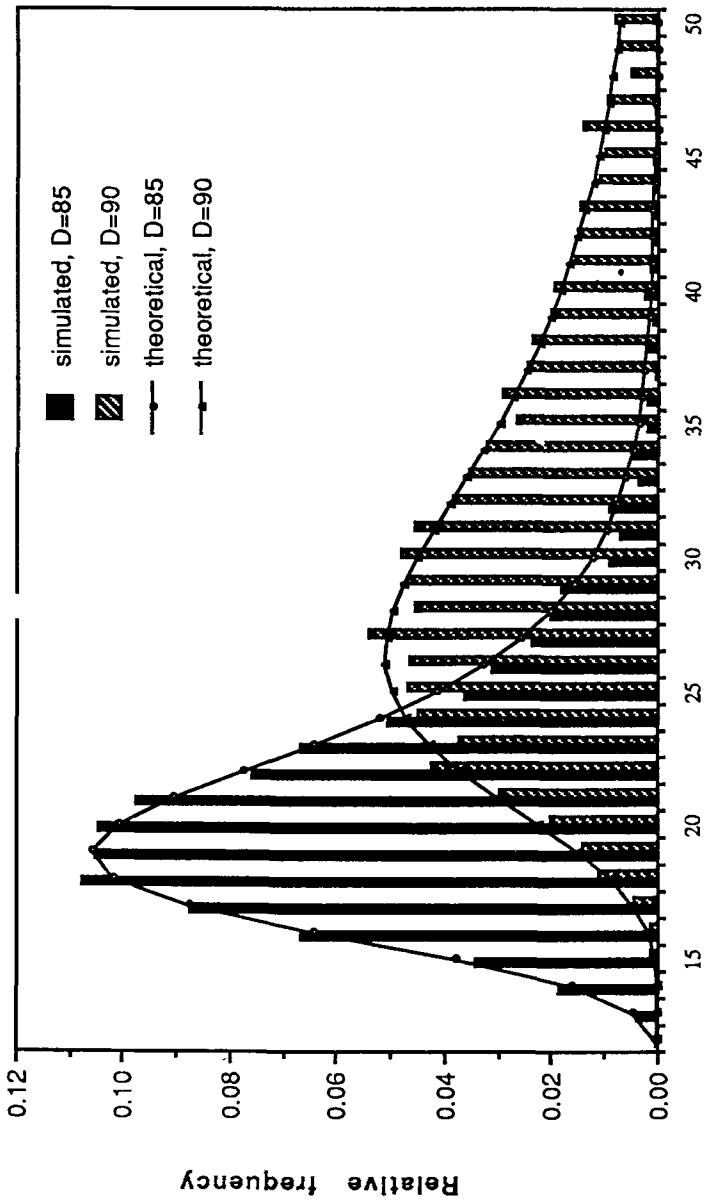

Fig. 1 Comparison of the distribution of the shelf-life ($t_{90(25)}$) generated using Monte Carlo methods (histograms) against the theoretical distribution (continuous line). Assay standard deviation is taken as 3%

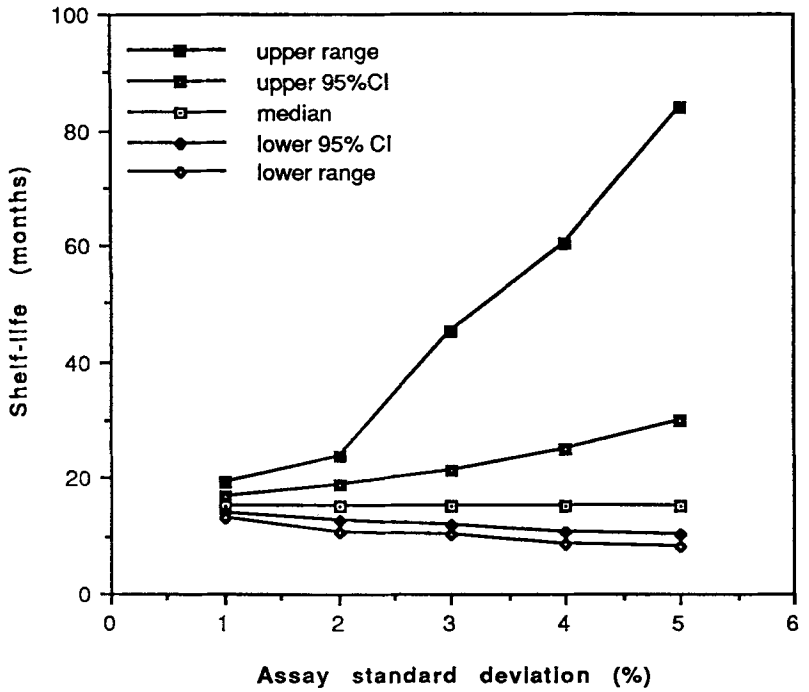

Fig. 2 Effect of the standard deviation of assay shelf-life ($t_{90(25)}$). $D = 80$, see text for calculation

