The method presented, when applied with care, should be a valuable tool in drug absorption investigations. It seems to have some distinct advantages over the method proposed by Cutler (6). The simulation data generated by Cutler do not differentiate between the two approaches. However, these data represent rather ideal, somewhat unrealistic cases. More realistic tests might result in a definite distinction of the methods. The relative accuracy of the two approaches in practice is closely related to the issue of whether the impulse response data (*i.e.*, the intravenous bolus data) are best approximated by a polynomial or a polyexponential expression. It seems quite evident that the latter is the case, so the present method should be generally more accurate. Although this method seems promising, it appears from previous theoretical considerations (4) that the other deconvolution method presented by the author (5) is superior. However, the new method has the advantage that the rate and the amount (cumulative) of input are summarized in a simple polynomial form (Tables V and VI).

REFERENCES

(1) V. F. Smolen, *J. Pharmacokinet. Biopharm.*, **4**, 337 (1976).

(2) P. Veng Pedersen, *J. Pharm. Sci.*, **67**, 187 (1978).

(3) J. G. Wagner, "Fundamentals of Clinical Pharmacokinetics," Drug Intelligence Publications, Hamilton, Ill., 1975.

(4) P. Veng Pedersen, *J. Pharm. Sci.*, **69**, 298 (1980).

(5) *Ibid.*, **69**, 305 (1980).

(6) D. J. Cutler, *J. Pharmacokinet. Biopharm.*, **6**, 243 (1978).

(7) W. R. Boland and C. S. Duris, *BIT*, **11**, 139 (1971).

(8) N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, "SPSS, Statistical Package for the Social Sciences," McGraw-Hill, New York, N.Y., 1975.

(9) W. J. Dixon and M. B. Brown, "BMDP Biomedical Computer Programs," University of California Press, Berkeley, Calif., 1977.

(10) C. Daniel and F. S. Wood, "Fitting Equations to Data," Wiley, New York, N.Y., 1971, p. 86.

(11) P. Veng Pedersen, *J. Pharmacokinet. Biopharm.*, **5**, 513 (1977).

(12) C. M. Metzler, "A Users Manual for NONLIN," Tech. Rept. 7293/7292/005, Upjohn Co., Kalamazoo, Mich., 1969.

(13) G. Dahlquist and Å. Björk, "Numerical Methods," Prentice-Hall, Englewood Cliffs, N.J., 1974.

(14) B. M. Brown, "The Mathematical Theory of Linear Systems," Chapman and Hall, London, England, 1965.

Novel Approach to Bioavailability Testing: Statistical Method for Comparing Drug Input Calculated by a Least-Squares Deconvolution Technique

PETER VENG PEDERSEN

Received January 29, 1979, from the *Department of Pharmacy, School of Pharmacy, University of California, San Francisco, CA 94143.* Accepted for publication July 18, 1979.

Abstract □ A novel approach to bioavailability testing is presented. The approach is model independent because it does not assume a specific pharmacokinetic model and does not use absorption, distribution, or elimination rate constants or a volume term. The method, which requires intravenous administration, is compared to classical bioavailability evaluation methods. Evaluation of drug input is based on the same assumptions required for using the area under the curve. No extrapolation beyond the last data point is required. Two statistics are derived that enable a comparison of the rate and the cumulative amount of input of two inputs for various times. A differential confidence profile is calculated that allows a more detailed and intrinsic bioavailability comparison than previous methods. The approach is demonstrated on simulated data containing random noise and shows satisfactory performance.

Keyphrases □ Bioavailability—testing, drug input, least-squares deconvolution technique □ Deconvolution—least-squares technique, bioavailability testing, drug input □ Drug availability—testing, least-squares deconvolution technique

The quality of a drug product as a drug delivery system is determined by the rate and extent of delivery of the active form to the biological environment responsible for the pharmacological effect. In most cases, this environment is neither known nor may be sampled for the drug. However, a close relationship usually exists between the drug concentration in a sampled environment (*e.g.*, blood) and its pharmacological response. Therefore, this environment may be a useful indicator of the drug input into the response environment.

Drug delivery usually is characterized in bioavailability terms with the blood as the sampleable environment. Bioavailability commonly is defined by the rate and the extent of drug input into the systemic circulation (1). Bioavailability comparisons usually are based on three parameters from a single-dose blood level curve: (*a*) the area under the curve, AUC (extrapolated); (*b*) the time of the peak concentration, t_{max}; and (*c*) the peak concentration, C_{max} (2).

These parameters are associated conceptually with the extent (AUC) and the rate (t_{max} and C_{max}) of input. This association is related to linear pharmacokinetic assumptions. For example, AUC is a proper measure of the total input only if the response (concentration) is linear with respect to the input (3, 4). If the system is not linear, a larger AUC does not guarantee greater input. Comparisons of the drug input on the basis of AUC may be inaccurate for several reasons:

1. The tail area must be estimated by extrapolation. Consequently, this area can be determined only by a model-dependent approach that assumes a certain functional form for the tail or the total curve.

2. The tail area frequently is estimated from the last observations (*e.g.*, log-linear extrapolation), based on the assumption that these points predict the behavior in the tail. The tail area usually is determined poorly in this way due to low information density of the terminal set of points. This problem is complicated further by constraints in time and the number of samples when dealing with human subjects. The experimenter must decide whether more samples should be taken in the terminal phase, where little or no input takes place, to predict the tail area better or whether these samples would be more valuable in the input phase where the real information about the input is present and where C_{max} and t_{max} are to be estimated.

3. The problem is complicated by the fact that the tail

0022-3549/80/0300-0318$01.00/0

area is determined by an asymptotic-type infinite extrapolation that is more complex than an intercept-type extrapolation.

A new approach to bioavailability testing is presented that enables a more detailed comparison of the rate and the extent of drug input than conventional methods. The method does not use *AUC* and does not require extrapolation beyond the last data point. The approach is model independent in the classical linear pharmacokinetic sense because it does not assume a specific pharmacokinetic model and does not use absorption, distribution, or elimination rate constants or a volume term. This method of evaluating drug input is based on the same assumptions as those required for using *AUC* as a proper measure of the drug input. These assumptions are: (*a*) the input evaluated is a noninteracting primary input, and (*b*) there is a time-invariant linearity between the response (concentration) measured in the sampled environment (the blood) and the input (4). These assumptions can be tested experimentally as discussed previously (4). Their violation invalidates the classical linear approaches to bioavailability testing (5) under the current bioavailability definition (1, 6).

