DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 13(15), 2823-2838 (1987)

The Application of Multiresponse Estimation to Drug Stability Studies

K. Kowalski, M. Beno, C. Bergstrom and H. Gaud

Searle Research and Development
Division of G. D. Searle & Co.
4901 Searle Parkway
Skokie, Illinois 60077

Abstract

This paper focuses on the statistical methodology for the simultaneous analysis of stability data obtained on both parent compound and degradation products. A mathematical model based on kinetic theory is developed to describe both the disappearance of parent compound and appearance of degradation products. Analysis of data obtained from accelerated testing of drug stability is accomplished by extending the kinetic model to include the Arrhenius relationship. A simulated data set generated from parameter estimates obtained from an actual analysis is used as an example to illustrate the statistical techniques.

Introduction

Several recent papers[1-15] in the literature discuss the statistical treatment of data from drug stability studies. None of these papers discusses the simultaneous analysis of data collected on the parent compound and its degradation products. When the mechanism of drug decomposition is well understood and a model can be generated from kinetic theory, a unified approach to the analysis of stability data collected on the disappearance of the parent compound and the appearance of the degradation products can be taken. A method of multiresponse estimation can be used to obtain estimates of kinetic rate constants which relate the degradation of the parent

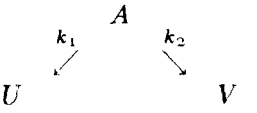

Figure 1: Illustration of Parallel First-Order Reactions

compound with the formation of the degradation products. The model can be further enhanced to treat data obtained from accelerated testing of drug stability using the Arrhenius relationship.

This paper will focus on a hypothetical drug decomposition mechanism to illustrate the method of multiresponse estimation. This mechanism has been postulated as effectively describing the stability of an actual pharmaceutical compound. The mathematical model and statistical methodology are developed for this postulated mechanism, and a simulated data set with known parameters is used to illustrate the statistical calculations. This methodology is also compared to an approach cited in the FDA draft guidelines. The paper concludes by noting that multiresponse estimation can be applied to a larger class of drug decomposition mechanisms.

Mathematical Model

Kinetic Theory

Consider an active drug, A, which degrades according to a parallel first-order mechanism to degradation products U and V. Figure 1 illustrates the mechanism, where k_1 and k_2 denote the first-order rate constants for the appearance of U and V, respectively. This mechanism is represented mathematically by the system of differential equations:

$$\frac{\partial A}{\partial t} = -(k_1 + k_2)A$$

$$\frac{\partial U}{\partial t} = k_1 A$$

$$\frac{\partial V}{\partial t} = k_2 A$$

where A, U and V denote concentrations at time, t. Letting A_0, U_0 and V_0 denote the respective initial concentrations, the solution to this system of differential equations yields the kinetic model:

$$A = A_0 e^{-(k_1+k_2)t} \tag{1a}$$

$$U = U_0 + \frac{k_1 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right) \tag{1b}$$

$$V = V_0 + \frac{k_2 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right) \tag{1c}$$

Arrhenius Relationship

A common model used to describe elementary chemical reaction rate constants as functions of temperature is the Arrhenius equation:

$$k_T = Fe^{-E_a/RT} \tag{2}$$

where k_T is the rate constant at temperature T (^oK), F is the pre-exponential or frequency factor, E_a is the activation energy and R is the ideal gas constant. A re-parameterization of Equation (2) is given by the expression:

$$k_T = k_0 e^{-E\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{3}$$

where k_0 is the rate constant at reference temperature T_0 and $E = E_a/R$. We assume Arrhenius behavior for each of the degradation pathway rate constants, k_1 and k_2, but not necessarily for the sum $(k_1 + k_2)$, since each pathway in a parallel degradation mechanism may have associated with it a different activation energy. Extending the kinetic model to include the Arrhenius relationship yields the mathematical model given by Equations (1a-1c) where the rate constants, k_1 and k_2, are given by the expressions:

$$k_1 = k_{1,0} e^{-E_1\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{4a}$$

$$k_2 = k_{2,0} e^{-E_2\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{4b}$$

$k_{1,0}$ and $k_{2,0}$ denote the rate constants for k_1 and k_2, respectively, at reference temperature T_0, and E_1 and E_2 denote the activation energy-to-gas constant ratios for reactions 1 and 2, leading to products U and V, respectively.

Statistical Methodology

Stochastic Model

For purposes of statistical exposition, an additive error structure is introduced so that the deterministic model given by Equations (1a-1c) and

(4a-4b) yields the stochastic model:

$$A = A_0 e^{-(k_1+k_2)t} + \varepsilon_A \tag{5a}$$

$$U = U_0 + \frac{k_1 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right) + \varepsilon_U \tag{5b}$$

$$V = V_0 + \frac{k_2 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right) + \varepsilon_V \tag{5c}$$

$$k_1 = k_{1,0} e^{-E_1\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{5d}$$

$$k_2 = k_{2,0} e^{-E_2\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{5e}$$

where ε_A, ε_U and ε_V represent the stochastic terms for A, U and V, respectively. These error terms are assumed to be mutually stochastically independent, each having mean zero and variances σ_A^2, σ_U^2 and σ_V^2.