The EXPERIMENT frame defines the experimental conditions, such as run length and initial conditions, and saves the simulated observations. In the present study, 50 replications starting from random integer seeds were generated.

The STATE subroutine is embedded in Fortran within a Fortran module called NAME·FOR. The subroutine defines the state equations (the integral equations) defining the model. If differential equations are given, then numerical integration is performed using Runge-Kutta's algorithm. Two approaches were used in the present study. Firstly, from the random observations of residual drug content, the STATE subroutine performs a simple algebraic substitution to calculate random $t_{90(25)}$ values using Eq. (4), the state equation. The second approach includes Eq. (8) in the STATE subroutine. For any fixed residual drug content D, integration of Eq. (8) with respect to $t_{90(25)}$ yields the cumulative density function for the shelf-life ($t_{90(25)}$).

The theoretical probability distribution for $t_{90(25)}$ can be plotted using Eq. (8).

Prior to running the program, the MODEL frame and the EXPERIMENT frames are compiled and linked. The Fortran file including the STATE subroutine is also compiled and the program run.

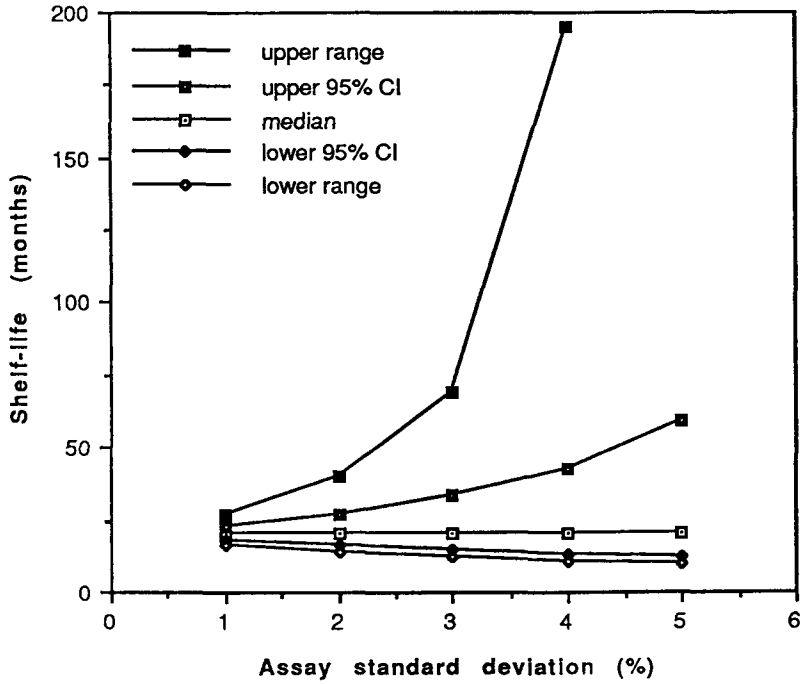

Fig. 3 Effect of the standard deviation of assay on shelf-life ($t_{90(25)}$). $D = 85$, see text for calculation

Results and discussion

Figure 1 shows probability distributions of the product shelf-lives using both Monte Carlo (histograms) simulation and deterministic modelling (continuous lines). The simulation assumed an assay standard deviation of 3% and two residual drug contents of 85 and 90%. The dramatic effects which the drug content at the sampling point has on the reliability of the shelf-life estimates is obvious from the diagram. The concordance in the results using the stochastic and deterministic approaches validates the underlying models used.

Using the approaches described, it is easy to compute the 95% confidence interval and other relevant statistics. We did this using a range of assay precisions (1 to 5%) at two residual drug contents (80 and 85%) and the results are shown in Figs 2 and 3. The non-symmetrical nature of the confidence interval is shown in the Figs 2 and 3.