THEORY

It was shown (4) that if the input rate into a linear response environment is given in polynomial form by[1]:

$$f(t) = \sum_{i=1}^{N} \beta_i t^{i-1} \qquad \text{(Eq. 1)}$$

and the unit impulse response (the characteristic response) of the environment is given by:

$$c_u(t) = \sum_{i=1}^{n} a_i e^{\lambda_i t} \qquad \text{(Eq. 2)}$$

then the response to $f(t)$ is:

$$c(t) = \sum_{i=1}^{N} \beta_i \phi_i(t) \qquad \text{(Eq. 3)}$$

which is linear with respect to the unknown parameters $\beta_1, \beta_2, \ldots, \beta_N$, and:

$$\phi_i(t) = (i-1)! \sum_{j=1}^{n} \frac{a_j}{\lambda_j^i} \left[e^{\lambda_j t} - \sum_{k=0}^{i-1} \frac{(\lambda_j t)^k}{k!} \right] \qquad \text{(Eq. 4)}$$

Without loss of generality, two brands or drug delivery systems may be compared by assuming the blood to be the response environment, the response to be the drug concentration, $c(t)$, and $c_m(t_{mi})$ to be the response measured at the time t_{mi}, $i = 1, 2, \ldots, M_m$, for the mth drug product ($m = 1, 2$). To facilitate the statistical analysis, the measured responses can be written in the form:

$$y_{mi} = \sum_{j=1}^{N_m} \beta_{mj} x_{mij} + \epsilon_{mi} \qquad \begin{matrix} i = 1, 2, \ldots, M_m \\ m = 1, 2 \end{matrix} \qquad \text{(Eq. 5)}$$

where:

$$x_{mij} = \phi_{mj}(t_{mi}) \qquad \begin{matrix} i = 1, 2, \ldots, M_m \\ j = 1, 2, \ldots, N_m \\ m = 1, 2 \end{matrix} \qquad \text{(Eq. 6)}$$

and:

$$y_{mi} = c(t_{mi}) \qquad \begin{matrix} i = 1, 2, \ldots, M_m \\ m = 1, 2 \end{matrix} \qquad \text{(Eq. 7)}$$

Equations 1 and 4 may be written similarly:

$$f_m(t) = \sum_{i=1}^{N_m} \beta_{mi} t^{i-1} \qquad m = 1, 2 \qquad \text{(Eq. 8)}$$

$$\phi_{mi}(t) = (i-1)! \sum_{j=1}^{n_m} \frac{a_{mj}}{\lambda_{mj}^i} \left[e^{\lambda_{mj} t} - \sum_{k=0}^{i-1} \frac{(\lambda_{mj} t)^k}{k!} \right] \qquad m = 1, 2 \qquad \text{(Eq. 9)}$$

It is of interest in bioavailability studies to test the significance of differences in the rate and/or extent of drug input between drug products. This is equivalent to testing the null hypothesis:

$$H_0: \quad \mu_1|t_0 = \mu_2|t_0 \qquad \text{(Eq. 10)}$$

against:

$$H_1: \quad \mu_1|t_0 \neq \mu_2|t_0 \qquad \text{(Eq. 11)}$$

where:

$$\mu_m|t_0 = f_m(t_0) = \sum_{i=1}^{N_m} \beta_{mi} t_0^{i-1} \qquad m = 1, 2 \qquad \text{(Eq. 12)}$$

for various values of $t = t_0$.

In matrix–vector notation, Eq. 5 is:

$$\mathbf{y}_m = \mathbf{X}_m \boldsymbol{\beta}_m + \boldsymbol{\epsilon}_m \qquad m = 1, 2 \qquad \text{(Eq. 13)}$$

The statistics to be derived for testing H_0 will be based on a normality assumption about the errors; *i.e.*, $\boldsymbol{\epsilon}_m$ is a random vector distributed $N(\mathbf{0}, \sigma_m^2 \mathbf{I})$, where $\mathbf{I}$ is the identity matrix and σ_m^2 is unknown. The unknown $N_m \times 1$ parameter vector, $\boldsymbol{\beta}_m$, is estimated by a multiple linear regression technique (4). It is assumed that the observed data are sufficient to ensure that the rank of the $M_m \times N_m$ matrix, $\mathbf{X}_m$, is $N_m \leq M_m$.

For this statistical model the least-squares estimates, denoted by $\hat{}$, are given by:

$$\hat{\boldsymbol{\beta}}_m = \mathbf{V}_m \mathbf{X}_m' \mathbf{y}_m \qquad m = 1, 2 \qquad \text{(Eq. 14)}$$

and:

$$\hat{\sigma}_m = \mathbf{e}_m' \mathbf{e}_m / (M_m - N_m) \qquad m = 1, 2 \qquad \text{(Eq. 15)}$$

where:

$$\mathbf{V}_m = (\mathbf{X}_m' \mathbf{X}_m)^{-1} \qquad m = 1, 2 \qquad \text{(Eq. 16)}$$

and:

$$\mathbf{e}_m = \mathbf{y}_m - \mathbf{X}_m \hat{\boldsymbol{\beta}}_m \qquad m = 1, 2 \qquad \text{(Eq. 17)}$$

are best linear unbiased estimates (BLUE) of $\boldsymbol{\beta}_m$ and σ_m^2. They are sufficient, efficient, complete, and consistent estimators. Furthermore, $\hat{\boldsymbol{\beta}}_m$ and $\hat{\sigma}_m$ are stochastically independent and $\hat{\boldsymbol{\beta}}_m$ is distributed $N(\boldsymbol{\beta}_m, \sigma_m^2 \mathbf{V}_m)$. Consequently, the least-squares estimate of the input rate at $t = t_0$ is the best linear unbiased estimate and is distributed:

$$\hat{f}_m(t_0) = \boldsymbol{\tau}_m' \hat{\boldsymbol{\beta}}_m \sim N(\mu_m|t_0, \sigma_m^2 \boldsymbol{\tau}_m' \mathbf{V}_m \boldsymbol{\tau}_m) \qquad m = 1, 2 \qquad \text{(Eq. 18)}$$

where:

$$\boldsymbol{\tau}_m' = (1, t_0, t_0^2, \ldots, t_0^{N_m - 1}) \qquad m = 1, 2 \qquad \text{(Eq. 19)}$$

