J. Pharm. Pharmacol. 1993, 45: 934–936
Communicated August 28, 1992

© 1993 J. Pharm. Pharmacol.

Anomalous parameter estimates in the one-compartment model with first-order absorption

R. D. PURVES, *Department of Pharmacology, University of Otago Medical School, PO Box 913, Dunedin, New Zealand*

Abstract—A difficulty sometimes encountered in least-squares fitting of a one-compartment model with first-order absorption is that estimated values (k_a and k_e) of the rate constants of absorption and elimination are almost identical, are highly correlated and have very large standard deviations. This anomaly is explained by the existence of a class of data sets for which least-squares estimates of the rate constants are complex quantities. Such data sets may arise either from an unfortunate combination of random (e.g. assay) errors in the concentration values if k_a and k_e are sufficiently similar in magnitude, or from delayed absorption.

The concentration-time profile for the one-compartment model with first-order absorption is given by

$$C(t) = \frac{k_a F\ Dose}{V\ (k_a - k_e)} (e^{-k_e t} - e^{-k_a t}) \quad (1)$$

where k_a and k_e are the rate constants for absorption and elimination, V is the apparent volume of distribution and F is the bioavailability. In the least-squares fitting of equation 1 to data, an anomalous outcome is sometimes obtained (Bialer 1980; Chan & Miller 1983; Wijnand 1988) with the estimates of k_a and k_e nearly equal in value. The usual recommendation in such cases, given without discussion of the cause, has been to use the special equation

$$C(t) = \frac{k\ F\ Dose\ t}{V} e^{-kt} \quad (2)$$

which is easily derived from equation 1 as the limiting form with $k_a \rightarrow k_e = k$.

The characteristics of data-sets giving rise to the anomaly have not hitherto been explored; consequently the phenomenon cannot even be reproduced and studied at will. The purpose of this study was to investigate the nature and causes of the anomalous outcome and the range of options available in dealing with it.

Nature of the anomaly

Numerical investigations were made with the Macintosh computer program MINIM[1]. This least-squares program can estimate parameters by Marquardt's method, by singular value decomposition or by the Nelder-Mead polytope (simplex) method, with the fitted function defined explicitly, given as the right-hand side of a differential equation, or defined as a Laplace transform. Fitting criteria used in the present study were ordinary unweighted least-squares (OLS), logarithmic least-squares (OLS[ln]) in which the logarithm of the function is fitted to the logarithm of the data, weighted least-squares with weights inversely proportional to the data (WLS[y^{-1}]), and iteratively reweighted least-squares with weights inversely proportional to the square of the fitted values (IRLS[f^{-2}]).

Anomalous estimates (for example, Table 1) are characterized by some or all of the following, depending on the parameter-estimation algorithm, the floating-point precision, and other details of programming:

—slow convergence with the estimated values of k_e and k_a in equation 1 frequently interchanging their rank order; if an insufficiently powerful algorithm is used, there may be outright failure to converge (Graves et al 1990)

—near identity of the final estimated values for k_e and k_a, typically within 10^{-3}

—strong correlation between parameter estimates as determined from the variance-covariance matrix; typically the absolute values of the correlations are >0.999

—strong linear dependence between (i.e. collinearity of) the columns of the Jacobian matrix J of partial derivatives; as shown by singular value decomposition, the ratio of largest to smallest singular value is typically $> 10^3$; the matrix J^TJ has two equal or nearly equal eigenvalues

—very large estimated standard deviations for the fitted parameters.

Table 1. Example data-set and anomalous OLS parameter estimates.

t	0·3	0·6	0·9	1·2	1·5	2·0	3·0	5·0	7·5	10·0
C	0·13	0·3	0·36	0·37	0·38	0·31	0·2	0·08	0·02	0·01

Parameter	Value	Standard deviation
V/F	1·0111	2638·97
k_a	0·7901	2062·08
k_e	0·7902	2062·28
Dose	1·0000	(fixed)

Singular values of Jacobian matrix: 1·082 0·432 6.38×10^{-6}
Ratio (largest/smallest): 169621

Approximate correlation matrix:

1·0000		
1·0000	1·0000	
−1·0000	−1·0000	1·0000

The anomaly may disappear with minor changes to the data; for example if the first concentration value in Table 1 is changed from 0·13 to 0·18. If equation 2 is fitted in place of equation 1, the anomaly always disappears and the estimated value of k is almost identical with k_e and k_a, but has a small standard deviation.