Multiresponse Estimation

Ordinary Least Squares Consider a set of stability data wherein assay measurements for A, U and V are collected over time at various temperatures. Statistical modeling of this stability data using the postulated model given by Equations (5a-5e) is a multiresponse estimation problem involving the joint estimation of seven parameters, namely, A_0, U_0, V_0, $k_{1,0}$, $k_{2,0}$, E_1 and E_2 from the three response variables, A, U and V. A commonly employed technique to perform the parameter estimation is the method of ordinary least squares.

The method of ordinary least squares as it applies to the example in this paper is as follows. Define $Q = Q_A + Q_U + Q_V$ where Q_A, Q_U and Q_V denote the error sums of squares for A, U and V, respectively, and are given by the expressions:

$$Q_A = \sum_{t,T} \varepsilon_A^2 = \sum_{t,T} \left(A - A_0 e^{-(k_1+k_2)t}\right)^2$$

$$Q_U = \sum_{t,T} \varepsilon_U^2 = \sum_{t,T} \left(U - U_0 - \frac{k_1 A_0}{k_1 + k_2}(1 - e^{-(k_1+k_2)t})\right)^2$$

$$Q_V = \sum_{t,T} \varepsilon_V^2 = \sum_{t,T} \left(V - V_0 - \frac{k_2 A_0}{k_1 + k_2}(1 - e^{-(k_1+k_2)t})\right)^2$$

and where $\sum$ denotes summation over the entire data set (across time and temperature). The method of ordinary least squares can be formulated as the determination of values $\hat{A}_0$, $\hat{U}_0$, $\hat{V}_0$, $\hat{k}_{1,0}$, $\hat{k}_{2,0}$, $\hat{E}_1$ and $\hat{E}_2$ which minimize Q. Since Q is nonlinear in the parameters, numerical techniques must be

employed to determine the solution. The numerical techniques will not be discussed here. It suffices to say that several statistical computer packages are available to perform nonlinear least squares estimation.

The method of ordinary least squares in this setting has its limitations. Quite often with analytical methods, the variation associated with assay determinations increases with increasing magnitude of the response. Since marketable pharmaceutical products tend to have good stability, the magnitude of the parent compound is much greater than the magnitude of the degradation products. Thus, the variance of the parent compound tends to be greater than the variances for the degradation products. For our example, this suggests that σ_A^2 is much greater than σ_U^2 and σ_V^2. Consequently, Q_A comprises a greater proportion of Q than do Q_U and Q_V. Hence, A has more influence on the determination of the parameter estimates than does either U or V. Therefore, the method of ordinary least squares is undesirable in this setting.

Weighted Least Squares To alleviate this problem, weighted least squares is used such that all three components contribute equally to the parameter estimation. To formulate the weighted least squares technique, define $Q^* = Q_A^* + Q_U^* + Q_V^*$ where Q_A^*, Q_U^* and Q_V^* denote the weighted error sums of squares for A, U and V, respectively, and are given by the expressions

$$Q_A^* = \sum_{t,T} w_A \varepsilon_A^2 = \sum_{t,T} w_A \left(A - A_0 e^{-(k_1+k_2)t}\right)^2$$

$$Q_U^* = \sum_{t,T} w_U \varepsilon_U^2 = \sum_{t,T} w_U \left(U - U_0 - \frac{k_1 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right)\right)^2$$

$$Q_V^* = \sum_{t,T} w_V \varepsilon_V^2 = \sum_{t,T} w_V \left(V - V_0 - \frac{k_2 A_0}{k_1 + k_2}\left(1 - e^{-(k_1+k_2)t}\right)\right)^2$$

where w_A, w_U and w_V are appropriately chosen weights for A, U and V, respectively. The weighted least squares estimates denoted as $\hat{A}_0^*$, $\hat{U}_0^*$, $\hat{V}_0^*$, $\hat{k}_{1,0}^*$, $\hat{k}_{2,0}^*$, $\hat{E}_1^*$ and $\hat{E}_2^*$ are those parameter values which minimize Q^*. Having supplied w_A, w_U and w_V, the problem reduces to one of numerical analysis just as in the ordinary least squares problem.