Since $[\hat{f}_1(t_0) - \mu_1|t_0]$ and $[\hat{f}_2(t_0) - \mu_2|t_0]$ are stochastically independent, the random variable Q_1 given by:

$$Q_1 = \frac{[\hat{f}_1(t_0) - \hat{f}_2(t_0)] - (\mu_1|t_0 - \mu_2|t_0)}{[\sigma_1^2 \boldsymbol{\tau}_1' \mathbf{V}_1 \boldsymbol{\tau}_1 + \sigma_2^2 \boldsymbol{\tau}_2' \mathbf{V}_2 \boldsymbol{\tau}_2]^{1/2}} \qquad \text{(Eq. 20)}$$

is standard and normally distributed. Furthermore, $(M_m - N_m)\hat{\sigma}_m^2 / \sigma_m^2$ is χ^2-distributed with $(M_m - N_m)$ degrees of freedom. Thus, the random variable Q_2 given by:

$$Q_2 = (M_1 - N_1)\hat{\sigma}_1^2 / \sigma_1^2 + (M_2 - N_2)\hat{\sigma}_2^2 / \sigma_2^2 = \frac{\mathbf{e}_1' \mathbf{e}_1}{\sigma_1^2} + \frac{\mathbf{e}_2' \mathbf{e}_2}{\sigma_2^2} \qquad \text{(Eq. 21)}$$

is χ^2-distributed with $(M_1 - N_1) + (M_2 - N_2)$ degrees of freedom because $(M_1 - N_1)\hat{\sigma}_1^2 / \sigma_1^2$ and $(M_2 - N_2)\hat{\sigma}_2^2 / \sigma_2^2$ are stochastically independent. It can be shown that the standard normal random variable Q_1 and the χ^2-distributed random variable Q_2 are stochastically independent. Consequently, by definition, the random variable $T = Q_1 / [Q_2 / (M_1 - N_1 + M_2 - N_2)]^{1/2}$ is t-distributed with $(M_1 - N_1 + M_2 - N_2)$ degrees of freedom. If the two population variances are equal, $\sigma_1^2 = \sigma_2^2$, then T becomes independent of σ^2, so that the H_0-hypothesis can be tested by a t-test:

$$T = \frac{\hat{f}_1(t_0) - \hat{f}_2(t_0)}{\left[\dfrac{\mathbf{e}_1' \mathbf{e}_1 + \mathbf{e}_2' \mathbf{e}_2}{M_1 - N_1 + M_2 - N_2} (\boldsymbol{\tau}_1' \mathbf{V}_1 \boldsymbol{\tau}_1 + \boldsymbol{\tau}_2' \mathbf{V}_2 \boldsymbol{\tau}_2) \right]^{1/2}} \qquad \text{(Eq. 22)}$$

[1] Note that the notational expressions for the input rate (Eq. 1), the impulse response (Eq. 2), and the input response (Eq. 3) were changed from previous notations (4) to facilitate the present statistical analysis.

Table I—Origin of the Simulated Data Used

Data Set	Unit Impulse Response $a_1e^{\lambda_1 t} + a_2e^{\lambda_2 t}$				Input Rate $f(t) = kKe^{-kt}$	
	a_1	a_2	λ_1	λ_2	K	k
1	1	1	−4	−1	Unit impulse	
2	1.727	7.523	−16.30	−2.134	Unit impulse	
3	1	1	−4	−3	Unit impulse	
4	1	1	−4	−1	3.0	1.5
5	1.727	7.523	−16.30	−2.134	1.0	1.0
6	1	1	−4	−3	3.0	1.5

with a critical region (H_0 rejection region) $|T| > t\ (\alpha/2, M_1 - N_1 + M_2 - N_2)$ for a type 1 error probability α.

The equality of the population variances can be tested according to an F-test:

$$F = \frac{\mathbf{e}_1'\mathbf{e}_1}{\mathbf{e}_2'\mathbf{e}_2} \qquad \text{(Eq. 23)}$$

with a critical region $F > F\ (\alpha, M_1 - N_1, M_2 - N_2)$, since $\mathbf{e}_1'\mathbf{e}_1/\sigma^2$ and $\mathbf{e}_2'\mathbf{e}_2/\sigma^2$ are stochastically independent, χ^2-distributed, random variables with the respective degrees of freedom.

If the F-test shows that the variances are not equal, then the t-test (Eq. 22) should not be used. For such cases, the H_0-hypothesis can be tested using the following statistic (see *Appendix*):

$$w = \frac{\hat{f}_1(t_0) - \hat{f}_2(t_0)}{\left[\dfrac{\mathbf{e}_1'\mathbf{e}_1\tau_1'\mathbf{V}_1\tau_1}{M_1 - N_1} + \dfrac{\mathbf{e}_2'\mathbf{e}_2\tau_2'\mathbf{V}_2\tau_2}{M_2 - N_2}\right]^{1/2}} \qquad \text{(Eq. 24)}$$

The w-statistic is approximated by a t-distribution with degrees of freedom, DF, given by[2]:

$$DF = \left[\frac{u^2}{M_1 - N_1} + \frac{1 - u^2}{M_2 - N_2}\right]^{-1} \qquad \text{(Eq. 25)}$$

where:

$$u = \frac{\mathbf{e}_1'\mathbf{e}_1\tau_1'\mathbf{V}_1\tau_1/(M_1 - N_1)}{\mathbf{e}_1'\mathbf{e}_1\tau_1'\mathbf{V}_1\tau_1/(M_1 - N_1) + \mathbf{e}_2'\mathbf{e}_2\tau_2'\mathbf{V}_2\tau_2/(M_2 - N_2)} \qquad \text{(Eq. 26)}$$

The critical region (H_0 rejection region) is $|w| > t(\alpha/2, DF)$. A $100(1 - \alpha)$-percent confidence interval for the drug input rate at time $t = t_0$ is:

$$\hat{f}_m(t_0) - t\left(\frac{\alpha}{2}, M_m - N_m\right)\left[\frac{\mathbf{e}_m'\mathbf{e}_m\tau_m'\mathbf{V}_m\tau_m}{M_m - N_m}\right]^{1/2} \le \mu_m|t_0 \le \hat{f}_m(t_0) + t\left(\frac{\alpha}{2}, M_m - N_m\right)\left[\frac{\mathbf{e}_m'\mathbf{e}_m\tau_m'\mathbf{V}_m\tau_m}{M_m - N_m}\right]^{1/2} \qquad m = 1, 2 \quad \text{(Eq. 27)}$$

This treatment dealt with a statistical analysis of the *rate* of drug input.

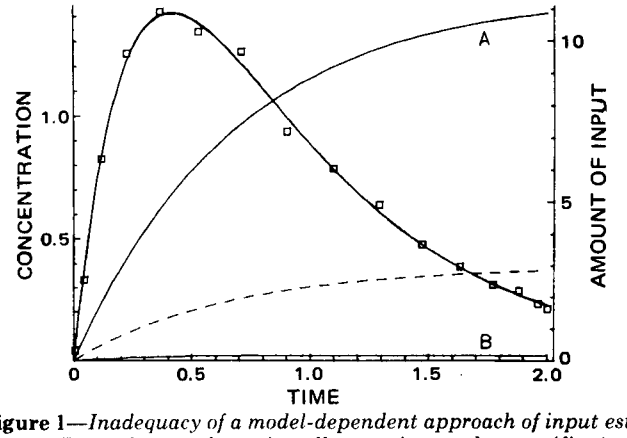

Figure 1—*Inadequacy of a model-dependent approach of input estimation. Examples are of two virtually superimposed curves (fits 1 and 2 in Table IV) fitted to Data Set 6. Curves A and B are the inputs calculated from the two fitted curves. The broken curve is the true input.*

[2] The degrees of freedom thereby is a real number, requiring interpolation in the t-table. However, most computer routines for the t-distribution permit noninteger degrees of freedom.