The difficulties above might be thought to arise from loss of numerical significance in the computation of equation 1, since as $k_a \rightarrow k_e$ the quantities subtracted in both numerator and denominator become nearly identical. However, numerical considerations are easily shown not to be the prime cause of the anomaly, although they strongly influence the closeness of the estimates of k_a and k_e. The occurrence of the anomaly is little affected by the floating-point precision, as shown by trials with single- and double-precision arithmetic, nor can it be circumvented by use of different methods for evaluating the derivatives in the Jacobian matrix (explicit formula or numerical approximation by forward or central differences) or by use of the polytope method, which is derivative-free. Most tellingly, the anomaly is still present when the defining differential equation

$$\frac{dC(t)}{dt} = \frac{k_a F\ Dose}{V} e^{-k_a t} - k_e C(t),\ t > 0;\ C(t) = 0,\ t = 0 \quad (3)$$

is integrated numerically to generate fitted values of C(t), with no reference to its analytical solution (eqn 1).

[1] Details of the Macintosh computer parameter-estimation program MINIM may be obtained from the author.

Cause of the anomaly

It is well known that parameter estimation in equation 1 (with non-anomalous data) may, depending on the initial guesses provided to the estimation algorithm, produce one of two results (with $k_a > k_e$ or $k_e > k_a$), both having the same sum of squared errors and both representing global minima of the fitting criterion. In the absence of other evidence it is conventional to choose the outcome with $k_a > k_e$; if necessary the two values are interchanged and the data refitted to ensure the desired result. A natural inclination then is to suppose that the outcomes of parameter estimation should be classified in three types: $k_a > k_e$ (normal), $k_a = k_e$ (anomalous) and $k_a < k_e$ (flip-flop). On this view the anomalous case arises from a chance equality of two estimated parameters, but strict floating-point equality of parameter estimates (or even equality within 10^{-3}) is so improbable that the anomalous solution would virtually never be observed. A simple numerical experiment shows that anomalous solutions can be common.

Data-sets were generated from equation 2 with F = V = Dose = k = 1 at sample times t = 0·3, 0·6, 0·9, 1·2, 1·5, 2, 3, 5, 7·5 and 10. Each noise-free value was modified by adding to it a pseudorandom normal variate of mean zero and standard deviation equal to 10% of the noise-free value. Twenty such data-sets were generated, and equation 1 fitted to each set by Marquardt's method with OLS, OLS[ln], WLS[y^{-1}] and IRLS[f^{-2}]. The outcome was regarded as anomalous if the ratio of largest to smallest singular value of the Jacobian matrix exceeded 2000. As indicated in Table 2, a substantial fraction of the data sets (roughly 50%) gave anomalous outcomes.

Two further groups of 20 data-sets with the sample times were generated from equation 1 with $F = V = Dose = k_a = 1$ and $k_e = 0{\cdot}5$ or 0·3. The noise-free data-sets were modified to include pseudorandom errors and analysed as above. The number of anomalous outcomes (Table 2) decreased sharply as the value of k_e departed from that of k_a. It is of some interest that the four fitting criteria used here gave different numbers of anomalous outcomes; moreover many data-sets were anomalous only for one or two of the fitting criteria. The last row of Table 2 shows the number of data-sets that were anomalous for all criteria. The smallness of these numbers indicates that many anomalies may be evaded merely by use of a different fitting criterion.

The preceding results show that a sizeable class of anomalous data-sets exists, and that the probability of obtaining an anomalous set depends on the ratio k_a/k_e of the parent noise-free data, but they do not help to identify which property of an individual data-set is responsible for the mismatch with equation 1.

In the course of identifying this property it proves useful to rewrite equation 1 in the form

$$C(t) = \frac{A}{\lambda_2 - \lambda_1}(e^{-\lambda_1 t} - e^{-\lambda_2 t}) \qquad (4)$$

in which the rate constants λ_1 and λ_2 appear symmetrically. Equations 1 and 4 are equivalent in the sense that when parameters A, λ_1 and λ_2 are estimated by fitting equation 4 to data, the sum of squared errors is identical with that obtained from fitting equation 1, and the values and standard deviations of λ_1 and λ_2 are identical with those of k_a and k_e. The Laplace transform of equation 4 is

$$\bar{C}(s) = \frac{A}{s^2 + bs + c} \qquad (5)$$

where A, b and c are real. Three cases may be distinguished.

Case 1. $b^2 > 4c$. The roots of the quadratic expression in the denominator are real and distinct. The inverse transform is given by equation 4 with

$$\lambda_1, \lambda_2 = \tfrac{1}{2}(b \pm \sqrt{b^2 - 4c}) \qquad (6)$$

Case 2. $b^2 = 4c$. There are two coincident roots. The inverse transform has a form similar to that of equation 2:

$$C(t) = A\,t\,e^{-\lambda t} \qquad (7)$$

where $\lambda = b/2$.

