Statistical Prediction of Drug Stability Based on Nonlinear Parameter Estimation

SHANG-YING P. KING *, MIN-SHYA KUNG ‡, and HO-LEUNG FUNG *x

Received November 15, 1982, from the *Department of Pharmaceutics and ‡Department of Statistics, State University of New York at Buffalo, Amherst, NY 14260. Accepted for publication April 28, 1983.

Abstract □ The classical approach in Arrhenius prediction of drug stability uses two sequential steps of linear regression involving (*a*) a function of drug content *versus* time to obtain the rate constants (*k*) at several elevated temperatures and (*b*) the relationship of logarithm of mean *k versus* reciprocal temperature to predict the room temperature rate constant and hence the shelf-life of the drug. Uncertainties in drug content determinations are often neglected in the second regression. The classical approach also provides a wide and unsymmetrical 95% confidence interval for the predicted shelf-life. We have developed equations which allow for direct statistical prediction of shelf-life using observed values of drug content, time, and temperature. Nonlinear regression analysis was employed to provide parameter estimates of drug shelf-life and the energy of activation. The developed approach was shown to provide good estimates of shelf-life with meaningful statistics of reactions over a wide range of stability and energetics, with various kinetic orders, with different levels of noise in the data, and with different types of data structure. Comparison between the nonlinear approach and the classical approach showed that the nonlinear approach provided better mean estimates of shelf-life with much smaller and more symmetrical 95% confidence intervals than the classical approach. The method appears sufficiently robust and wide-ranging as to be potentially applicable for the prediction of the drug stability of pharmaceutical products.

Keyphrases □ Drug stability—statistical prediction of shelf-life, nonlinear parameter estimation, comparison with classical Arrhenius prediction □ Nonlinear parameter estimation—drug stability, statistical prediction of shelf-life, comparison with classical Arrhenius prediction □ Shelf-life—statistical prediction of drug stability, nonlinear parameter estimation, comparison with classical Arrhenius prediction

The establishment of the projected expiration date (or shelf-life) of a pharmaceutical product is of importance to the pharmaceutical industry. Accelerated stability testing using the Arrhenius equation is often employed to treat stability data obtained under elevated temperature storage conditions for the prediction of the shelf-life of drug products. The classical approach consists of two sequential steps which include (*a*) the choice of a proper order for the degradation reactions to describe the drug content (amount or concentration) *versus* time relationship for determining rate constants at each elevated temperature and (*b*) the application of the Arrhenius equation to predict (or extrapolate) the room temperature rate constant, and hence the shelf-life of the drug, through the use of the mean rate constants obtained at several elevated temperatures. Linear regression analysis (1) is the statistical method employed in each of these two steps.

Although this classical linear approach has been used quite frequently, the method of data treatment appears to suffer from two statistical problems. First, since Arrhenius treatment often employs only the mean rate constants obtained at elevated temperatures to predict drug shelf-life, errors associated with the determination of drug content are not included (and thus they are often ignored) in the analysis of accelerated stability data. Second, the 95% confidence interval of the predicted room temperature rate constant derived from the confidence band of the best-fitted straight line may be so wide as to make the predicted shelf-life of little value (2). Because of these inherent problems in the classical approach, it seems desirable to develop an alternate approach for the treatment of accelerated stability data so as to provide an improved prediction of drug shelf-life.

This study was initiated to develop and test mathematical relationships among drug content, time, and temperature, which take into account the errors in the drug content data for the prediction of shelf-life of drug products. A one-step nonlinear approach is proposed for the treatment of accelerated stability data. Comparison of this approach with the classical method is reported.

BACKGROUND

The shelf-life or expiration date of a drug product is defined as the time period required for a drug to decompose to a specified fraction of the labeled drug content, usually 90%. This shelf-life may be determined by directly monitoring the drug content in the drug product as a function of time at some ambient storage conditions, *e.g.*, 25°C and certain humidity and light levels. The proper order for the degradation reactions was determined usually by examining the best fit of plots of drug content *versus* time (for a zero-order reaction) or of the logarithm of drug content *versus* time (for a first-order reaction). Linear regression analysis may then be used to estimate the rate constant at the specified storage conditions. The time for the lower 95% confidence limit curve about the fitted straight line to reach 90% of the labeled drug content is assigned as the shelf-life of the pharmaceutical product (3, 4).

Although the shelf-life of the drug product can be accurately determined by the previously described method, it is usually a time-consuming process. To obtain a predicted shelf-life in a shorter period of time, accelerated stability testing is employed. Most frequently, accelerated stability testing is coupled with the use of the Arrhenius equation for the prediction of drug stability (5, 6). This approach involves two steps. First, a correct order of degradation reaction is determined to describe the functional relationship between drug content and time at several elevated temperatures. The equations for zero-, first-, and second-order degradation reactions are respectively:

$$C = C_0 - kt \qquad \text{(Eq. 1)}$$

$$C = C_0 \exp(-kt) \qquad \text{(Eq. 2)}$$

$$\frac{1}{C} - \frac{1}{C_0} = kt \qquad \text{(Eq. 3)}$$

where k is the rate constant. C_0 and C are the drug contents at time zero and time t, respectively. Both reactants are assumed to have the same initial contents for the second-order reaction. The equation for the first-order reaction can also be expressed in a linearized form:

$$\ln C = \ln C_0 - kt \qquad \text{(Eq. 4)}$$

The linear regression method may then be used to obtain the rate constant at each elevated temperature by using Eq. 1, 4, or 3 for each order of reaction.

In the second step, the Arrhenius equation (Eq. 5 or 6) is employed to relate the logarithm of mean k linearly to the reciprocal of the elevated temperature, T.

$$k = A \exp\left(\frac{-E_a}{RT}\right) \qquad \text{(Eq. 5)}$$

$$\ln k = \ln A - \frac{E_a}{RT} \qquad \text{(Eq. 6)}$$

where R is the gas constant.