The issue central to the weighted least squares problem is the choice of weights. An appropriate choice of weights would be such that $w_A\sigma_A^2 = w_U\sigma_U^2 = w_V\sigma_V^2$. This equality would insure that each component would contribute equally to the weighted error sums of squares. For example, the weights defined as $w_A = 1/\sigma_A^2$, $w_U = 1/\sigma_U^2$ and $w_V = 1/\sigma_V^2$ would yield such an equality. However, since σ_A^2, σ_U^2 and σ_V^2 are unknown, estimates must be obtained from the data. Initial weights can be obtained by first performing ordinary least squares. Letting $\hat{A}$, $\hat{U}$ and $\hat{V}$ denote the least squares predicted

values (obtained by substitution of the ordinary least squares parameter estimates into the deterministic equation) for A, U and V, respectively, we obtain $\hat{Q} = \hat{Q}_A + \hat{Q}_U + \hat{Q}_V$ where:

$$\hat{Q}_A = \sum_{t,T}(A - \hat{A})^2 \tag{6a}$$

$$\hat{Q}_U = \sum_{t,T}(U - \hat{U})^2 \tag{6b}$$

$$\hat{Q}_V = \sum_{t,T}(V - \hat{V})^2 \tag{6c}$$

Equations (6a-6c) provide a natural decomposition of the error sums of squares ($\hat{Q}$) such that estimates for σ_A^2, σ_U^2 and σ_V^2 can be obtained. These estimates denoted as $\hat{\sigma}_A^2$, $\hat{\sigma}_U^2$ and $\hat{\sigma}_V^2$ are:

$$\hat{\sigma}_A^2 = \hat{Q}_A/N_A$$
$$\hat{\sigma}_U^2 = \hat{Q}_U/N_U$$
$$\hat{\sigma}_V^2 = \hat{Q}_V/N_V$$

where N_A, N_U and N_V denote the numbers of observations in the summation for $\hat{Q}_A$, $\hat{Q}_U$ and $\hat{Q}_V$, respectively. The weights are then defined as $w_A = 1/\hat{\sigma}_A^2$, $w_U = 1/\hat{\sigma}_U^2$ and $w_V = 1/\hat{\sigma}_V^2$. Using these initial weights we can obtain the weighted least squares parameter estimates and consequently, the predicted values, $\hat{A}^*$, $\hat{U}^*$ and $\hat{V}^*$. Substituting the weighted least squares predicted values for the ordinary least squares predicted values in Equations (6a-6c) yields:

$$\hat{Q}_A^* = \sum_{t,T}(A - \hat{A}^*)^2 \tag{7a}$$

$$\hat{Q}_U^* = \sum_{t,T}(U - \hat{U}^*)^2 \tag{7b}$$

$$\hat{Q}_V^* = \sum_{t,T}(V - \hat{V}^*)^2 \tag{7c}$$

This leads to new estimates for the variances and the weights. This procedure can be performed for successive iterations until convergence in the weights is achieved. The following 5 step algorithm summarizes this iteratively re-weighted least squares procedure.

Step I: Perform ordinary least squares to obtain parameter estimates.

Step II: Compute variances and weights from Equations (6a-6c).

Step III: Perform weighted least squares to obtain parameter estimates.

Step IV: Compute variances and weights from Equations (7a-7c).

Step V: Check convergence of the weights. If the weights have converged, terminate the iterative process and use estimates from final iteration. Otherwise, go to Step III using updated weights from Step IV.

Confidence Intervals

Having performed the multiresponse estimation procedure on a given set of stability data, point estimates for predicted values for A, U and V for a given time and temperature can be made. Furthermore, based on a lower expiry limit for A and upper expiry limits for U and V, shelf-life estimates can be obtained by solving for the times at which expiry limits are met for A, U and V. The assay prediction estimates and the shelf-life estimates are subject to error (statistical variation). It is desirable therefore, to be able to quantify these errors and construct interval estimates with specified levels of confidence of containing the true values. In order to construct such intervals, an assumption about the distributional properties of the error terms is necessary. The stochastic error terms ε_A, ε_U and ε_V introduced in Equations (5a-5c) will be assumed to be normally distributed.

Prediction In general, exact expressions do not exist for the variance of a nonlinear function of parameter estimates; however, an approximation known as the linearization technique[16] or as the propagation of errors can be used to obtain an approximate variance. To illustrate this technique, suppose it is of interest to compute the variance of $\hat{A}^*$ at a fixed time t_1 and temperature T_1. We will denote the point estimate as $\hat{A}^*(t_1, T_1)$ which is given by the expression:

$$\hat{A}^*(t_1, T_1) = \hat{A}_0^* e^{-(\hat{k}_1^* + \hat{k}_2^*)t_1}$$

where

$$\hat{k}_1^* = \hat{k}_{1,0}^* e^{-\hat{E}_1^*(1/T_1 - 1/T_0)}$$

$$\hat{k}_2^* = \hat{k}_{2,0}^* e^{-\hat{E}_2^*(1/T_1 - 1/T_0)}$$

The approximate variance of $\hat{A}^*(t_1, T_1)$, denoted $S^2_{\hat{A}^*}$, is given by the expression:

