Analysis and comparison of sigmoidal curves: application to dose-response data

J. B. MEDDINGS, R. B. SCOTT, AND G. H. FICK
Departments of Medicine, Pediatrics, and Community Health Sciences, University of Calgary, Calgary, Alberta T2N 4N1, Canada

MEDDINGS, J. B., R. B. SCOTT, AND G. H. FICK. *Analysis and comparison of sigmoidal curves: application to dose-response data.* Am. J. Physiol. 257 (Gastrointest. Liver Physiol. 20): G982–G989, 1989.—A number of physiological or pharmacological studies generate sigmoidal dose-response curves. Ideally, data analysis should provide numerical solutions for curve parameters. In addition, for curves obtained under different experimental conditions, testing for significant differences should be easily performed. We have reviewed the literature over the past 3 years in six journals publishing papers in the field of gastrointestinal physiology and established the curve analysis technique used in each. Using simulated experimental data of known error structure, we have compared these techniques with nonlinear regression analysis. In terms of their ability to provide accurate estimates of ED_{50} and maximal response, none approached the accuracy and precision of nonlinear regression. This technique is as easily performed as the classic methods and additionally provides an opportunity for rigorous statistical analysis of data. We present a method of determining the significance of differences found in the ED_{50} and maximal response under different experimental conditions. The method is versatile and applicable to a variety of different physiological and pharmacological dose-response curves.

data analysis; curve fitting; drug dose-response; smooth muscle contractility; intestinal transport

IN THE FIELD of gastroenterology, physiological investigations of smooth muscle contractility, intestinal transport, and the effect of hormones generate sigmoidal dose-response curves. In the majority of cases these curves are assumed to follow a hyperbolic relationship; each curve can be described by two parameters, the maximal response (T^m) and the concentration that produces a half-maximal response (ED_{50}, Ref. 11). For the physiologist, accurately estimating these two parameters is often crucial; however, the most efficient means by which to perform this analysis is not readily apparent. It is now clear that simply inspecting such data and generating a subjective estimate of T^m and ED_{50} is not adequate because no objective estimate of error can be made. In response to this, a wide variety of data analysis techniques have emerged (6, 11–13). These include transformations that linearize the dose-response relationship, thereby allowing the use of simple linear regression techniques or directly fitting data to the original relationship with nonlinear regression techniques. The relative merits of these two different approaches are not well established despite several publications suggesting that nonlinear regression analysis is the more useful technique for analysis of transport processes (2, 3). The problem becomes even more complex when two curves are considered, as the investigator invariably wishes to know if the curves are different from each other and, if so, is it the ED_{50} or the T^m that is significantly different. Once again it has been unclear how to approach this perplexing but common problem. It is common practice today to evaluate differences between curves by performing multiple comparisons of the response at different concentrations, a technique that cannot possibly take advantage of all the information contained within the curve (6).

Therefore, the aims of this paper are *1*) to review the methods currently utilized by gastrointestinal physiologists to analyze dose-response curves, *2*) determine the accuracy and precision of currently utilized linear and nonlinear methods for curve analysis, and *3*) present a method to statistically evaluate differences between families of dose-response data obtained under different experimental conditions.

METHODS

A Medline search using the key words smooth muscle, contractile response, drug, and dose-response was performed covering the years 1986–1989. To ensure some uniformity of quality the search was limited to six journals that commonly publish papers concerned with gastrointestinal physiology. These were the *American Journal of Physiology*, *Canadian Journal of Physiology and Pharmacology*, *Digestive Diseases and Sciences*, *Gastroenterology*, *Journal of Clinical Investigation*, and the *Journal of Pharmacology and Experimental Therapeutics.* Papers were classified as to whether data were presented either *1*) descriptively as a dose-response curve, *2*) graphically with visual estimation of maximal response and ED_{50}, *3*) transformed for linear regression analysis (linear regression analysis of the linear portion of sigmoidal

curve, logit, probit, or the Lineweaver-Burk double reciprocal plot), or *4*) whether nonlinear regression analysis was employed.