In employing the Arrhenius approach, several assumptions are implicit:

0022-3549/84/0500-0657$01.00/0

(*a*) that the kinetic model is valid, (*b*) that analytical accuracy is not compromised over the time of the stability study, and (*c*) that the Arrhenius equation is valid, *i.e.*, linear extrapolation beyond the observed data is feasible. In practice, after obtaining the estimates of the energy of activation (E_a) and frequency factor (A) through the use of Eq. 6 *via* the least-squares technique, the rate constant at room temperature and, subsequently, the shelf-life can then be predicted. The 95% confidence interval of the predicted room temperature rate constant can be derived from the confidence band of the regressed straight line (2, 7). The linearity of the Arrhenius plot can be checked statistically using the F ratio test (8, 9).

Besides the classical linear approach described, other methods using the Arrhenius relationship were also reported for stability prediction. Lordi and Scott (10) designed stability charts to facilitate the analysis of accelerated stability data. Kirkwood (11) suggested a maximum likelihood nonlinear approach for the prediction of the shelf-life of biological products using all the available drug potency *versus* time data. Nonisothermal kinetic methods (12–18) have also been developed for stability prediction.

Carstensen and Su (7) suggested a nonlinear approach for estimating the parameters of E_a and A in the Arrhenius equation (Eq. 5). The equations were derived by substituting the Arrhenius equation (Eq. 5) into either the equation of a zero-order reaction (Eq. 1) or that of a first-order reaction (Eq. 2). The form of the zero-order equation (Eq. 7) and first-order equation (Eq. 8) were, respectively:

$$C = C_0 - tA \exp\left(\frac{-E_a}{RT}\right) \quad \text{(Eq. 7)}$$

$$C = C_0 \exp\left[-tA \exp\left(\frac{-E_a}{RT}\right)\right] \quad \text{(Eq. 8)}$$

The Gauss-Newton iterative technique was used to obtain the parameter estimates of E_a and A. However, a digital computer was not used and the computational steps associated with the estimation of A and E_a may be too tedious.

Recently, Scher (19) proposed a method of kinetic ratio parameter for product stability calculation. Both Eqs. 7 and 8 were used to derive equations for the calculation of the shelf-life (t_{90}) of the drug. Equations 9 and 10 were employed for the zero-order and first-order reactions, respectively:

$$t_{90} = (C_0 - C) \Big/ \left[A \exp\left(\frac{-E_a}{RT}\right)\right] \quad \text{(Eq. 9)}$$

$$t_{90} = \ln (C_0/C) \Big/ \left[A \exp\left(\frac{-E_a}{RT}\right)\right] \quad \text{(Eq. 10)}$$

Davies et al. (20, 21) suggested the use of the logarithmic forms of Eqs. 8 and 10 to calculate drug shelf-life. The statistical properties of the parameter estimates of E_a and A in the equation for the first-order reaction (Eq. 8) were evaluated through the use of the Monte Carlo method of computer simulations (22). The simulation results showed that estimates of E_a and A were robust, and meaningful statistics of these two parameter estimates could be obtained. The relative standard deviation of the frequency factor (A) was large in comparison with the prescribed standard deviation of normal distribution for generating random errors in the data. After obtaining estimates of A and E_a, the shelf-life (t_{90}) may then be estimated using Eq. 10. The "approximate" standard deviation of the shelf-life (t_{90}) may be calculated by using the standard Taylor expansion (23) expressed as:

$$(SD_{t_{90}})^2 = (t_{90,A} SD_A)^2 + (t_{90,E_a} SD_{E_a})^2 + 2(t_{90,A} t_{90,E_a} \mathrm{COV}_{E_a,A}) \quad \text{(Eq. 11)}$$

where $t_{90,A}$ and t_{90,E_a} are, respectively, the partial derivatives of t_{90} with respect to A and E_a; SD_A and SD_{E_a} are the standard deviations of A and E_a, respectively; and $\mathrm{COV}_{E_a,A}$ is the covariance of E_a and A. No simulation was attempted to evaluate the statistical properties of the predicted shelf-life (t_{90}) as estimated by this nonlinear approach.

In this report, the nonlinear approach was further modified to facilitate a direct prediction of drug stability using observed values of drug content, time, and temperature. Extensive computer simulations were made to evaluate the statistical properties of the parameter estimates of E_a and t_{90}. The applicability and versatility of this nonlinear approach were evaluated as well.

THEORETICAL

The assumptions made in the subsequent mathematical derivations are that (*a*) the energy of activation (E_a) is not a function of temperature and (*b*) normal storage condition at room temperature is set at 25°C. Since a direct prediction of the shelf-life of the drug at 25°C is desired, modification of the Arrhenius equation appears necessary. The degradation rate constant at 25°C (k_{298}) can be obtained from the Arrhenius equation:

$$k_{298} = A \exp\left(-\frac{E_a}{R298}\right) \quad \text{(Eq. 12)}$$

Rearrangement of Eq. 12 yields:

$$A = k_{298} \exp\left(\frac{E_a}{R298}\right) \quad \text{(Eq. 13)}$$

Substituting Eq. 13 for A into the Arrhenius equation (Eq. 5) yields:

$$k = k_{298} \exp\left(\frac{E_a}{R298}\right) \exp\left(-\frac{E_a}{RT}\right) \quad \text{(Eq. 14)}$$

If drug degradation follows a first-order reaction, Eq. 2 can be used to describe the functional relationship between content and time. Substitution of Eq. 14 for k into Eq. 2 and rearrangement yields:

$$C = C_0 \exp\{-k_{298} t \exp[(E_a/R)(1/298 - 1/T)]\} \quad \text{(Eq. 15)}$$

For a first-order reaction, t_{90} can be determined by using the following relationship:

$$t_{90} = 0.1054/k_{298} \quad \text{(Eq. 16)}$$

Substituting Eq. 16 for k_{298} into Eq. 15 yields:

$$C = C_0 \exp\{-t(0.1054/t_{90}) \exp[(E_a/R)(1/298 - 1/T)]\} \quad \text{(Eq. 17)}$$