Given the non-symmetrical distribution of the shelf-life ($t_{90(25)}$), the mean is not the most appropriate measure of central tendency. We therefore computed the median and its associated confidence interval using the Monte Carlo simulated data. The intervals are compared with those for the mean shelf-lives in Fig 4. Increasing assay variance amplifies the non-symmetry of the distribution, suggesting that the median shelf-life would be a more appropriate statistic for defining shelf-life estimates. Indeed, the relative insensitivity of the median to assay variance can be seen in Fig. 4.

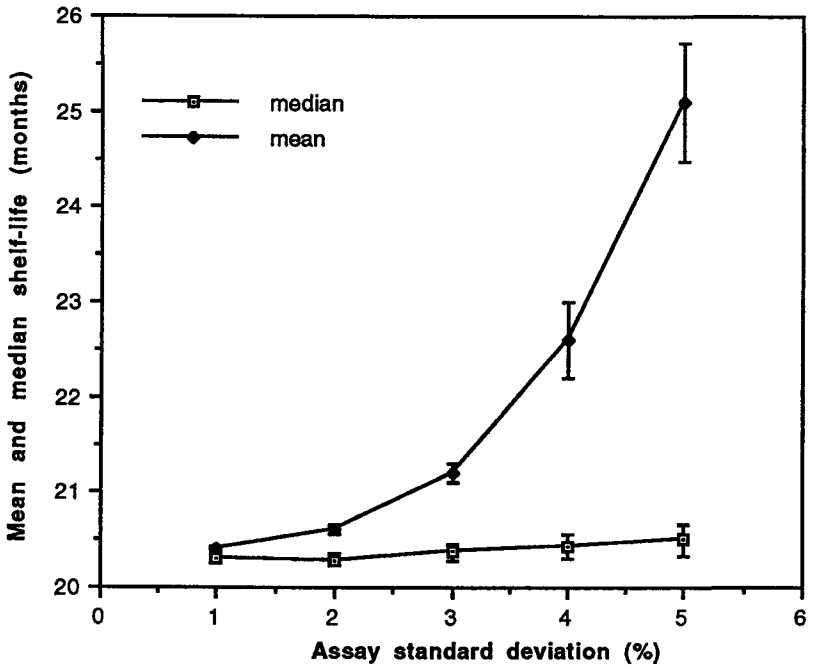

Fig. 4 Effect of assay standard deviation on mean and median (error bars refer to standard error $n = 100$), $D = 85$

Figure 5 further illustrates the differences in the mean and median estimates. The higher the variance, the greater the degree of non-symmetry when the mean shelf-life is used. The $t_{90(25)}$ distribution is skewed to the right, with approximately 60% of the observations lying between the lower 95% confidence interval and the mean when the standard deviation was 5%.

Experimental data

To evaluate the performance of our simulations, the experimental data reported by Slater *et al.* [5] on the stability of vitamin *A* were used. Those authors

reported a first-order rate constant (mean±95% CI) of 0.00729±0.00232 weeks^{-1} for vitamin *A* decomposition at 40°C (Product F Table 1). Using those statistics, we simulated 500 observations for the rate constant and, for each value, we calculated a corresponding *D* value at 16 weeks, as used by Slater *et al.* [5]. The 500 sample points gave a residual drug content (mean±*SD*) of 87.70±3.85%. Using an activation energy, E_a, of 26.07 kcal/mol (109.1 kJ·mol^{-1}), the corresponding shelf-life ($t_{90(25)}$) can be calculated, as previously described, using either Monte Carlo simulation or deterministic modelling. The activation energy was calculated using Eq. (10). The associated 95% confidence interval was then calculated by integrating the probability distribution or by excluding the extreme 5% observations. The results are shown in Table 1. For comparison, we also quote results obtained by King *et al.* [2] using classical linear regression and non-linear regression for estimating the shelf-lives. We can see that the Monte Carlo sampling method described in this study produces results which compare well with the non-linear regression method reported by King *et al.* [2]. The confidence interval is narrower and symmetrical.