$$S^2_{\hat{A}^*} = \underline{d}'_A \Sigma \underline{d}_A \tag{8}$$

where

$$\underline{d}'_A = \left(\frac{\partial A}{\partial A_0}, \frac{\partial A}{\partial U_0}, \frac{\partial A}{\partial V_0}, \frac{\partial A}{\partial k_{1,0}}, \frac{\partial A}{\partial k_{2,0}}, \frac{\partial A}{\partial E_1}, \frac{\partial A}{\partial E_2} \right) \Bigg|_{(t_1, T_1)}$$

and Σ is a 7×7 variance-covariance matrix of the parameter estimates. In practice, Σ is usually printed by nonlinear least squares computer packages. The elements of $\underline{d}_A$ are the partial derivatives of A evaluated at the estimated values of the parameters. In a similar fashion, variances for $\hat{U}^*(t_1, T_1)$ and $\hat{V}^*(t_1, T_1)$ can also be obtained by computing vectors of partial derivatives d_U and d_V to replace d_A in Equation (8).

An approximate lower one-tailed 95% confidence limit for A is given by the expression:

$$\hat{A}_l^*(t_1, T_1) = \hat{A}^*(t_1, T_1) - t(.95, N - 7)S_{\hat{A}^*}$$

where $t(.95, N - 7)$ is the 95th percentile of the Student's t-distribution with $N - 7$ degrees of freedom ($N = N_A + N_U + N_V$). A lower one-sided 95% confidence band for $A(t, T_1)$ can be constructed by computing $\hat{A}_l^*(t, T_1)$ for several values of t. Approximate upper one-tailed 95% confidence limits for $U(t_1, T_1)$ and $V(t_1, T_1)$ are given by the expressions:

$$\hat{U}_u^*(t_1, T_1) = \hat{U}^*(t_1, T_1) + t(.95, N - 7)S_{\hat{U}^*}$$
$$\hat{V}_u^*(t_1, T_1) = \hat{V}^*(t_1, T_1) + t(.95, N - 7)S_{\hat{V}^*}$$

Shelf-Life An estimate of the shelf-life for A at temperature T_1, given a lower expiry limit A_l, can be obtained by solving for t in the expression $\hat{A}^*(t, T_1) = A_l$. The solution for this expression, denoted $t_A(A_l, T_1)$, is given by the expression:

$$\hat{t}_A(A_l, T_1) = \frac{\ln \hat{A}_0^* - \ln A_l}{\hat{k}_1^* + \hat{k}_2^*} \tag{9}$$

Similarly, shelf-life estimates for U and V can be obtained by solving for t in the expressions: $\hat{U}^*(t, T_1) = U_u$ and $\hat{V}^*(t, T_1) = V_u$.

A method for obtaining a lower one-sided confidence limit for the shelf-life involves constructing a lower one-sided confidence band for A (or upper one-sided confidence bands for U and V). The shelf-life confidence limit is determined by the time point at which the confidence band (A, U or V) intersects the specified expiry limit. This is analogous to the approach that the FDA guidelines[17] suggest for shelf-life interval estimation in simple linear regression situations. An alternative method which requires fewer computer resources again makes use of the linearization technique. The linearization technique can be applied to Equation (9) to obtain the approximate variance of $\hat{t}_A(A_l, T_1)$ denoted as $S^2(\hat{t}_A)$. The approximate lower one-tailed 95% confidence limit for $t_A(A_l, T_1)$ is given by the expression:

$$\hat{t}_{A_l}(A_l, T_1) = \hat{t}_A(A_l, T_1) - t(.95, N - 7)S(\hat{t}_A)$$

Table 1
Parameter Values For Simulated Example Data Set

Parameter	Value
A_0 (%)	101.17
U_0 (%)	0.61
V_0 (%)	0.69
$k_{1,30}$ (weeks^{-1})	9.34×10^{-5}
$k_{2,30}$ (weeks^{-1})	2.72×10^{-5}
E_1 (°K)	13962.9
E_2 (°K)	14325.2
σ^2_A	3.210
σ^2_U	0.193
σ^2_V	0.042

Approximate variances and lower one-tailed confidence limits for the shelf-life of U and V can be obtained in a similar fashion. The value for the overall shelf-life can be taken to be the minimum of the three shelf-life confidence limits.

Example

Data Simulation

To illustrate the multiresponse estimation procedure, a data set has been simulated according to the stochastic model given in Equations (5a-5c). The parameter values used in the simulation are estimates obtained from an actual drug stability data set via the procedure described in this paper. The parameter values are listed in Table 1. The design and simulated data values for A, U and V in units of "percent of label claim" are shown in Table 2. These data were simulated on an IBM 4361 computer using the SAS[18] (Statistical Analysis System) function NORMAL, to simulate the normally distributed errors, ε_A, ε_U, and ε_V.