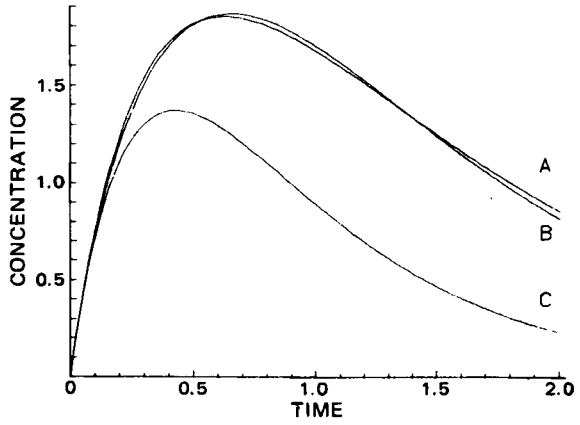

Figure 2—*Response curves A, B, and C from which the simulated Data Sets 4, 5, and 6, respectively, are derived by adding 5% random, normally distributed error. Curves A and B have widely different inputs; curves A and C have the same input (Table I).*

However, the *extent* of drug input given by:

$$A_m(t) = \int_0^t f_m(t)\,dt = \sum_{i=1}^{N_m} \frac{\beta_{mi}}{i!}\,t^i \qquad m = 1, 2 \qquad \text{(Eq. 28)}$$

can be treated similarly by replacing $f_m(t)$ with $A_m(t)$ in the respective equations and redefining τ_m' (Eq. 19) as:

$$\tau_m' = \left(t_0, \frac{t_0^2}{2}, \frac{t_0^3}{3}, \ldots, \frac{t_0^{Nm}}{N_m}\right) \qquad m = 1, 2 \qquad \text{(Eq. 29)}$$

The analysis was based on the ordinary least-squares Gauss–Markov statistical model. It is extended readily by standard means (7) to the generalized linear regression model involving correlated and/or heteroscedatic errors. The Gauss–Markov statistical model assumes no error in the independent variables, which is not true in the present case. However, the independent variables are stochastically independent random variables that do not depend on the β parameters, so the above estimation, testing, and prediction still apply. Only two aspects of the analysis are affected: (*a*) the probability in the t- and F-tests is not exact but only approximate, and (*b*) the power of the tests is different. The effect of these modifications becomes smaller the more accurately the unit impulse response is determined.

EXPERIMENTAL

The rate and the cumulative amount of input were determined by the least-squares deconvolution method presented previously (4). The data used for the demonstration of the proposed approach were generated from an arbitrarily chosen linear model with a two-exponential unit impulse response and a first-order input (Table I). Random, normally dis-

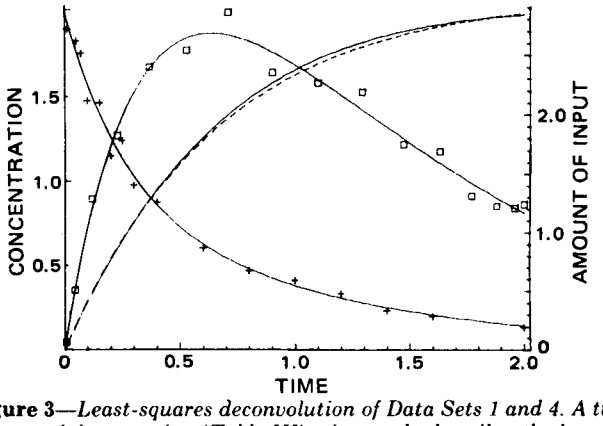

Figure 3—*Least-squares deconvolution of Data Sets 1 and 4. A two-exponential expression (Table III) adequately describes the impulse response data (+). The input function Eq. 1 is estimated subsequently by fitting Eq. 3 to the response data (□) by multiple linear regression. The continuously increasing curve is the calculated input amount (Table V). The broken curve is the exact input.*

Table II—Data Used in Simulations [a]

Unit Impulse Response				First-Order Input Response			
	Concentration				Concentration		
Time	Data Set 1	Data Set 2	Data Set 3	Time	Data Set 4	Data Set 5	Data Set 6
0.01	1.89	8.42	1.79	0.00482	0.0402	0.0467	0.0407
0.05	1.82	8.21	1.79	0.0431	0.352	0.343	0.331
0.07	1.75	6.99	1.52	0.118	0.895	0.891	0.826
0.10	1.47	6.76	1.43	0.227	1.27	1.25	1.25
0.15	1.46	5.59	1.06	0.366	1.67	1.54	1.42
0.20	1.15	4.97	1.04	0.529	1.77	1.88	1.34
0.25	1.24	4.49	0.812	0.710	1.99	1.96	1.26
0.30	0.977	4.40	0.714	0.902	1.64	1.74	0.938
0.40	0.873	3.54	0.519	1.10	1.58	1.57	0.787
0.60	0.603	1.95	0.270	1.29	1.53	1.42	0.640
0.80	0.468	1.32	0.130	1.47	1.22	1.39	0.479
1.00	0.413	0.926	0.0753	1.63	1.18	1.11	0.389
1.20	0.333	0.562	0.0389	1.77	0.916	1.03	0.313
1.40	0.233	0.385	0.0157	1.88	0.855	0.999	0.290
1.60	0.200	0.230	0.00997	1.96	0.845	0.841	0.237
2.00	0.137	0.112	0.00303	2.00	0.868	0.847	0.214

[a] The data were generated from the equations in Table I, and random normally distributed noise was added.

tributed noise was added to the exact concentration data at the 5% level, *i.e.*, with an expected coefficient of variation of 5%, using a pseudo-random normal generator based on the polar method (8) (Table II). To demonstrate the method, the unit impulse response parameters for Data Set 2 (Table I) were calculated by a nonlinear optimization technique so that Data Sets 4 and 5 (having different input) would not differ significantly. The parameters for Data Set 6 were chosen for the same reason so that Data Sets 4 and 6 would differ significantly but would have the same input (Tables I and II). The 16 sampling times for Data Sets 4–6 were chosen according to a Tschebyscheff sampling (Eq. 20 in Ref. 4) to avoid the "Runge effect" occurring in some linear regressions with fairly equally spaced sampling points.