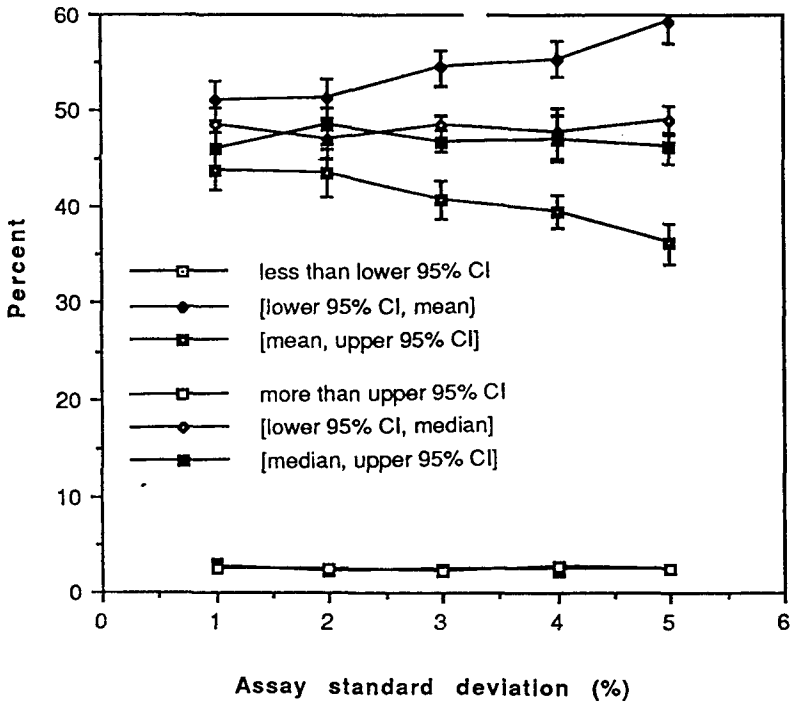

Fig. 5 Percentage of simulated shelf-lives ($t_{90(25)}$) in various intervals within the probability distribution. The error bars are corresponding standard deviations ($n = 20$) sets

Table 1 Comparison of different methods for predicting room temperature (25°C) shelf-lives from accelerated (40°C) data

Method	$t_{90(25)}$ / week	
	Mean	95% CI
Classical linear regression[a]	113.3	49.3–262.2
Nonlinear regression[a]	109.0	61.7–156.2
Monte Carlo sampling[b]	122.1	89.5–154.6
Integration of probability distribution[c]	106.2	62.9–286.0

[a]These values were obtained from Table 10 in King *et al.* [2]
[b]These values were calculated by Monte Carlo simulation using data from Slater et al. [5]
[c]These values were obtained by the integration of Eq. (9)

SIMAN

The graphical facilities offered by SIMAN are shown by Figs 6 and 7. The latter shows the probability distributions obtained for $t_{90(25)}$ using assay standard deviations ranging from 1 to 5%, given a fixed (20%) extent of decomposition, while Fig. 6 shows fifty simulated $t_{90(25)}$ values using Monte Carlo simulation, again starting with 80% residual drug content, but with a fixed assay standard deviation of 2%.

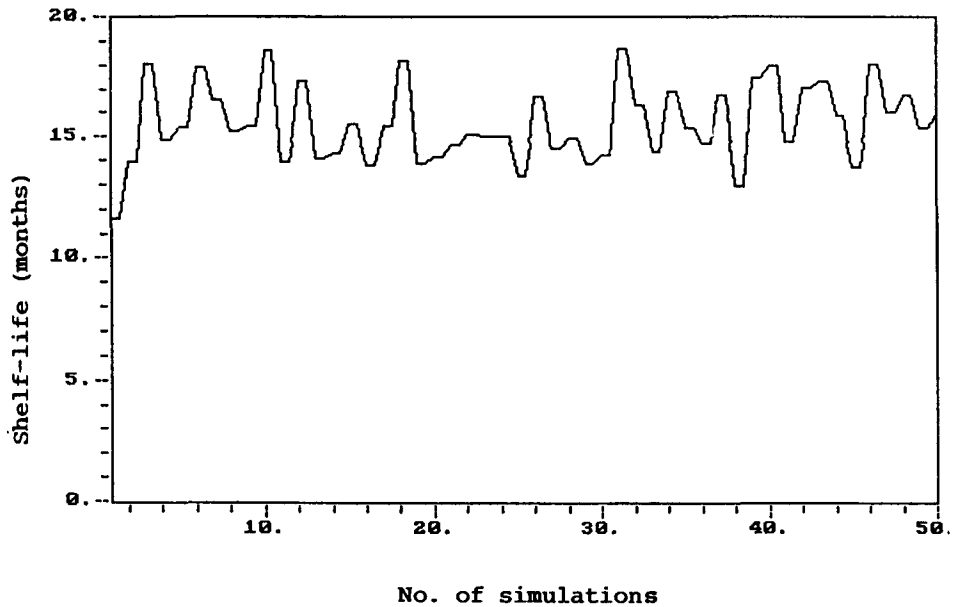

Fig. 6 Graphical display of simulated shelf-life using SIMAN (D = 80, SD = 2%)