Second, to critically evaluate these techniques several sets of simulated experimental data were generated. Each set was derived from hypothetical values that closely approximated actual experimental data describing the tension produced in a strip of jejunal muscle under the stimulation of graded doses of carbachol. For the purposes of this experiment the tissue was assumed to generate tension as a hyperbolic function of the carbachol dose, an assumption commonly employed in such studies. Therefore, the tension (T) at any dose of carbachol (C) is given by

$$T = \frac{T^m C}{ED_{50} + C} \qquad (1)$$

In this equation T^m represents the maximal tension developed and ED_{50} represents the dose that produces half-maximal tension. The true maximal response was predefined as 140 g and the true ED_{50} as 0.016 μM. For the hypothetical experiment, a dose range typical of those used in the literature was selected ranging from 10^{-8} to 10^{-4} M. Nine concentrations were selected within this range and the tension of the tissue at each concentration calculated from *Eq. 1*. These nine perfect data points represented the basis of each data set, which were constructed by generating 20 random numbers within a well-defined range (discussed below) around each perfect data point. Normal random numbers were generated using the statistical software package Systat (Systat, Evanston, IL). Therefore, each data set had 20 observations of tension at each of nine concentrations of carbachol. Six data sets were selected to evaluate the most common error structures seen in experimental data. These included fixed ranges of tension values at each concentration of either 20 or 40% of the perfect point. A third data set was constructed with variable ranges that increased linearly from 20% at the lowest concentration to 40% at the highest. Because experimental data invariably includes outliers, these three data sets served as starting material for a further three. These had the same data points, but for every concentration 10% of the tension values were randomly selected to be outliers. Outliers had a value three times the standard deviation of the sample, added or subtracted at random. Therefore, the six data sets used were *1*) range = 20%, *2*) range = 40%, *3*) range = variable, *4*) range = 20% plus outliers, *5*) range = 40% plus outliers, and *6*) range = variable plus outliers. Each set contained 180 data points.

Data sets were evaluated by four analytical methods. In each case the best estimates of T^m and ED_{50} were obtained and 95% confidence intervals constructed for each parameter. For transformations in which the parameter in question was a ratio of the coefficients, Fieller's theorem (1, 8, 9) was used to construct 95% confidence intervals. All calculations were performed using a microcomputer and commercially available statistical software (Systat and Minitab; Minitab, State College, PA).

RESULTS

Prevalent methods. The Medline search identified 74 articles published since January 1986 in the journals selected. Of these, 27% presented dose-response data graphically. Most simply discussed the data in descriptive or qualitative terms; only 8 of 20 attempted to test for differences between curves. This attempt was always limited to testing individual data points in each curve by either a *t* test or the analysis of variance.

Thirty-five percent of the papers cited visually determined the ED_{50} for each curve from a graphical presentation of the data. Of this group, <25% attempted a statistical evaluation of differences between curves. In all cases this was once again performed by comparing the tension at a single constant concentration in the two curves and evaluating the difference with a *t* test.

Only 24% of the papers quoted evaluated their data with an aim to establishing a T^m and ED_{50} utilizing one of the linear transformations. Of these approximately one quarter performed a regression on the linear portion of the sigmoid curve to establish the ED_{50}. Another quarter utilized either probit or logit analysis, again to establish an ED_{50}. The remaining investigators used a double reciprocal transformation (Lineweaver-Burk), with most citing the methodology of Tallarida and Murray (14).

A minority, just 14% of authors, utilized nonlinear regression analysis to estimate curve parameters.

Evaluation of methods. The six data sets outlined in the METHODS section were used to assess the ability of four techniques to establish precise and accurate estimates for both parameters in the presence of known error structures. Figure 1 illustrates the relationship between tension and carbachol dose for each data set. Figure 1*A* contains the data set named range = 20%. With an allowed range of ±20% around each data point, it is the tightest of the data sets. Figure 1*B* contains range = 40% and is equivalent to Fig. 1*A* except that the allowed range on each side of the data point is now 40%. Range = var (Fig. 1*C*) represents a composite of the first two. At low concentrations of agonist the data points are tightly grouped but with increasing concentration the range increases, a pattern reminiscent of experimental data. Figure 1, *D*, *E*, and *F*, represent the same data points found in Fig. 1, *A*, *B*, and *C*, respectively, with the exception that 10% of the points at each concentration were randomly selected to be outliers and had a value of three times the standard deviation of the sample either added or subtracted at random.

The techniques evaluated were selected to include the methods prevalent in the literature.

1) Probit analysis as described by Tallarida and Jacob (13). In this analysis, data for tension are expressed as a percentage of maximum tension developed. Each point is then converted to the corresponding probit and graphed as a function of log C. Linear regression analysis of the linear portion of the curve (identified visually in most cases) yields an equation from which the concentration that would produce a probit of 5.0 can be calculated. This value is taken to be the ED_{50}. It should be pointed out that implicit to this analysis is the acknowl-