Equation 17 can be used now to provide direct estimates of t_{90} and E_a *via* nonlinear regression analysis. Time (t) and temperature (T) are the two independent variables. Similarly, equations for either zero- (Eq. 18) or second-order (Eq. 19) degradation can be derived as follows:

$$C = C_0\{1 - t(0.1/t_{90}) \exp[(E_a/R)(1/298 - 1/T)]\} \quad \text{(Eq. 18)}$$

$$C = \frac{C_0}{1 + t(0.11/t_{90}) \exp[(E_a/R)(1/298 - 1/T)]} \quad \text{(Eq. 19)}$$

EXPERIMENTAL

Generation of Data With or Without Error—Data were generated by using Eqs. 17–19 for first-, zero-, and second-order reactions, respectively. Different sets of input theoretical values of E_a and t_{90} were assigned to each case. A single temperature sequence [323 K, 333 K, 343 K, and 353 K (50–80°C)] and a specified time scale for each temperature were employed to generate, at first, errorless drug content data. Drug content data were expressed as percent of content at time zero. Random numbers, which were also expressed as percent, were selected from a normal distribution with a mean of zero and a prescribed standard deviation (24, 25). These selected random numbers (used as noise) were added to the errorless drug content data to obtain simulated experimental "raw data." The normality of the generated error distribution was tested by using the chi-square goodness-of-fit test (26). The mean and standard deviation of the error distribution were also computed to ascertain that the mean was close to zero and that the standard deviation was close to that prescribed.

Calculation of Initial Estimates for Nonlinear Regression Analysis—Since the nonlinear regression method uses the iterative technique, the proper initial estimates of the parameters should be provided to initiate the iteration (27). The initial estimates of E_a and t_{90} were calculated in the following manner: (*a*) the functional relationships between drug content and time (Eqs. 1, 4, and 3) were employed to estimate the rate constants at these temperatures; (*b*) the linearized form of the Arrhenius relationship (Eq. 6) was used to obtain the values of E_a and A; and (*c*) the shelf-lives (t_{90}) can be calculated using Eqs. 20–22 for zero-, first-, and second-order reactions, respectively:

$$t_{90} = (0.1C_0/A) \exp\left(\frac{E_a}{R298}\right) \quad \text{(Eq. 20)}$$

$$t_{90} = (0.1054/A) \exp\left(\frac{E_a}{R298}\right) \quad \text{(Eq. 21)}$$

Table I—Specification of Computer Simulation Using a Random Error Structure

Order of Reaction	Parameter E_a	Parameter t_{90}	n^a	N^b
First	25.00	124.0	32	30
	25.00	260.0	32	30
	10.00	124.0	32	30
	10.00	260.0	32	30
Zero	25.00	124.0	32	30
Second	25.00	124.0	32	30

[a] Number of data points in each set of data. [b] Number of data sets.

Table II—Influence of Data Noise on Final Parameter Estimates of a First-Order Reaction with Standard Deviations of 2.5, 5.0, and 7.5% [a]

	Standard Deviation					
	2.5%		5.0%		7.5%	
	E_a	t_{90}	E_a	t_{90}	E_a	t_{90}
Estimate[b]	25.02 ± 0.28 (1.1%)	124.6 ± 7.6 (6.1%)	25.06 ± 0.57 (2.3%)	126.4 ± 15.5 (12.3%)	25.09 ± 0.88 (3.5%)	128.1 ± 24.8 (19.3%)
Range of Estimate	24.61-25.57	113.0-139.6	24.14-26.19	101.9-159.7	23.70-26.91	91.9-185.8
SD of Estimate	0.29 ± 0.04	7.50 ± 1.12	0.60 ± 0.07	15.6 ± 2.9	0.88 ± 0.11	23.4 ± 5.7
Range of *SD* of Estimate	0.23-0.39	5.57-10.8	0.48-0.76	11.0-20.6	0.70-1.12	14.5-39.5
Bias[c]	0.23 (0.9%) ± 0.16	6.03 (4.9%) ± 4.53	0.44 (1.8%) ± 0.34	12.1 (9.8%) ± 9.8	0.68 (2.7%) ± 0.55	18.7 (15.1%) ± 16.4
Range of Bias	0.020-0.57	0.343-15.6	0.040-1.19	0.689-35.7	0.030-1.9	1.10-61.8

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 weeks. [b] Mean ± *SEM*; *CV* in parentheses. [c] Percent error from the theoretical values in parentheses.

$$t_{90} = \left(\frac{0.11}{C_0 A}\right) \exp\left(\frac{E_a}{R298}\right) \qquad \text{(Eq. 22)}$$

Statistical Methods—The Monte Carlo method of computer simulations (28) was used to evaluate the reliability of the statistics for the estimates of E_a and t_{90}. The NONLIN computer program (29) was employed to obtain values of final estimates of E_a and t_{90} in the nonlinear approach. After obtaining the final estimates of E_a and t_{90} for each set of data, a certain sampling distribution for either E_a or t_{90} was formed. The mean, standard deviation, and bias of the sampling distribution of the parameter estimates were determined. The bias, in absolute values, is defined as the difference between the parameter estimate and the theoretical value. The statistics of the parameter estimates were evaluated by the closeness of the mean of the sampling distribution of each parameter estimate to its theoretical value, and the relative magnitudes of the standard deviation and the bias.