Case 3. $b^2 < 4c$. There is a pair of complex conjugate roots. Although correct (real) values may be computed from equations 4 and 6, for example with MINIM's complex arithmetic package, a more convenient form of the inverse transform is:

$$C(t) = \frac{A}{\lambda_{imag}} \sin[\lambda_{imag}\,t]\,e^{-\lambda_{real} t} \qquad (8)$$

where

$$\lambda_{real} = b/2 \quad \text{and} \quad \lambda_{imag} = \tfrac{1}{2}\sqrt{4c - b^2} \qquad (9)$$

The three cases may be combined if coefficients A, b and c above are treated as parameters to be estimated, either directly by numerical inversion of the Laplace transform in equation 5, or by a function definition which specifies C(t) to be equation 4 (Case 1), 7 (Case 2) or 8 (Case 3) according to the current values of b and c. When such a function was fitted to the 3 groups of 20 data-sets described earlier, a remarkable result was obtained: no outcome was anomalous. Every data-set that previously gave a normal outcome from fitting equation 1 belonged to Case 1, no data-set belonged to Case 2, and every data-set that previously gave an anomalous outcome in equation 1 belonged to Case 3. Furthermore, the goodness-of-fit (as determined by the smallness of the sum of squares) of Case 3 outcomes was in every

Table 2. Number of data-sets (out of 20) giving anomalous parameter estimates in equation 1.

	k = 1	$k_a = 1$ $k_e = 0{\cdot}5$	$k_a = 1$ $k_e = 0{\cdot}3$
OLS	11	7	1
OLS[ln]	7	1	0
WLS[y^{-1}]	10	3	1
IRLS[f^{-2}]	8	1	0
Anomalous with all fitting criteria	3	0	0

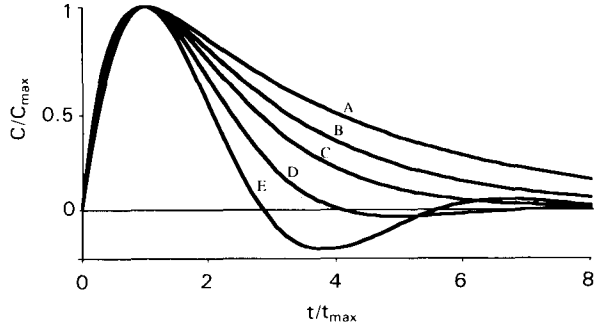

FIG. 1. Concentration profiles predicted by equations 5, 7 and 8. The curves are normalized to the same t_{max} and C_{max}. Curve A: $\lambda_1 = 4\lambda_2$. Curve B: $\lambda_1 = 2\lambda_2$. Curve C: $\lambda_1 = \lambda_2 = \lambda$. Curve D: $\lambda_{imag} = 0{\cdot}5\lambda_{real}$. Curve E: $\lambda_{imag} = \lambda_{real}$. Curves A and B correspond to Case 1, C to Case 2, and D and E to Case 3.

Table 3. Number of data sets out of 20 (with $k_a = 1$, $k_e = 0{\cdot}5$) giving anomalous WLS[y^{-1}] parameter estimates in equation 1 and equation 10.

T_{lag}	Equation 1	Equation 10
0·0	3	3
0·1	13	3
0·2	18	3
0·3	20	3

instance better than the corresponding anomalous outcomes in equations 1 or 4. Thus the property of an individual data-set that is responsible for the anomaly is that the least-squares estimates of the rate constants in equation 4 are complex quantities.

The concentration profiles due to equations 5, 7 and 8 (Fig. 1) give guidance to the nature of data-sets that produce anomalous parameter estimates. Data whose profile approximates curves A or B in Fig. 1 are likely to be normal. Although the full time-course of curves D or E is not observable (since negative concentrations are inadmissable), data approximating the positive portions of these curves is likely to be anomalous. Indeed any set of three or more noise-free data points calculated from equation 8 (with $\lambda_{imag} \neq 0$) at distinct positive values of t is found to be anomalous by any fitting criterion. Thus equation 8 provides for the first time a numerical recipe for generating anomalous data sets.