The data matrix for the linear regression was generated from the simulated data as discussed previously (4). The matrix was executed stepwise with the BMDP-77 linear regression program P9R with a zero-intercept option and the option METHOD = NONE resulting in double-precision computations (9). The regression subsets were chosen on the basis of Mallow's criterion (10).

None of the BMDP linear regression programs provides output of the unscaled variance–covariance matrix required in the statistical computations. Therefore, a program was written to assemble the cross-product matrix and invert it using a Cholesky decomposition. A subroutine was written in FORTRAN IV to calculate and compare the rate and the cumulative amount of input of two inputs. The routine tests for heteroscedatic errors (Eq. 23) and chooses the appropriate type of statistic (Eq. 22 or 24) to calculate the significance of the input difference for various times. The α-values for the t-statistic were calculated according to the algorithm presented by Hill (11). The α-values for the F-statistic were calculated according to algorithms 26.6.15 and 26.2.17 reported in the literature (12). The parameters in Tables III and IV were obtained by nonlinear regression using the interactive program FUNFIT (13). All drawings were done by a penplotter[3] driven by a computer[4] using a software package written by the author.

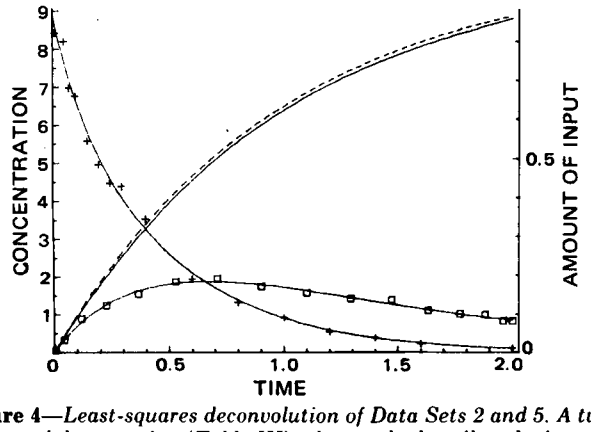

Figure 4—*Least-squares deconvolution of Data Sets 2 and 5. A two-exponential expression (Table III) adequately describes the impulse response data (+). The input function Eq. 1 is estimated subsequently by fitting Eq. 3 to the response data (□) by multiple linear regression. The continuously increasing curve is the calculated input amount (Table V). The broken curve is the exact input.*

RESULTS AND DISCUSSION

Simulated Model—Data Sets 1–3 (Tables I and II) can be considered to simulate, in the classical linear compartmental sense, intravenous bolus data normalized with respect to the dose for three subjects showing two-compartmental pharmacokinetic behavior with respect to a given drug. Data Sets 4–6 (Tables I and II) can be considered similarly as simulated first-order absorption blood concentration data adjusted for a possible lag time for the three subjects.

Comparison of Model-Dependent and Model-Independent Approaches—It is often impossible to administer a drug intravenously. It is of interest to see how accurately the drug input can be determined by a model-dependent approach from absorption blood level data alone, without knowledge of the system's response to intravenous administration of the drug. Data Sets 4–6 are generated from the convolution of $a_1 \exp(\lambda_1 t) + a_2 \exp(\lambda_2 t)$ and $kK \exp(-kt)$. The three data sets are based on:

$$c(t) = kK\left[\frac{a_1}{\lambda_1 + k}(e^{\lambda_1 t} - e^{-kt}) + \frac{a_2}{\lambda_2 + k}(e^{\lambda_2 t} - e^{-kt})\right] \qquad \lambda_1, \lambda_2 < 0 \quad \text{(Eq. 30)}$$

where K corresponds to FD (F = fraction of dose D available) and k corresponds to k_a (the first-order absorption rate constant) in classical linear compartmental modeling. This equation was fitted several times to Data Set 6 using the FUNFIT program (13) with the initial parameter estimates randomly chosen in a feasible parameter space. It is evident (Table IV and Fig. 1) that the parameters for the rate, k, and the extent, K, of input describe the input inadequately, even though the correct model is fitted to the data, the fits are excellent (Table IV and Fig. 1), and the data are as accurate as can be expected in a pharmacokinetic experiment. The input is determined poorly primarily because the parameters a_1, a_2, λ_1, and λ_2 from the unit impulse response $a_1 \exp(\lambda_1 t) + a_2 \exp(\lambda_2 t)$ (the characteristic response) are free to take nearly any values in the

Table III—Least-Squares Unit Impulse Response Parameters Used in Deconvolutions

Data Set	a_1	a_2	λ_1	λ_2
1	1.1055	0.88427	−3.7725	−0.91283
2	1.3170	7.5800	−9.0096	−2.1257
3	0.94748	0.99355	−3.3894	−3.2941

[3] Tektronix 4662.
[4] IBM 370.

Table IV—Estimates of the Input $f(t) = kKe^{-kt}$ obtained in a Model-Dependent Approach by Fitting Eq. 30 to Data Set 6

Fit	K	k	a_1	a_2	λ_1	λ_2	$RSS \times 10^{2a}$
1	0.170	3.31	10.8	6.10	−2.08	−1.13	1.0659
2	11.5	1.46	0.491	0.0797	−3.70	−3.64	1.0758
3	1.85	1.46	3.05	0.500	−3.69	−3.70	1.0758
4	0.233	3.65	9.39	1.85	−1.55	−1.18	1.0733
5	3.17	1.46	1.08	0.992	−3.70	−3.68	1.0758
6[b]	1.04	3.69	0	2.50	0	−1.46	1.0758
True values	3.0	1.5	1	1	−4	−3	

[a] Residual sum of squares. [b] The parameters a_1 and λ_1 were fixed as zero in the fitting procedure corresponding to a one-compartment model.

curve-fitting procedure at the expense of k and K. Furthermore, the nonlinear parameters are not unique due to the presence of multiple minima of the residual sum of the squares function (13). (The six fits in Table IV represent different local minima.)

The input could have been determined more accurately if the unit impulse response parameters had been determined first from an intravenous administration and then used as constants in the fitting of Eq. 30. However, this approach has the disadvantage of being model dependent since it assumes a first-order input $f(t) = kK \exp(-kt)$. Since many variables affect the input *via* the GI route (gastric emptying, GI motility, pH, blood flow, *etc.*), it may seem unreasonable that the input is described in such a smooth, functional form. The present least-squares deconvolution approach is more general. The input is approximated by a polynomial. As such, it is more flexible and can adapt to whatever fluctuations or functional form the input may have. Furthermore, the input is determined by linear rather than nonlinear regression, so it is easier to apply computationally (*e.g.*, no initial parameter estimates are required). The problem with multiple minima is eliminated, and the approach is more suitable for a statistical analysis of the input.