Data Analysis

The multiresponse estimation procedure was performed on the data shown in Table 2 using the SAS[19] procedure, NLIN. The parameter estimates and their respective standard errors are shown in Table 3. The simulated values are also shown in Table 3 for comparison. To illustrate the model fit, Figures 2-4 show the observed and predicted values for A, U and V, respectively.

Table 2
Example Data Set

Temp. (°C)	Time (weeks)	A (%)	U (%)	V (%)
	0	101.55	0.15	0.70
30	4	98.06	0.57	0.56
30	8	100.91	0.06	0.42
30	13	99.56	0.26	0.91
30	26	99.60	1.07	0.86
30	52	101.27	1.52	0.54
30	78	99.68	2.24	0.75
30	104	100.50	2.18	0.71
40	4	101.60	1.01	0.89
40	8	99.60	0.75	0.60
40	13	98.02	1.44	0.82
40	26	99.49	1.66	1.37
40	52	97.79	2.58	1.01
40	78	100.19	3.30	1.77
40	104	91.77	4.32	2.15
55	4	98.42	1.30	1.09
55	8	96.44	3.17	1.53
55	13	93.19	3.64	2.16
55	26	93.51	8.06	3.02
55	52	80.34	15.47	5.27
70	4	90.42	8.32	3.02
70	8	83.30	14.87	5.58
70	13	70.43	22.56	8.20

Table 3
Parameter Estimates and Standard Errors
for the Example Data Set

Parameter	Theoretical (Simulated)	Estimate	Standard Error
A_0 (%)	101.17	100.54	0.37
U_0 (%)	0.61	0.56	0.12
V_0 (%)	0.69	0.65	0.05
$k_{1,30}$ (weeks^{-1})	9.34×10^{-5}	9.26×10^{-5}	7.10×10^{-6}
$k_{2,30}$ (weeks^{-1})	2.72×10^{-5}	2.77×10^{-5}	2.91×10^{-6}
E_1 (°K)	13962.9	13961.9	210.1
E_2 (°K)	14325.2	14296.1	284.4
σ_A^2	3.210	2.735	
σ_U^2	0.193	0.174	
σ_V^2	0.042	0.034	

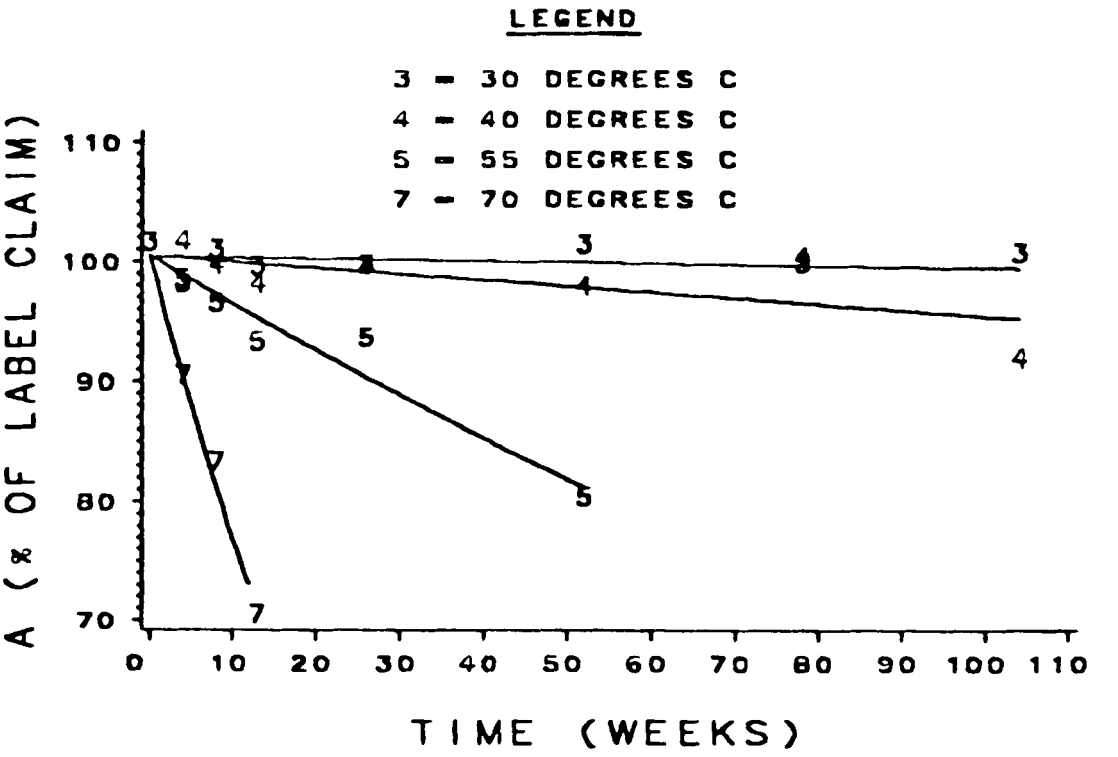

FIGURE 2: OBSERVED AND PREDICTED VALUES (A)