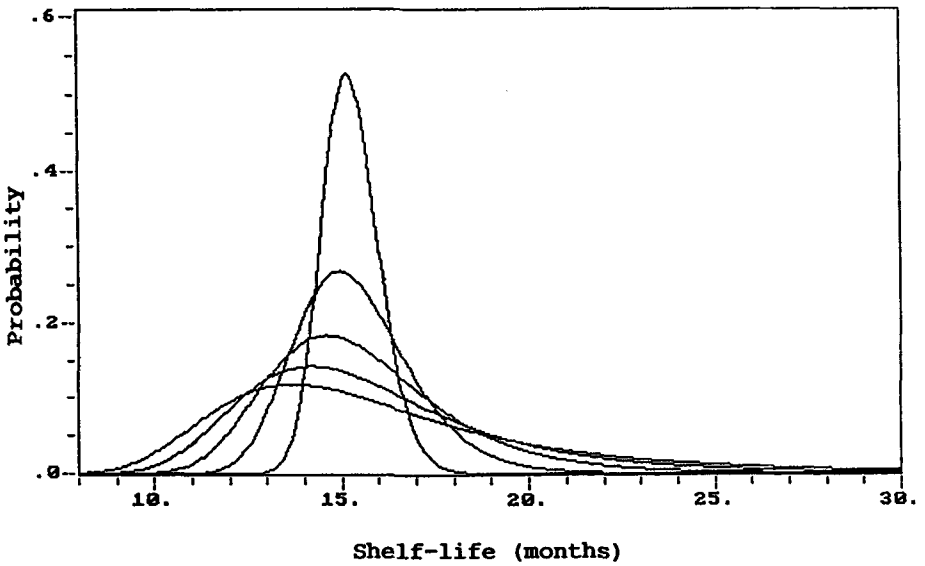

Fig. 7 Graphical display of probability distribution of shelf-life (D = 80, SD = 1 to 5%). The larger the SD the broader the distribution

Conclusion

The use of SIMAN for predicting the shelf-life at room temperature from accelerated stability data has been demonstrated. The package permits both stochastic and deterministic modelling and can handle both continuous and discrete models. Using SIMAN and both simulated and literature data, it is shown that with Monte Carlo methods the predicted shelf-life has a narrower and more symmetrical distribution than can be obtained with integral methods. The results were similar to those obtained by King *et al.* [2] using non-linear estimation. The data suggest that the median may be a more reliable estimate of shelf-life, particularly if assay variance is high. With Monte Carlo methods and SIMAN the approximate 95% confidence interval can readily be estimated.

Despite its usefulness, SIMAN suffers from extremely poor use-friendliness and is expensive. A Fortran compiler is required to run the programme. It is not a package to be recommended to novices in computing.

References

1 A. Li Wan Po, N. A. Elias and W. J. Irwin, Acta Pharm. Suecica., 20 (1983) 277.
2 S. P. King, M. Kung and H. Kung, J. Pharm. Sci., 73 (1984) 657.
3 S. Yoshioka, Y. Aso and M. Uchiyama, J. Pharm. Sci., 76 (1987) 794.

4 S. Yoshioka, Y. Aso and Y. Takeda, Chem. Pharm. Bull., 38 (1990) 1760.
5 J. G. Slater, H. A. Stone, B. T. Palermo and R. N. Duvall, J. Pharm. Sci., 68 (1979) 49.

Zusammenfassung — Das Computer-Simulationsprogramm SIMAN® wurde für Monte-Carlo-Simulationen verwendet, um die Lagerfähigkeit von pharmazeutischen Produkten zu schätzen und vorherzusagen. Die Input-Angaben berücksichtigen die Probenstreuung und niedrige Zersetzungsmengen, typisch für die von Pharmafirmen beim Antrag auf Produktezulassungen eingereichten Angaben. Es wurde gezeigt, daß Monte Carlo Methoden Schätzungen der vorausgesagten Lagerfähigkeit mit einer schmaleren und symmetrischeren Verteilung liefern als dies mittels Integralverfahren möglich Fall ist. Die Resultate zeigen, daß der Zentralwert gerade bei großer Probenstreuung eine bessere Schätzung der Lagerfähigkeit abgeben kann als der Mittelwert. Trotz des Nutzens ist SIMAN ein schwierig zu handhabendes Softwarepaket und wird nicht generell empfohlen.