The present method requires the unit impulse response parameters to be determined by nonlinear regression. However, the nonuniqueness of these parameters and their actual values as such have no influence on the accuracy of the input determination. The input is not calculated from the parameters individually but collectively in the way that together they represent the unit impulse response. In estimating these parameters from data from an intravenous bolus or intravenous infusion administration, two or more sets of widely different parameters may be obtained that give nearly superimposed curves as in Fig. 1 (Table IV). However, contrary to a model-dependent approach, the individual parameter values are irrelevant. The important consideration is how well the curve fits the data. There are several methods for automatically handling this curve-fitting problem (14, 15).

Deficiencies in Using C_{max}, t_{max}, and AUC in Bioavailability Testing—Data Sets 4, 5, and 6 are generated from Fig. 2 curves A, B, and C, respectively, by adding 5% normally distributed noise to the exact concentration data (Table II). Curves A and B are very close, but the rate and the extent of input resulting in these curves differ substantially (Table I; A: $k = 1.5$, $K = 3.0$; B: $k = 1.0$, $K = 1.0$). Curves A and C differ widely, but they result from the same input (Table I; $k = 1.5$, $K = 3.0$). A comparison of the extent and the rate of input for Data Sets 4 and 5 (A and B, Fig. 2) on the basis of AUC, t_{max}, and C_{max} will not show a significant difference in their input. Furthermore, a comparison of Data Sets 4 and 6 (A and C, Fig. 2) by AUC, t_{max}, and C_{max} is likely to show a significant difference in both the rate and the extent of input when, in fact, there is no difference.

Because AUC, t_{max}, and C_{max} depend both on the input and the response properties of the system, they can be misleading measures of bioavailability if the comparison is not done within subjects, *i.e.*, for a constant unit impulse response. The deconvolution approach does not have this disadvantage. The response properties of the system are deconvoluted from the measured response to give a pure measure of the rate and the extent of input.

Deconvolution Approach—Deconvolution of Data Sets 4, 5, and 6 (corresponding to A, B, and C in Fig. 2) is done (Figs. 3, 4, and 5, respectively) using the responses measured for the three systems from a bolus input (Data Sets 1, 2, and 3, respectively)[5]. There is excellent agreement between the exact input and that calculated by deconvolution (Figs. 3–5). The regression function, Eq. 3, fits Data Sets 4–6 well. A Durbin–Watson analysis (16) and a runs test (17) show no systematic deviations among the residuals. The Tschebyscheff sampling strategy for Data Sets 4–6 appears to give a more stable behavior for the fitted curve (Eq. 3) and the calculated input for an increase in N (Eqs. 1 and 3) than the sampling scheme used previously (4).

Statistical Comparison of Input—The input rate calculated from Data Sets 1 and 4 and 3 and 6 (A and B, respectively, Fig. 6) does not appear to differ significantly throughout the input period $t = 0$–2. The differential confidence profile (curve C, Fig. 6) is largest in the initial input phase but shows no significant [$100(1 - \alpha) < 95$] difference in the input rate. The same is the case in comparing the cumulative input amount (Fig. 7). The input rates calculated from Data Sets 2 and 5 and 3 and 6 (A and B, respectively, Fig. 8) differ significantly in most of the input period. The respective cumulative input amounts differ significantly during the entire input period (Fig. 9). The input calculated from Data Sets 1 and 4 and 2 and 5 shows the same general behavior as the input calculated from Data Sets 2 and 5 and 3 and 6 (Figs. 10 and 11).

The method proposed performs satisfactorily. The inputs are well determined[6], and the method properly differentiates between the input

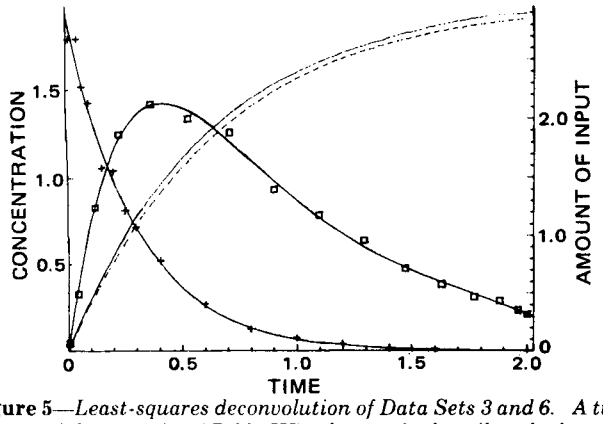

Figure 5—*Least-squares deconvolution of Data Sets 3 and 6. A two-exponential expression (Table III) adequately describes the impulse response data (+). The input function Eq. 1 is estimated subsequently by fitting Eq. 3 to the response data (□) by multiple linear regression. The continuously increasing curve is the calculated input amount (Table V). The broken curve is the exact input.*

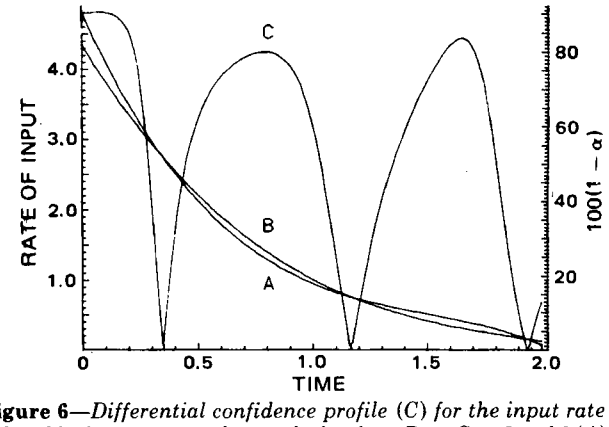

Figure 6—*Differential confidence profile (C) for the input rate calculated by least-squares deconvolution from Data Sets 3 and 6 (A) and 1 and 4 (B).*

[5] An infusion input also could have been used to obtain the unit impulse response parameters.

[6] The input might not have been approximated so well if the analysis had been performed using data extending substantially beyond the absorption phase. However, a beginning oscillatory behavior around zero of $f(t)$ in the postabsorption phase would result, indicating that the fitting is extended too far.