Confidence Intervals

Prediction Using the linearization technique to obtain approximate variances for prediction estimates, a 95% one-sided lower confidence limit for A, and 95% one-sided upper confidence limits for U and V were computed at 156 weeks and 30°C. Table 4 shows the prediction estimates, standard errors and confidence limits for A, U and V. For comparison, the theoretical values obtained from the deterministic portion of the model used to simulate the data are also shown.

Shelf-Life To compute shelf-life estimates, it is necessary to specify a lower expiry limit for A, and upper expiry limits for U and V. Consider for this example, a lower expiry limit for A of 90% and upper expiry limits for U and V of 3% and 2%, respectively. Table 5 shows the shelf-life estimates, standard errors and 95% one-sided lower confidence limits for A, U and V at 30°C. In addition, the theoretical shelf-life values are also shown. For this example, the overall expiry which is taken to be the minimum of the three confidence limits is 237 weeks.

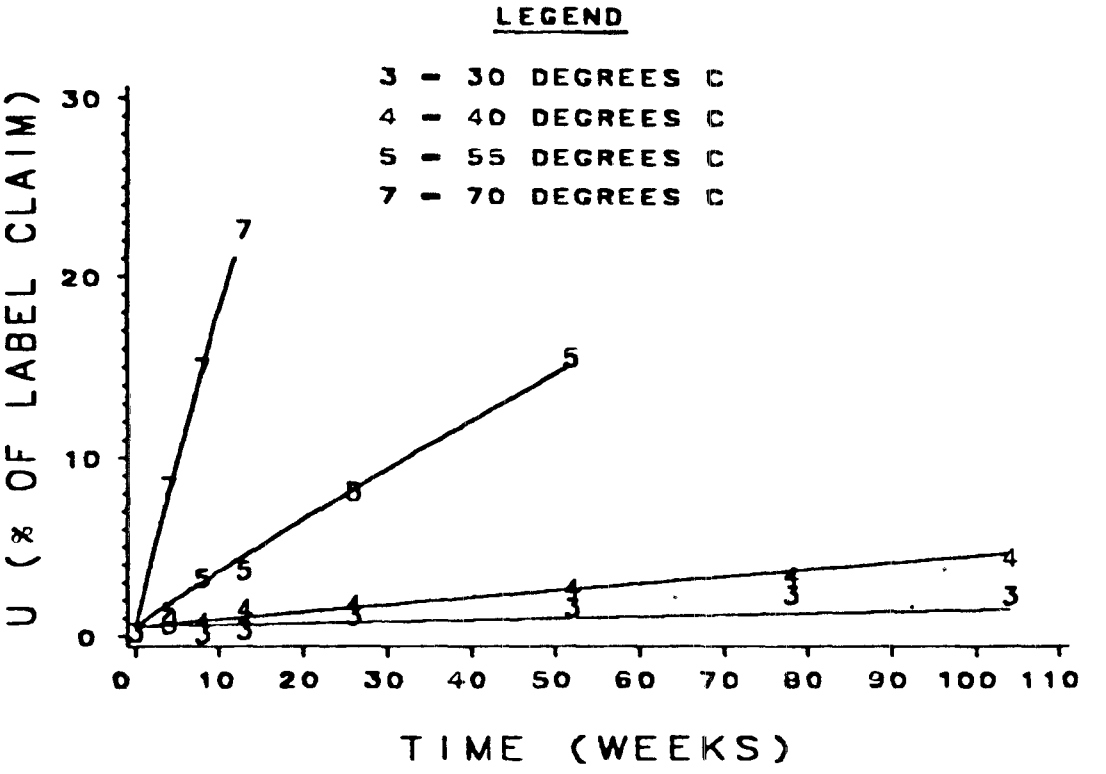

FIGURE 3: OBSERVED AND PREDICTED VALUES (U)

Discussion

An alternate method proposed by the FDA in their draft guidelines[17] is to fit simple linear regression models for the parent compound and degradation products over time for a given temperature. The methodology described in this paper is relatively complex and computer intensive as compared to the approach suggested by the FDA. However, it is the authors' opinion that the potential for improvements in accuracy and precision of the predictions warrants its consideration. To demonstrate the utility of this multiresponse estimation procedure, we shall compare the results of the analysis of the example data set from the previous section with results obtained by an alternate method. The alternate method is the approach proposed by the FDA with the added assumption that the zero-order rate constants for the disappearance of the parent compound follow the Arrhenius relationship.