For comparison, the classical linear approach was also used to obtain the predicted shelf-life of the drug. The shelf-life was estimated through steps involving (*a*) the use of the linearized form of drug content-time relationship to obtain the rate constants at elevated temperatures, (*b*) the application of the Arrhenius equation (Eq. 6) to estimate E_a and A, (*c*) the extrapolation from straight line with the best fit in the Arrhenius plot to obtain the room temperature rate constant and its 95% confidence interval, and (*d*) the transformation of the room temperature rate constant and its 95% confidence interval to the corresponding value of shelf-life and its 95% confidence interval. Three computer programs written in FORTRAN IV were employed to generate the simulated experimental data, to calculate the initial estimates for nonlinear regression analysis, and to obtain values of the shelf-life and its 95% confidence interval through the use of the classical linear approach.

RESULTS

Before performing the computer simulations, Eqs. 17–19 were fitted to errorless data. In every case, NONLIN gave final estimates of E_a and t_{90} exactly identical to the input theoretical values. Standard deviations of these two estimates were close to zero. However, they were found to be highly correlated, with a correlation coefficient of 0.97 in all cases. Estimates of E_a and t_{90} obtained from the simulated "raw data" with prescribed errors also had the same high correlation coefficient. The high dependence between parameter estimates might complicate the separate estimation of each parameter (30). Metzler and Tong (31) indicated that this high correlation between parameters may imply meaningless statistics of the final estimates.

To examine this high correlation problem between the computer estimates of E_a and t_{90} and to evaluate the applicability of this "one-step" nonlinear approach, the Monte Carlo method of computer simulations was employed to probe into the statistical properties of these two final estimates. The simulations were divided into four parts which examined (*a*) the influence of data noise on the final estimates of E_a and t_{90}, (*b*) the influence of different theoretical values on the estimates of E_a and t_{90}, (*c*) the influence of various orders of reaction on the final estimates of E_a and t_{90}, and (*d*) the influence of the "extensiveness" of stability data on the estimates of E_a and t_{90}. It was reasoned that if estimates of these two parameters with meaningful statistical properties could be obtained under a wide range of conditions and different distributions of random errors in the data set, this high correlation between them may be assumed to be attributed to the nature of the mathematical function employed for estimating these parameters.

The simulated experimental "raw data" employed for examining the influence of data noise, different theoretical values, and various orders of reaction on these two final parameter estimates were generated using a single temperature sequence of 50–80°C and a variable time scale for each temperature. The time scales were selected to achieve ≥60% degradation of drug content at the lowest temperature and ≥80% degradation of content for three higher temperatures. The specifications of the computer simulations is shown in Table I. A wide range of stability and energetics values were chosen to test this nonlinear mathematical relationship. Thirty sets of data with 32 data points per set were generated for each case.

Table III—Influence of Initial Estimates on Final Parameter Estimates of a First-Order Reaction [a]

	Initial Estimate		Final Estimate	
SD	E_a	t_{90}	E_a	t_{90}
2.5%	17.50	219.0	25.17 ± 0.27	127.4 ± 7.0
	27.11	185.1	25.17 ± 0.27	127.5 ± 7.0
	32.50	40.00	25.17 ± 0.27	127.5 ± 7.0
5.0%	32.50	40.00	24.85 ± 0.60	120.5 ± 14.9
	24.19	108.6	24.85 ± 0.60	120.5 ± 14.9
	17.50	210.0	24.85 ± 0.60	120.5 ± 14.9
7.5%	17.50	210.0	25.07 ± 0.77	125.5 ± 19.9
	24.75	118.1	25.07 ± 0.77	125.6 ± 19.9
	32.50	40.00	25.07 ± 0.77	125.6 ± 19.9

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 week.

The results for simulations which examined the influence of data noise on final estimates of E_a and t_{90} are shown in Tables II–IV. At all three noise levels, mean estimates for E_a and t_{90}, obtained as the average from 30 individual estimates given by NONLIN, were very close (within 3.3%) to the assigned theoretical values. The standard deviations of the estimates given by NONLIN, on the average, were found to be acceptable estimates of the standard errors of mean estimates for both E_a and t_{90} in all cases. Both standard errors of mean estimates and average biases of the estimates were small and responded well to the change of the noise level in the constructed data. Estimates of E_a and t_{90} also did not vary excessively among these 30 sets of data.

The effect of initial estimates on the computed estimates of E_a and t_{90} was also tested for each set of data. The data, shown in Table III, represented simulation trials from one representative run for each noise level. One of these three sets of initial estimates was calculated using the linear approach, as described in the experimental section. The other two sets of initial estimates were arbitrarily chosen to provide different starting points to initiate the iteration step. The results clearly indicated that the locations of the initial estimates had no apparent effect on final estimates of E_a and t_{90}.

The examination of the residuals in the regression analysis revealed that the standard deviations of residuals obtained were also consistent with the prescribed standard deviations (Table IV). The r^2 values, which can be used as measures of good fit (27), were close to 1.0 in all cases. The randomness of the residuals in a representative set of data is shown in Fig. 1. The composite results showed, therefore, that estimates of E_a and t_{90} with meaningful statistics can be obtained with Eq. 17 using simulated data with different levels of data noise.

Results for studies which examined the influence of different theoretical values and of various orders of reaction on the parameter estimates of E_a and t_{90} are shown in Tables V and VI, respectively. The statistics of the parameter estimates for E_a and t_{90} were shown to be satisfactory for these conditions.

The influence of the "extensiveness" of stability data on the final parameter estimates of E_a and t_{90} was also tested. Specification of different types of data

Table IV—Goodness of Fit of Data from a First-Order Reaction with Standard Deviations of 2.5, 5.0, and 7.5% [a]

	Standard Deviation		
	2.5%	5.0%	7.5%
SD of Residuals	2.4 ± 0.3	5.0 ± 0.6	7.4 ± 1.0
r^2	0.989 ± 0.0003	0.954 ± 0.013	0.900 ± 0.031

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 weeks.