Table V—Input Rate Estimated by the Least-Squares Deconvolution Method, $0 \leq t \leq 2$

Data Sets	
1 and 4	$f(t) = 4.35 - 5.37t + 2.43t^2 - 0.403t^3$
2 and 5	$f(t) = 0.937 - 0.754t + 0.184t^2$
3 and 6	$f(t) = 4.81 - 7.32t + 4.45t^2 - 0.988t^3$

for the similar Data Sets 4 and 5. It also shows no significant difference between the widely different Data Sets 4 and 6 having the same input. However, the statistics appear rather sensitive (Figs. 6 and 7), as expected, since they are based on normality assumptions. Therefore, it would be appropriate to choose a more conservative hypothesis testing such as:

$$H_1: \quad \mu_1|t_0 \leq \mu_2^*|t_0 \qquad \text{(Eq. 31)}$$

against:

$$H_0: \quad \mu_1|t_0 > \mu_2^*|t_0 \qquad \text{(Eq. 32)}$$

$$\mu_2^*|t_0 = \frac{100 + D}{100}\mu_2|t_0 \qquad \text{(Eq. 33)}$$

where D, the differential index, is chosen appropriately. For example, with the differential index equal to 25, the hypothesis that the rate or cumulative amount of input differs by more than 25% could then be tested at various times by a simple extension of the present method. This approach seems to be more rational considering the large intersubject variations observed in the pharmacological response for many drugs. The application of the method to analyze real pharmacokinetic data is in progress.

APPENDIX

Several statistics have been proposed to test the equality of means for normal populations with unequal variances (18–21). Monte Carlo simulation studies comparing several of these statistics (21) indicate that the statistic proposed by Welch (19) is the most robust and appropriate for small samples.

This statistic can be stated for a two population problem in the following simplified form:

Theorem I: Let x_{ij} be the jth observation from the ith population, where $j = 1, \ldots, n_i$ and $i = 1, 2$. If x_{ij} are stochastically independent normal variates with expected values of μ_i and the variance σ_i^2, then the probability distribution of the statistic:

$$w = \frac{\bar{x}_1 - \bar{x}_2}{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^{1/2}} \qquad \text{(Eq. A1)}$$

is approximated by a t-distribution with degrees of freedom given by:

$$DF = \left[\frac{u^2}{n_1 - 1} + \frac{1 - u^2}{n_2 - 1}\right]^{-1} \qquad \text{(Eq. A2)}$$

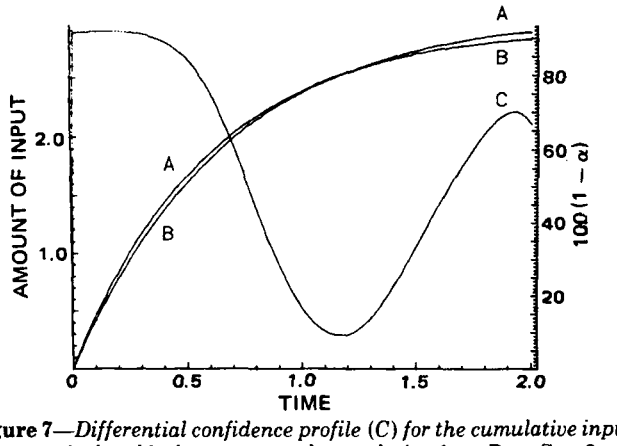

Figure 7—*Differential confidence profile (C) for the cumulative input amount calculated by least-squares deconvolution from Data Sets 3 and 6 (A) and 1 and 4 (B).*

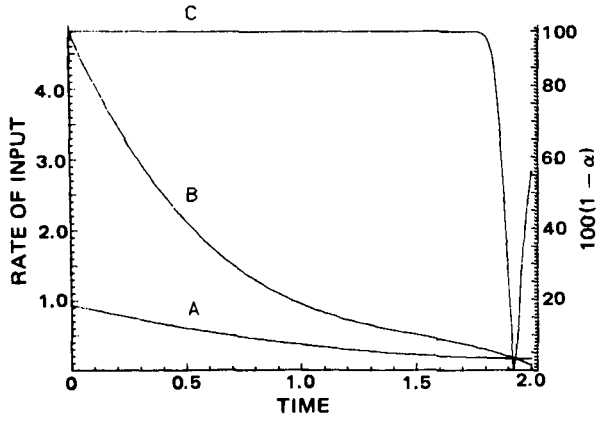

Figure 8—*Differential confidence profile (C) for the input rate calculated by least-squares deconvolution from Data Sets 2 and 5 (A) and 3 and 6 (B).*

where:

$$u = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2} \qquad \text{(Eq. A3)}$$

$$s_i^2 = \sum_{j=1}^{n_i} (\bar{x}_i - x_{ij})^2/(n_i - 1) \qquad i = 1, 2 \qquad \text{(Eq. A4)}$$

and:

$$\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i \qquad i = 1, 2 \qquad \text{(Eq. A5)}$$

This theorem deals with a differently formulated statistical problem than that of interest. However, the following extension of the theorem can be applied to test H_0 (Eq. 10) in the case of unequal population variances:

Corollary I: Let z_1 and z_2 be stochastically independent random variables distributed $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Further, let s_{z1}^2 and s_{z2}^2 be stochastically independent random variables with moment-generating functions:

$$G_1(t) = \left(1 - 2\frac{\sigma_1^2}{\nu_1}t\right)^{-\nu_1/2} \qquad \nu_1 > 0 \qquad \text{(Eq. A6)}$$

$$G_2(t) = \left(1 - 2\frac{\sigma_2^2}{\nu_2}t\right)^{-\nu_2/2} \qquad \nu_2 > 0 \qquad \text{(Eq. A7)}$$

Then the probability distribution of the statistic:

$$w = \frac{z_1 - z_2}{(s_{z1}^2 + s_{z2}^2)^{1/2}} \qquad \text{(Eq. A8)}$$

is approximated by a t-distribution with degrees of freedom given by:

$$DF = \left[\frac{u^2}{\nu_1} + \frac{1 - u^2}{\nu_2}\right]^{-1} \qquad \text{(Eq. A9)}$$

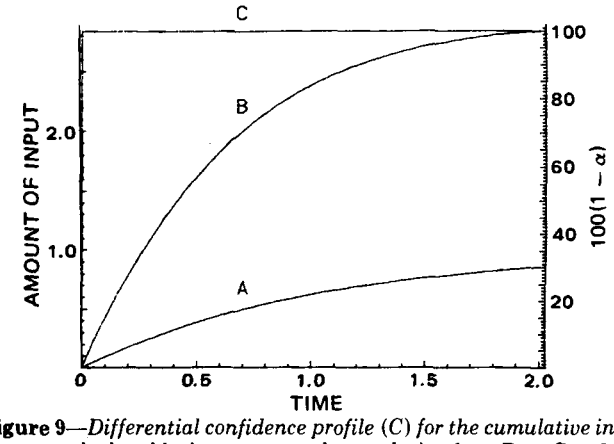

Figure 9—*Differential confidence profile (C) for the cumulative input amount calculated by least-squares deconvolution from Data Sets 2 and 5 (A) and 3 and 6 (B).*