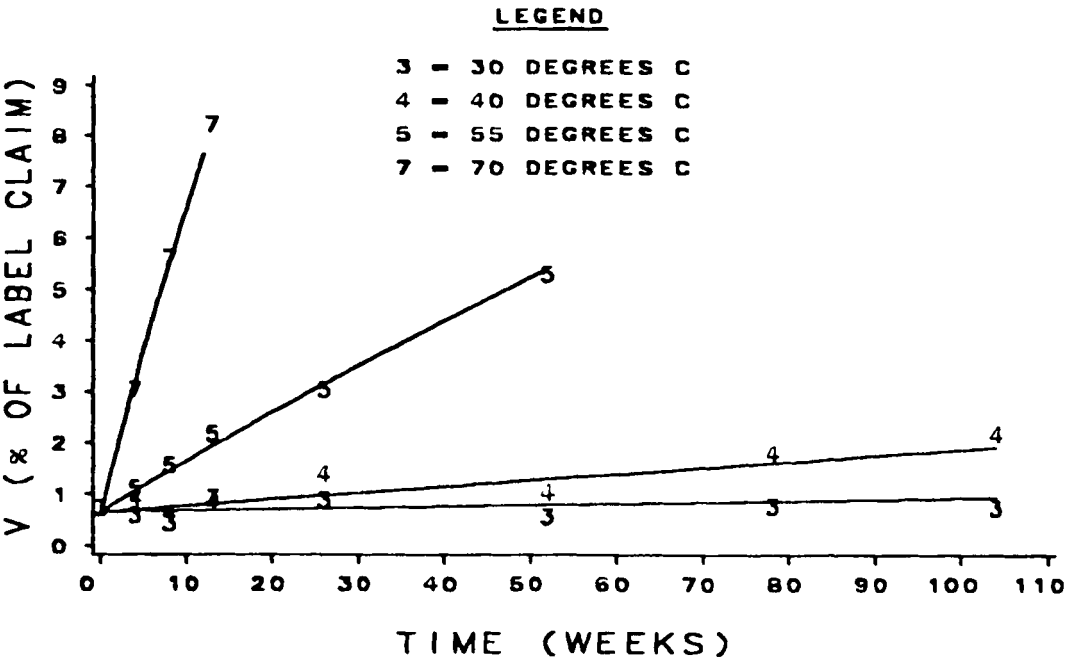

FIGURE 4: OBSERVED AND PREDICTED VALUES (V)

Table 4
Prediction Estimates and Confidence Limits
at 30°C and 156 weeks

	A	*U*	*V*
Theoretical (%)	99.28	2.07	1.12
Estimate (%)	98.67	1.99	1.08
Std.Error (%)	0.37	0.11	0.05
95% Conf. Limit (%)	98.05	2.18	1.16

Table 5
Shelf-Life Estimates and Confidence Limits at 30°C

	A	*U*	*V*
Theoretical (weeks)	970	257	490
Estimate (weeks)	921	267	499
Std.Error (weeks)	63	18	46
95% Conf. Limit (weeks)	815	237	422

Table 6
Prediction Estimates and Confidence Limits at 30°C and 156 Weeks Obtained From Alternate Method

	A	U	V
Theoretical (%)	99.28	2.07	1.12
Estimate (%)	98.49	3.65	0.75
Std.Error (%)	0.44	0.37	0.22
95% Conf. Limit (%)	97.72	4.37	1.18

From the plots of the observed values in Figures 2-4 one would expect that a linear approximation would be adequate. When data are obtained on the parent compound from accelerated testing of drug stability, it is commonly assumed that the rate constants at various temperatures follow the Arrhenius relationship. However, this assumption should not be made without first considering the kinetics of the degradation pathways. If it is reasonable to assume that the activation energies of the two pathways are equal, the Arrhenius relationship can be applied to the parent compound since the sum of the two rate constants will also follow the Arrhenius relationship. In our example, the simulated activation energies differ by less than one kcal/mole, and thus it appears to be a reasonable assumption.

A zero-order Arrhenius model was fit to the data for the parent compound in Table 2 using nonlinear least squares. For the degradation products, separate linear regression analyses were performed using only the 30°C data in Table 2. Table 6 shows the prediction estimates, standard errors and 95% one-sided confidence limits for A, U and V at 30°C and 156 weeks. The theoretical values are also shown in Table 6. Comparing Table 6 with Table 4, it can be observed that the prediction estimates for the multiresponse estimation method are more accurate than the estimates obtained by the alternate method. The prediction estimates in Table 6 are biased because the linear regression model is only a linear approximation to the nonlinear model. Furthermore, the standard errors of the prediction estimates using the linear regression approach are greater than those obtained using the multiresponse estimation procedure. The standard errors in Table 6 are inflated due to a lack of fit component, and because not all of the data for the degradation products were used. Consequently, the confidence intervals are wider and are less reliable.