Table V—Influence of Different Theoretical Values on Final Parameter Estimates of a First-Order Reaction using Data with a Standard Deviation of 5.0% [a]

	Set 1		Set 2		Set 3	
	E_a	t_{90}	E_a	t_{90}	E_a	t_{90}
Estimate[b]	24.99 ± 0.57 (2.3%)	259.9 ± 32.8 (12.6%)	9.99 ± 0.59 (5.9%)	124.0 ± 16.2 (13.1%)	9.99 ± 0.60 (6.1%)	260.0 ± 34.7 (13.4%)
Range of Estimate	24.09-26.30	215.7-343.4	8.98-11.38	101.8-166.0	8.96-11.42	212.5-351.1
SD of Estimate	0.61 ± 0.09	32.9 ± 6.6	0.61 ± 0.09	15.8 ± 3.2	0.62 ± 0.09	33.8 ± 7.0
Range of *SD* of Estimate	0.49-0.78	22.6-51.7	0.46-0.79	10.8-25.0	0.47-0.80	22.9-53.4
Bias	0.44 (1.8%) ± 0.35	25.0 (9.6%) ± 20.5	0.46 (4.6%) ± 0.37	12.5 (10.1%) ± 10.1	0.47 (4.7%) ± 0.38	26.7 (10.3%) ± 21.7
Range of Bias	0.010-1.3	0.291-83.4	0.020-1.4	0.340±42.0	0.030-1.4	1.20-91.1

[a] Theoretical values: (set 1) E_a = 25.00 kcal/mol, t_{90} = 260 weeks; (set 2) E_a = 10.00 kcal/mol, t_{90} = 124 weeks; (set 3) E_a = 10.00 kcal/mol, t_{90} = 260 weeks. [b] Mean ± *SEM; CV* in parentheses. [c] Percent error from the theoretical value in parentheses.

Table VI—Influence of Different Orders of Reaction on Final Parameter Estimates using Data with a Standard Deviation of 5.0% [a]

	Zero Order		Second Order	
	E_a	t_{90}	E_a	t_{90}
Estimate[b]	25.00 ± 0.16 (0.6%)	124.0 ± 10.8 (8.7%)	24.96 ± 0.77 (3.1%)	124.0 ± 21.1 (17.0%)
Range of Estimate	24.15-26.06	105.8-153.7	23.72-26.75	94.6-177.2
SD of Estimate	0.40 ± 0.06	10.2 ± 1.9	0.86 ± 0.13	21.3 ± 4.9
Range of *SD* of Estimate	0.30-0.53	7.06-13.0	0.69-1.1	14.7-37.3
Bias[c]	0.32 (1.3%) ± 0.26	8.22 (6.6%) ± 6.77	0.62 (2.5%) ± 0.45	16.5 (13.3%) ± 12.8
Range of Bias	0.0010-1.1	0.785-29.7	0.0020-1.8	0.138-53.2

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 weeks. [b] Mean ± *SEM; CV* in parentheses. [c] Percent error from the theoretical value in parentheses.

Table VII—Specification of Different Types of Data Structure

Structure	Temp. K	Time Scale, week Sampling Time	Last Sample	Percent Remaining of Drug Content[a]
Type I	323	10.0	80.0	16.8
	333	3.0	24.0	17.8
	343	1.0	8.0	17.7
	353	0.4	3.2	14.1
Type II	323	2.0	16.0	70.0
	333	0.7	5.6	66.9
	343	0.2	1.6	70.7
	353	0.07	0.56	71.0
Type III	323	0.4	3.2	93.1
	333	0.4	3.2	79.5
	343	0.4	3.2	50.1
	353	0.4	3.2	14.1

[a] At the last sampling time, using errorless data for estimation.

structure is shown in Table VII. The same temperature sequence (50-80°C) was used for all three types of data structure. "Raw data" with type I structure, representing an example of extensive data gathering, were generated employing a variable time scale to achieve ≥80% degradation in drug content at each temperature. Simulated "raw data" with type II structure were generated using a variable time scale for each temperature as well. However, these time scales were set to account for degradation of the drug content to only ~30% completion. Data with type III data structure were generated by employing a fixed sampling time and frequency for all four temperatures studied. Thus, data structures shown as type II and III represented situations in which only limited kinetic data were available, as often the case in real accelerated stability trials. Simulations using these types of data structures are shown in Table VIII. Despite considerable differences in the degree of "extensiveness" in the data base, it was shown that good estimates of E_a and t_{90} with satisfactory statistical properties could be obtained for each type of data structure.

Sometimes the initial drug content, C_0, is not known and should be estimated together with the estimation of E_a and t_{90}. Simulations for estimating these three parameters are shown in Table IX. The nonlinear approach is shown to be capable of providing good mean estimates and meaningful statistics for C_0, E_a, and t_{90}. Based on kinetic principles, Eqs. 17-19 can be easily modified to determine the shelf-life at any chosen content, *i.e.*, either as percent of initial or of labeled strength.

Comparison between the classical linear approach and the nonlinear approach for the prediction of drug stability was made by using a first-order degradation reaction as an example. The "raw data" generated for the study that examined the influence of data noise on the final estimates of E_a and t_{90} were employed for this comparison between the two methods. The shelf-lives (t_{90}) and their 95% confidence intervals as estimated by using either the classical linear approach or the present nonlinear approach for three representative sets of data (each at three noise levels) are shown in Fig. 2. The nonlinear approach appeared to provide better mean estimates of drug shelf-life with much smaller and more symmetrical 95% confidence intervals than did the classical approach in all cases examined.

It is desirable to compare the nonlinear approach with the classical linear approach using real stability data. Accelerated stability data of vitamin A in multivitamin tablets (2) and of NAD (32) were employed for this purpose; results of this comparison are shown in Table X. The linear and nonlinear methods provided identical mean estimates of shelf-life in both cases. Moreover, the nonlinear method also provided a more symmetrical and smaller 95% confidence interval in the prediction of vitamin A stability than did the classical linear approach. However, both linear and nonlinear methods predicted somewhat longer shelf-lives of these two compounds than that experimentally determined at room temperature. It is possible that both systems possess some degree of non-Arrhenius behavior, a problem that neither the linear nor the nonlinear approach can correct.