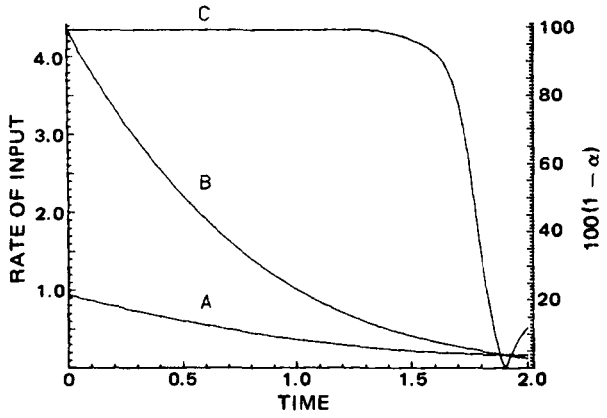

Figure 10—*Differential confidence profile (C) for the input rate calculated by least-squares deconvolution from Data Sets 2 and 5 (A) and 1 and 4 (B).*

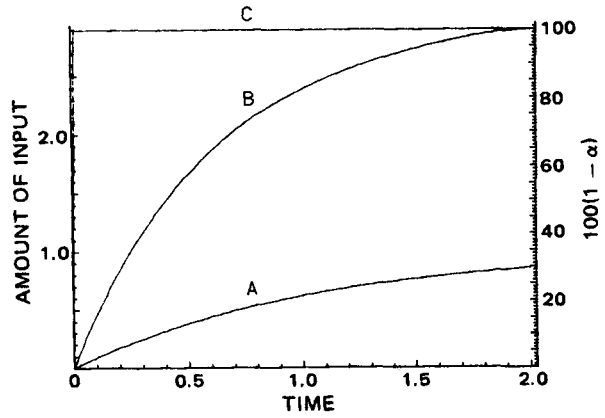

Figure 11—*Differential confidence profile (C) for the cumulative input amount calculated by least-squares deconvolution from Data Sets 2 and 5 (A) and 1 and 4 (B).*

where:

$$u = \frac{s_{z1}^2}{s_{z1}^2 + s_{z2}^2} \qquad \text{(Eq. A10)}$$

Proof: The random variables $z_i = \bar{x}_i$, $i = 1, 2$ (Eq. A5), are distributed $N(\mu_i, \sigma_i^2/n_i)$. The random variables $s_{zi}^2 = s_i^2/n_i$ have the moment-generating functions:

$$G_i(t) = \left(1 - 2\frac{\sigma_i^2/n_i}{n_i - 1}t\right)^{-(n_i-1)/2} \qquad i = 1, 2 \qquad \text{(Eq. A11)}$$

because $(n_i - 1)s_i^2/\sigma_i^2$ is χ^2-distributed with $n_i - 1$ degrees of freedom[7]. Equation A11 uniquely defines the probability distribution of s_{zi}^2. Letting $\nu_i = n_i - 1$ and replacing σ_i^2/n_i by σ_i^2 lead to the corollary.

The applicability of the corollary to test H_0 (Eq. 10) is seen as follows. With:

$$z_m = \hat{f}_m(t_0) \sim N(\mu_m | t_0, \sigma_m^2 \tau_m' \mathbf{V}_m \tau_m) \qquad m = 1, 2 \qquad \text{(Eq. A12)}$$

the moment-generating function of:

$$s_{zm}^2 = \hat{\sigma}_m^2 \tau_m' \mathbf{V}_m \tau_m = \frac{\mathbf{e}_m' \mathbf{e}_m}{M_m - N_m} \tau_m' \mathbf{V}_m \tau_m \qquad m = 1, 2 \qquad \text{(Eq. A13)}$$

is:

$$G_m(t) = \left(1 - 2\frac{\sigma_m^2 \tau_m' \mathbf{V}_m \tau_m}{M_m - N_m}t\right)^{-(M_m - N_m)/2} \qquad m = 1, 2 \qquad \text{(Eq. A14)}$$

because $(M_m - N_m)\hat{\sigma}_m^2/\sigma_m^2$ is χ^2-distributed with $M_m - N_m$ degrees of freedom. Comparison of Eqs. A12 and A14 in relation to Corollary I leads to Eqs. 24–26.

[7] The moment-generating function of a χ^2-distributed variable with n degrees of freedom is $(1 - 2t)^{-n/2}$.

REFERENCES

(1) "Guidelines for Biopharmaceutical Studies in Man," APhA Academy of Pharmaceutical Sciences, Washington, D.C., Feb. 1972.

(2) D. J. Chondos and A. R. DiSanto, "Basis of Bioavailability," The Upjohn Co., Kalamazoo, Mich., 1974.

(3) F. H. Dost, "Grundlagen der Pharmakokinetic," 2nd ed., Thieme, Stuttgart, West Germany, 1968, p. 155.

(4) P. Veng Pedersen, *J. Pharm. Sci.*, **69**, 312 (1980).

(5) J. G. Wagner, "Biopharmaceutics and Relevant Pharmacokinetics," Drug Intelligence Publications, Hamilton, Ill., 1971.

(6) *J. Pharmacokinet. Biopharm.*, **1**, 3 (1973).

(7) A. S. Goldberger, *JASA*, **57**, 369 (1962).

(8) D. E. Knuth, "The Art of Computer Programming," vol. 2, Addison-Wesley, Reading, Mass., 1969, p. 105.

(9) W. J. Dixon and M. B. Brown, "BMDP Biomedical Computer Programs," University of California Press, Berkeley, Calif., 1977.

(10) C. Daniel and F. S. Wood, "Fitting Equations to Data," Wiley, New York, N.Y., 1971, p. 86.

(11) G. W. Hill, *CACM*, **13**, 10 (1970).

(12) M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions," Dover, New York, N.Y., 1970, pp. 932, 947.

(13) P. Veng Pedersen, *J. Pharmacokinet. Biopharm.*, **5**, 513 (1977).

(14) D. W. Kammler, *J. Approximation Theory*, **16**, 384 (1976).

(15) D. W. Kammler and R. J. McGlinn, *J. Computational Appl. Math.*, **4**, 167 (1978).

(16) J. Durbin and G. S. Watson, *Biometrica*, **37**, 409 (1950).

(17) F. S. Sweed and C. Eisenhart, *Ann. Math. Stat.*, **14**, 66 (1943).

(18) B. L. Welch, *Biometrics Bull.*, **1**, 80 (1951).

(19) B. L. Welch, *Biometrica*, **38**, 330 (1951).

(20) G. S. James, *ibid.*, **38**, 19 (1951).

(21) M. B. Brown and A. B. Forsythe, *Technometrics*, **16**, 129 (1974).