For this example, the multiresponse estimation procedure performs better than the simple linear regression approach. This should come as no

surprise since the data were simulated according to the model used in the multiresponse estimation procedure. Surprising, however, is the magnitude of the improvement in accuracy and precision of the prediction estimates particularly for the degradation products. Therefore, when expiry dating must be based on individual degradation products and the drug decomposition mechanism is well understood, this multiresponse estimation procedure should be considered because of the potential improvements in accuracy and precision of the prediction estimates.

The multiresponse estimation procedure described in this paper was illustrated for a relatively simple drug decomposition mechanism. However, this estimation procedure can be applied to a larger class of drug decomposition mechanisms involving both parallel and sequential pathways.

References

1. Munguia, O., Parrilla, M.J. and Llabres, M. "The Statistical Interpretation of Accelerated Degradation Studies in Solutions," *J. Paren. Sci. & Tech.*, Vol. 40, No. 3 (1986), pp. 100-103.
2. Norwood, T., "Statistical Analysis of Pharmaceutical Stability Data," *Drug Dev. & Ind. Pharm.*, Vol. 12, No. 4 (1986), pp. 553-560.
3. Boudreau, C. F., "Handling Stability Data Via the Computer," *Drug Dev. & Ind. Pharm.*, Vol 10, No. 8 & 9 (1984), pp. 1527-1547.
4. Kmetec, V., Mrhar, A. Karba, R., and Kozjek, F., "Evaluation of Drug Stability Data by Analog-Hybrid Computer: Application to Lorazepam," *Int. J. Pharm.*, Vol. 21 (1984), pp. 211-218.
5. King, S. P., Kung, M., and Fung, H., "Statistical Prediction of Drug Stability Based on Nonlinear Parameter Estimation," *J. Pharm. Sci.*, Vol. 73, No. 5 (1984), pp. 657-662.
6. Bolton, S., "Factorial Designs in Pharmaceutical Stability Studies," *J. Pharm. Sci.*, Vol. 72, No. 4 (1983), pp. 362-366.
7. Nelson, P. R., "Stability Prediction Using the Arrhenius Model," *Computer Programs in Biomed.*, Vol. 16 (1983), pp. 55-60.
8. Tucker, I. G., and Owen, W. R., "Estimation of All Parameters from Nonisothermal Kinetic Data," *J. Pharm. Sci.*, Vol. 71, No. 9 (1982), pp. 969-974.
9. Yang, W., "Statistical Treatment of Stability Data," *Drug Dev. & Ind. Pharm.*, Vol. 7, No. 1 (1981), pp. 63-77.
10. Davies, O. L., and Budgett, D. A., "Accelerated Storage Tests on Pharmaceutical Products: Effect of Error Structure of Assay and Errors in Recorded Temperature," *J. Pharm. Pharmacol.*, Vol. 32, (1980), pp. 155-159.

11. Ogden, J. E., "A Weibull Shelf-Life Model for Pharmaceuticals," *Pharm. Tech.*, October (1978), pp. 45-49, 58.
12. Marubini, E., Milani, S., and Catalucci, R., "Criteria for Assigning the Order of Degradation Rate in the Accelerated Stability Tests: Statistical Aspects," *Il Farmaco (Ed. Pr.)*, Vol. 33, No. 3 (1978), pp. 95-110.
13. Thiers, R. E., Wu, G. T., Reed, A. H., and Oliver, L. K., "Sample Stability: A Suggested Definition and Method of Determination," *Clin. Chem.*, Vol. 22, No. 2 (1976), pp. 176-183.
14. Carstensen, J. T., and Su, K. S. E., "Statistical Aspects of Arrhenius Plotting," *Bull. Paren. Drug Assoc.*, Vol. 25, No. 6 (1971), pp. 287-302.
15. Bentley, D. L., "Statistical Techniques in Predicting Thermal Stability," *J. Pharm Sci.*, Vol. 59, No. 4 (1970), pp. 464-468.
16. Kendall, M. and Stuart, A., *The Advanced Theory of Statistics*, Vol. 1, 4th Edition, MacMillan, (1977), pp. 246-247.
17. Draft Guidelines For Stability Studies for Human Drugs and Biologics, Food and Drug Administration, Docket No. 84D-0115, August, 1985.
18. SAS User's Guide: Basics, Version 5 Edition, SAS Institute, (1985), pg. 260.
19. SAS User's Guide: Statistics, Version 5 Edition, SAS Institute, (1985), pp. 575-606.