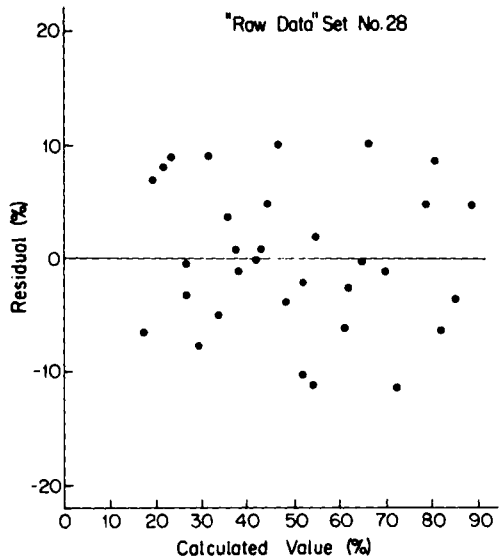


Figure 1—*Representative plot of residuals. Calculated values are expressed as percent of drug content at time zero. Residual is the difference between the observed value and calculated value of drug content for each data point.*

Table VIII—Influence of Different Types of Data Structure on Final Parameter Estimates of a First-Order Reaction using Data with a Standard Deviation of 2.5% [a]

	Data Structure					
	Type I		Type II		Type III	
	E_a	t_{90}	E_a	t_{90}	E_a	t_{90}
Estimate[b]	25.12 ± 0.32[c] (1.3%)	127.7 ± 8.6 (6.7%)	25.19 ± 0.73 (2.9%)	130.5 ± 18.6 (14.2%)	25.23 ± 0.84 (3.3%)	133.4 ± 26.7 (20.0%)
Range of Estimate	24.61-25.59	115.1-136.9	24.53-26.31	101.6-161.5	24.15-26.77	100.3-187.0
SD of Estimate	0.26 ± 0.03	6.90 ± 0.90	0.50 ± 0.05	13.2 ± 2.7	0.68 ± 0.07	22.1 ± 6.0
Range of *SD* of Estimate	0.23-0.30	5.55-8.22	0.43-0.59	8.96-17.8	0.56-0.80	14.0-34.0
Bias[d]	0.29 (1.1%) ± 0.17	7.76 (6.3%) ± 4.63	0.61 (2.4%) ± 0.41	15.5 (12.5%) ± 11.1	0.67 (2.7%) ± 0.51	21.2 (17.1%) ± 17.6
Range of Bias	0.10-0.48	0.490-16.2	0.029-1.3	0.0900-37.5	0.068-1.7	2.60-63.0

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 weeks. [b] Mean ± *SEM*; *CV* in parentheses. [c] N = 10. [d] Percent error from the theoretical value in parentheses.

Table IX—Values of Parameter Estimates for a First-Order Reaction, Including the Initial Drug Content, C_0, as a Parameter, using Data with a Standard Deviation of 2.5% [a]

	Parameter		
	E_a	t_{90}	C_0
Estimate[b]	24.97 ± 0.22 (0.9%)	122.3 ± 6.0 (4.9%)	100.2 ± 1.9 (1.9%)
Range of estimate	24.76-25.45	113.3-131.1	97.19-102.6
SD of estimate	0.33 ± 0.05	9.34 ± 1.51	1.66 ± 0.25
Range of *SD* of Estimate	0.26-0.41	7.94-11.8	1.30±1.99
Bias[c]	0.17 (0.7%) ± 0.13	5.26 (4.2%) ± 2.85	1.55 (1.6%) ± 0.89
Range of Bias	0.030-0.450	0.740-10.7	0.270-2.81

[a] Theoretical values: E_a = 25.00 kcal/mol; t_{90} = 124 weeks; C_0 = 100%. [b] Mean ± *SEM*; *CV* in parentheses. [c] Percent error from the theoretical value in parentheses.

DISCUSSION

The nonlinear approach shown in this report appears to be capable of predicting the shelf-life (t_{90}) directly and reproducibly. The uncertainties in the estimated t_{90} are the reflection of the errors inherent in the observed drug content data. The nonlinear approach can be applied to various orders of degradation reaction and a wide range of different theoretical values of E_a and t_{90}. It was shown that although some experimental "raw data" may not be very extensive, good estimation of shelf-life with meaningful statistical properties can still be obtained. Furthermore, this approach provides better mean estimates of t_{90} with much smaller 95% confidence intervals than the classical approach. Obviously, this method of drug stability prediction will be of little value without the aid of appropriate computer programs to carry out the nonlinear parameter estimation procedure. Fortunately, several computer programs including NONLIN (29), BMDP-P3R (33), SAAM (34), *etc.* are available for nonlinear regression analysis. The 1974 version of the NONLIN program was selected in this study because this particular computer program, which can be used to fit various models (27), is a widely used program in the area of pharmacokinetics and related pharmaceutical sciences.

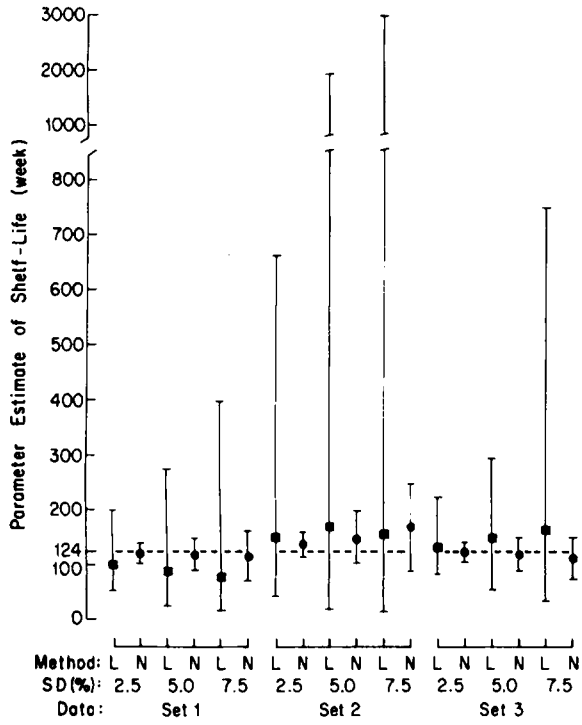


Figure 2—*Plot of the parameter estimate of the shelf-life (t_{90}) for each of the two methods of estimation in three sets of data with three different levels of noise. Data represent mean estimate and its 95% confidence interval after using the nonlinear approach (N, ●) and the classical linear approach (L, ■); (- - - -) represents the theoretical value of shelf-life of 124 weeks.*