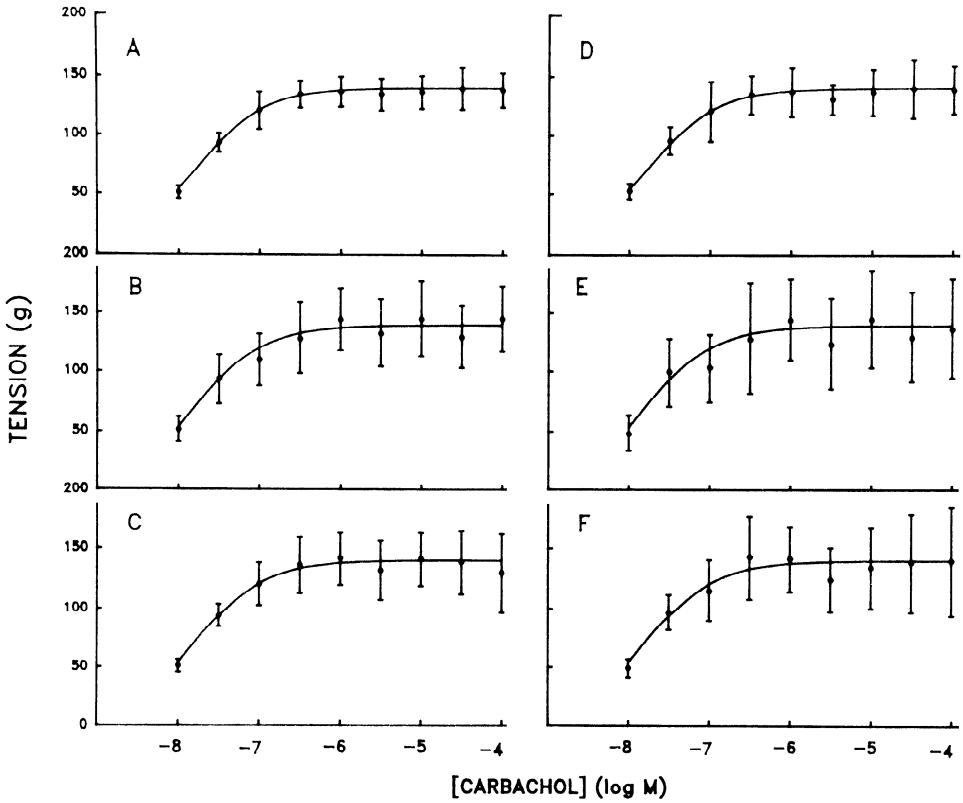

FIG. 1. Ranges for experimental data. In each case data were derived from a perfect data set where T^m and ED_{50} took the values of 140 g and 0.016 μM, respectively. The function described by the perfect data is represented by the solid line in each panel. *A* (range = 20%): the allowable range around each perfect point did not exceed ±20% of the perfect value. The mean of each point is shown and the error bars represent standard deviations. *B* (range = 40%): similar to *A* except that the allowable range was ±40%. *C* (range = var): allowable range that increases from ±20 to ±40% in a linear fashion over the concentration range shown. *D*, *E*, and *F*: same data as *A*, *B*, and *C*, respectively, except that 10% of the points at each concentration were changed to outliers by either randomly adding or subtracting a value equal to 3 standard deviations of the original data.

edgment of a maximal response estimated visually or by averaging the tensions generated by the highest doses of carbachol. The latter route was followed in our analysis with the 40 tensions generated by the two highest doses being averaged to obtain the maximal response from which probits were calculated. Because the error involved in this calculation affects the resulting line but is not incorporated into the standard calculations that follow, it is not clear how to construct confidence intervals for the ED_{50}. Therefore, for this group of data no confidence intervals have been reported.

2) Double reciprocal transformation. By taking the reciprocal of both sides of *Eq. 1* the following relationship emerges

$$\frac{1}{T} = \frac{1}{T^m} + \frac{ED_{50}}{T^m}\left(\frac{1}{C}\right) \qquad (2)$$

By plotting 1/T vs. 1/C a straight line is obtained with the *y*-intercept equaling $1/T^m$ and the slope representing ED_{50}/T^m. Linear regression analysis of the transformed data directly yields T^m and its 95% confidence intervals. The ED_{50} can be calculated from this value and the slope of the line, while confidence intervals may be constructed using Fieller's theorem, as described in the METHODS section. One important point concerning this transformation should be appreciated. If the error in tension at each concentration is normally distributed to begin with, it cannot remain so in the transformed state.