A useful prediction can be based on the observation that the anomalous curves have a small dispersion or coefficient of variation of the residence times, defined by $CVRT = VRT^{1/2}/MRT$, where VRT is the variance of the residence times and MRT is the mean residence time (Yamaoka et al 1978). The CVRT of curve C of Fig. 1 is $1/\sqrt{2}$. This is evidently a critical value: data-sets with a CVRT less than this produce an anomalous outcome. One circumstance in which data-sets with small CVRT values are obtained arises when an absorption lag time T_{lag} is ignored or underestimated, since T_{lag} increases MRT without affecting VRT. If the preceding speculation is soundly based, it should be possible to make a previously normal data-set become anomalous by shifting its concentrations to later times. To check this prediction, the 20 data-sets above with $k_a = 1$, $k_e = 0{\cdot}5$ were modified by adding to each value of t a lag time of 0·1, 0·2 or 0·3. Table 3 shows the number of anomalous outcomes of parameter estimation by WLS[y^{-1}]. The larger values of lag time cause a substantial increase in the fraction of anomalous outcomes, and it is readily found by experiment that any data-set whatsoever may be made anomalous for any fitting criterion by applying a sufficiently large lag time. The prediction is therefore confirmed. As shown in the last column of Table 3, the anomaly may be removed in most instances by the inclusion of T_{lag} as a parameter to be estimated:

$$C(t) = \frac{k_a\,F\,Dose}{V\,(k_a - k_e)}\{e^{-k_e(t-T_{lag})} - e^{-k_a(t-T_{lag})}\},\ t > T_{lag}$$
$$= 0,\ t \leqslant T_{lag}. \qquad (10)$$

Discussion

It has been shown in this study that anomalous parameter estimates in equation 1 result from the estimation of real rate constants in data-sets for which the least-squares values in equation 4 are complex. This circumstance may arise either from an unfortunate combination of random (e.g. assay) errors in the data if k_a and k_e are sufficiently similar in magnitude, or from delayed absorption.

The appearance of complex rate constants in equation 4 (or equivalently the use of equation 8) does not, however, correspond to any physically interpretable pharmacokinetic model, as is obvious from the negative portions of the damped oscillatory profiles D and E in Fig. 1. In practice it will, therefore, be preferable to use equations 1, 2, 5 or 7 to estimate the rate constants even in the anomalous cases, with implicit acceptance of a worse fit as the price of obtaining real values. Equation 8 is therefore unlikely to see service for the analysis of actual pharmacokinetic data. Its significance lies in providing a numerical recipe for generating data-sets all of which are anomalous, and in allowing curves D and E to be drawn in Fig. 1. In consequence, the results of Table 2 are easily understood. The addition of random errors to data-sets generated from equations 1 or 2 evidently causes a proportion of the sets to fall on the inappropriate side of curve C.

The above analysis leads to a further numerical recipe for generating anomalous data-sets, namely equation 10, and the surprising result that the anomaly may often be removed by inclusion of an extra parameter to be estimated—the opposite of the usual approach in which a parameter is discarded on replacing equation 1 with equation 2.

Thus the principal options available when equation 1 gives rise to anomalous estimates are: attempted evasion of the anomaly by trial of different fitting criteria; attempted removal of the anomaly by inclusion of an absorption lag-time parameter; acceptance of the (very close) values of k_a and k_e; and the fitting of equation 2 in place of equation 1, thus obtaining a single rate constant k of small standard deviation.

It does not seem possible to offer general advice as to which option should be pursued. Since their fitted concentration values are virtually identical, and $k_a \approx k_e \approx k$, the last two options are in most respects equivalent, but the last has the advantage of avoiding numerical difficulties and thereby assuring rapid convergence of the estimation algorithm.

An alternative approach to the anomaly has been suggested (Niedzwiecki & Simonoff 1990) in which the right-hand side of equation 1 is reformulated as $(\phi_1 + \phi_2 t^{-1})\,e^{-\phi_3 t}$. This approach is clearly invalid, since if $\phi_2 \neq 0$ the expression is unbounded as $t \to 0$. Any useful model of absorption kinetics must have the property $C \to 0$ as $t \to 0$.

References

Bialer, M. (1980) A simple method for determining whether absorption and elimination rate constants are equal in the one-compartment open model with first order processes. J. Pharmacokin. Biopharm. 8: 111–113

Chan, K. K. H., Miller, K. W. (1983) Nonlinear regression approach for determining whether absorption and elimination rate constants are equal in the one-compartment open model with first-order processes. J. Pharm. Sci. 72: 574–576

Graves, D. A., Locke, C. S., Muir, K. T., Miller, R. P. (1990) The influence of assay variability on pharmacokinetic parameter estimation. J. Pharmacokin. Biopharm. 17: 571–592

Niedzwiecki, D., Simonoff, J. S. (1990) Estimation and inference in pharmacokinetic models: the effectiveness of model reformulation and resampling methods for functions of parameters. J. Pharmacokin. Biopharm. 18: 361–377

Wijnand, H. P. (1988) Pharmacokinetic model equations for the one- and two-compartment models with first order process in which the absorption and exponential elimination or distribution rate constants are equal. J. Pharmacokin. Biopharm. 16: 109–128

Yamaoka, K., Nakagawa, T., Uno, T. (1978) Statistical moments in pharmacokinetics. J. Pharmacokin. Biopharm. 6: 547–558