A high correlation between E_a and t_{90} was found during the study of computer simulations. This high dependency between the parameter estimates was suggested by Metzler and Tong as a possible indication of meaningless statistics of these estimates (31). The Monte Carlo method of computer simulations was used to examine this high correlation problem and to evaluate the statistical properties of the final estimates of E_a and t_{90} in this report. During the regression analysis for each set of data by NONLIN, the final parameter estimates of E_a and t_{90} converged on the mathematical minimum after only 3-5 iteration steps without any computational difficulty.

The statistics of the parameter estimates can be evaluated by examining the sampling distributions of both E_a and t_{90}. The sampling distributions of both E_a and t_{90} were symmetrical about the mean estimates and had small standard deviations and biases. A large standard deviation is usually reflective of a poor determination of a parameter estimate (35, 36). Examination of the residuals (36) and of the correlation of the observations (r^2 value) can also be used to evaluate the statistics of the parameter estimates in a model. Our results indicated that (*a*) the standard deviations of the residuals were consistent with the prescribed standard deviations and (*b*) the plots of the residuals *versus* the calculated values for each set of data showed the random nature of the residuals. The r^2 values found were close to perfection (*i.e.*, 1.0) in all the data sets examined. Furthermore, it has been stated (37) that if the nonlinear regression method is correct, different sets of initial estimates for each set of data should converge on the same final estimates. Our simulation trials (Table III) showed this to be the case here. Therefore, it appears that good mean estimates of E_a and t_{90} with meaningful statistical properties are obtained using the nonlinear method, even though these two parameters are highly correlated.

Our observation of a correlation between the final estimates of E_a and t_{90} is not unique. Metzler and Tong (31) showed the high correlation between V_{max} and K_m in the parameter estimations of the Michaelis-Menten equation to be accompanied by meaningless statistics. They attributed this result to the peculiar nature of the mathematical function that relates these two variables, in that a change in the estimate of V_{max} may be compensated by a simultaneous change in K_m. However, Bard (22) had shown that meaningful statistical properties of the parameter estimates of A and E_a (Eq. 8) could be obtained despite the high correlation (*i.e.*, r = 0.98) between them. Thus,

Table X—Comparison Between the Classical Linear Approach and Nonlinear Method for the Prediction of Drug Stability using Literature Data

Literature Study	Shelf-Life, week			
	Linear Method		Nonlinear Method	
	Mean	95% CI[a]	Mean	95% CI[a]
I[b]	113.3[d]	49.25-262.2[d]	109.0	61.73-156.2
II[c]	6.49	5.67-7.44	6.63	5.54-7.71

[a] Confidence interval. [b] Prediction of vitamin A stability in multivitamin tablets. The accelerated stability data were obtained from Fig. 1 in the Ref. 2. [c] Prediction of the stability of NAD, taken from Ref. 32. [d] These values were calculated by using the reported rate constants.

it seems that the presence of a high correlation between parameter estimates may not necessarily indicate that the statistical properties of these final estimates are meaningless.

In conclusion, meaningful statistics of the final parameter estimates of E_a and t_{90} can be obtained by the nonlinear approach described in this paper. This approach is applicable over a wide range of different theoretical values of E_a and t_{90}, different orders of reaction, different levels of noise in data, and different types of data structure. The advantage of this nonlinear approach is that it uses data of drug content, time, and temperature to provide a direct estimation of shelf-life with relevant statistics. This method may be potentially useful for the realistic prediction of drug stability of pharmaceutical products.

REFERENCES

(1) N. R. Draper and H. Smith, "Applied Regression Analysis," Wiley, New York, N.Y., 1966, chap. 1.

(2) J. G. Slater, H. A. Stone, B. T. Palermo, and R. N. Duvall, *J. Pharm. Sci.*, **68**, 49 (1979).

(3) J. T. Carstensen and E. Nelson, *J. Pharm. Sci.*, **65**, 311 (1976).

(4) W.-H. Yang, *Drug Dev. Ind. Pharm.*, **7**, 63 (1981).

(5) E. R. Garrett, *J. Pharm. Sci.*, **51**, 811 (1962).

(6) S.-L. Lin, *Bull. Parent. Drug Assoc.*, **23**, 269 (1969).

(7) J. T. Cartensen and K. S. E. Su, *Bull. Parent. Drug Assoc.*, **25**, 287 (1971).

(8) C. J. Clark and H. F. Hudson, *Manuf. Chem. Aeros. News*, **39**, 25 (1968).

(9) D. J. Bentley, *J. Pharm. Sci.*, **59**, 464 (1970).

(10) N. G. Lordi and M. W. Scott, *J. Pharm. Sci.*, **54**, 531 (1965).

(11) T. B. L. Kirkwood, *Biometrics*, **33**, 736 (1977).

(12) S. P. Eriksen and H. Stelmach, *J. Pharm. Sci.*, **54**, 1029 (1965).

(13) M. A. Zoglio, J. J. Windheuser, R. Vatti, H. V. Maulding, S. S. Kornblum, A. Jacobs, and H. Hamot, *J. Pharm. Sci.*, **57**, 2080 (1968).

(14) H. V. Maulding and M. A. Zoglio, *J. Pharm. Sci.*, **59**, 333 (1970).

(15) M. A. Zoglio, H. V. Maulding, W. H. Streng, and W. C. Vincek, *J. Pharm. Sci.*, **64**, 1381 (1975).

(16) B. Edel and M. O. Baltzer, *J. Pharm. Sci.*, **69**, 287 (1980).

(17) J.-O. Waltersson and P. Lundgren, *Acta Pharm. Suec.*, **19**, 127 (1982).

(18) W.-H. Yang, *Drug Dev. Ind. Pharm.*, **8**, 189 (1982).

(19) M. Scher, *J. Pharm. Sci.*, **69**, 325 (1980).

(20) O. L. Davies and D. A. Budgett, *J. Pharm. Pharmacol.*, **32**, 155 (1980).

(21) O. L. Davies and H. E. Hudson, in "Statistics in the Pharmaceutical Industry," C. R. Buncher and J.-Y. Tsay, Eds., Dekker, New York, N.Y., 1981, chap. 15.

(22) Y. Bard, "Nonlinear Parameter Estimation," Academic, New York, N.Y., 1974, pp. 206–212.

(23) H. G. Boxenbaum, S. Riegelman, and R. M. Elashoff, *J. Pharmacokinet. Biopharm.*, **2**, 123 (1974).

(24) D. H. Lehmer, *Ann. Comp. Lab. Harvard Univ.*, **26**, 141 (1951).

(25) P. A. W. Lewis, A. S. Goodman, and J. M. Miller, *IBM Systems J.*, **8**, 136 (1969).

(26) G. W. Snedecor and W. G. Cochran, "Statistical Methods," 6th ed., The Iowa State University Press, Ames, Iowa, 1967, pp. 84–85.

(27) C. M. Metzler, *Pharmacol. Ther.*, **13**, 543 (1981).

(28) Y. Bard, "Nonlinear Parameter Estimation," Academic, New York, N.Y., 1974, chap. 3.

(29) C. M. Metzler, G. L. Elfring, and A. J. McEwen, "A Users Manual for NONLIN and Associated Programs," The Upjohn Co., Kalamazoo, Mich., 1974.

(30) G. P. Box and W. G. Hunter, *Technometrics*, **4**, 301 (1962).

(31) C. M. Metzler and D. D. M. Tong, *J. Pharm. Sci.*, **70**, 733 (1981).

(32) B. M. Colombo, P. Causa, P. Primavera, and V. Villano, *Il Farmaco, Ed. Pr.*, **27**, 327 (1972).

(33) J. Garcia-Pena and S. P. Azen, *Comput. Prog. Biomed.*, **10**, 185 (1979).

(34) M. Berman and M. F. Weiss, "Users Manual for SAAM," National Institutes of Health, Bethesda, Md., 1974.

(35) H. P. Wijnand and C. J. Timmer, *J. Pharmacokinet. Biopharm.*, **7**, 681 (1979).

(36) Y. Bard, "Nonlinear Parameter Estimation," Academic, New York, N.Y., 1974, chap. 7.

(37) K. T. Muir, *Comput. Biomed. Res.*, **13**, 307 (1980).

ACKNOWLEDGMENTS

Supported in part by NIH Grants GM 20852 and HL 22273.

Behavior of ATP Toward Phenothiazine Drugs

M. R. GASCO ˣ, M. E. CARLOTTI, and M. TROTTA

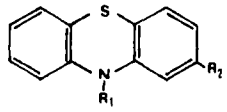
Received January 4, 1982, from the *Istituto di Chimica Farmaceutica e Tossicologica, Università di Torino, Corso Raffaello 31, 10125 Torino, Italy.* Accepted for publication February 21, 1983.

Abstract □ The tricyclic amines promazine, promethazine, chlorpromazine, triflupromazine, and trifluperazine form solid ion pairs with ATP in a 1:2 molar ratio. There is a good correlation between the measured K_{sp} and the apparent diffusion constants of the ion pairs with the critical micelle concentration (CMC) of the corresponding phenothiazines. Solid ion pairs are solubilized by phenothiazine micelles; the binding constants of ATP to drug micelles are calculated from solubility data at 25°C and can be related to the CMC of the phenothiazines.

Keyphrases □ ATP—behavior toward phenothiazine drugs, binding, micellar solubilization □ Binding—behavior of ATP toward phenothiazine drugs, micellar solubilization □ Phenothiazine drugs—behavior of ATP, binding, micellar solubilization □ Micellar solubilization—behavior of ATP toward phenothiazine drugs, binding

The mechanism of action of phenothiazine drugs is difficult to explain because of their great variety of biochemical and physiological effects. Membrane interactions seem to be important particularly in the case of chlorpromazine (1–4). Phenothiazines, like many tricyclic amines, have amphiphatic properties and are surface-active drugs (4–7). Their micelles are able to solubilize *in vitro* various high molecular weight drugs such as pteridine and porphyrin derivatives (8, 9); the binding constants of solubilized compounds with micelles are rather high. Previously, Blei (10) studied the decrease of chlorpromazine surface tension in the presence of ATP, while Moriguchi *et al.* (11) observed the formation of a 1:1 complex between the neuroleptic agent and ATP.

The aim of the present work is to investigate the behavior of five phenothiazines with different pharmacological activity



Promethazine	$R_1 = CH_2CH(CH_3)N(CH_3)_2$	$R_2 = H$
Promazine	$R_1 = CH_2CH_2CH_2N(CH_3)_2$	$R_2 = H$
Chlorpromazine	$R_1 = CH_2CH_2CH_2N(CH_3)_2$	$R_2 = Cl$
Triflupromazine	$R_1 = CH_2CH_2CH_2N(CH_3)_2$	$R_2 = CF_3$
Trifluperazine	$R_1 = CH_2CH_2CH_2$—N N—CH_3	$R_2 = CF_3$

0022-3549/84/0500-0662$01.00/0