3) Modified Eadie-Hofstee transformation (4, 7). Although not used extensively in dose-response studies, this transformation is commonly used in the transport literature. *Equation 1* can be rearranged to give

$$\frac{T}{C} = \frac{T^m}{ED_{50}} - \left(\frac{1}{ED_{50}}\right)T \qquad (3)$$

Plotting T/C vs. T, therefore, results in a line with slope equal to $-1/ED_{50}$ and intercept of T^m/ED_{50}. Additionally, the axes are interchangeable and T is often plotted as a function of T/C. These parameters and their confidence intervals can then be calculated in a manner analogous to *method 2.*

4) Nonlinear regression. With the advent of powerful microcomputers and statistical software, the ability to fit experimental data directly to nonlinear equations is available to all investigators. For the purpose of this analysis data were fitted directly to *Eq. 1* and the program allowed to converge on a solution that minimized the weighted sum of squared residuals. Data points were weighted in proportion to the reciprocal of the within concentration estimates of variance. In general, convergence occurred within 20 iterations using the nonlinear regression module of Systat.

Analysis of available methods. Table 1 illustrates the results of the data analysis for each method on all six data sets. A perfect analysis would return the starting values for T^m and ED_{50} of 140 and 0.016, respectively. Although probit analysis gave acceptable values for the ED_{50} in all data sets, it had two serious drawbacks. First, the T^m has to be evaluated separately from visual inspection of the data, which introduces subjective bias. In this analysis all 40 data points at the two highest concentrations were averaged, which helped to overcome this problem. However, the 95% confidence intervals remained unacceptably large. Second, since the value of

TABLE 1. *Parameter estimates by different analysis methods*

Method	Parameter	Data Set					
		Range = 20%	Range = 40%	Range = Var	Range = 20% + outliers	Range = 40% + outliers	Range = Var + outliers
Probit	T^m	138.7	137.3	134.1	138.7	133.2	139.1
		(131.1, 146.3)	(124.4, 150.1)	(120.1, 148.1)	(128.2, 149.3)	(114.7, 151.7)	(118.9, 159.3)
	ED_{50}	0.016	0.017	0.016	0.016	0.017	0.016
Double reciprocal	T^m	137.6	131.5	134.3	134.2	107.7	125.5
		(134.4, 141.1)	(124.9, 138.8)	(129.8, 139.1)	(129.0, 139.8)	(86.4, 143.0)	(114.5, 138.7)
	ED_{50}	0.017	0.017	0.016	0.016	0.017	0.016
		(0.016, 0.018)	(0.015, 0.019)	(0.015, 0.018)	(0.014, 0.017)	(0.009, 0.030)	(0.012, 0.019)
Eadie-Hofstee	T^m	143.3	157.1	151.4	147.6	179.8	165.5
		(140.3, 146.6)	(148.8, 167.7)	(145.6, 158.3)	(142.8, 153.2)	(161.4, 209.8)	(154.8, 179.6)
	ED_{50}	0.021	0.034	0.028	0.024	0.058	0.042
		(0.019, 0.023)	(0.029, 0.041)	(0.025, 0.033)	(0.022, 0.033)	(0.045, 0.083)	(0.035, 0.052)
Nonlinear	T^m	138.8	138.2	140.2	136.5	135.2	138.4
		(136.1, 141.5)	(134.1, 142.3)	(134.5, 145.8)	(133.1, 140.0)	(127.9, 142.5)	(131.2, 145.6)
	ED_{50}	0.016	0.017	0.017	0.015	0.017	0.017
		(0.015, 0.017)	(0.016, 0.018)	(0.016, 0.018)	(0.014, 0.016)	(0.013, 0.021)	(0.016, 0.018)

Values are parameter estimates calculated by the 4 methods listed. In each case the starting values were T^m, 140 and ED_{50} = 0.016. Values are the best estimate for each method. Underneath each estimate in parentheses are the upper and lower 95% confidence intervals. Significance of the data is discussed in the text.

T^m is critical to the evaluation of the ED_{50}, it was unclear how to evaluate error in the determination of the ED_{50}. Thus, although useful to approximate parameter estimates, probit analysis is of little use for critical evaluation.

The double-reciprocal transformation is one of the most commonly used transformations in the dose-response literature. From inspection of Table 1 it is evident that this technique was exquisitely sensitive to error in the data. With relatively tight data, such as presented with range = 20%, this technique provided useful approximations of the parameters. However, with increasing variability in the experimental data, especially with the addition of outliers, this technique became misleading.

The second linearizing transformation presented in Table 1, the modified Eadie-Hofstee plot, has the same drawbacks as the double-reciprocal transformation. In fact, estimates derived from this procedure were even more sensitive to experimental error.

Of the four methods presented in Table 1 only nonlinear regression provided accurate and precise parameter estimates under all types of experimental error. In all cases the 95% confidence interval bracketed the known correct value for each parameter and in all cases the range was relatively narrow, allowing confidence in the result.

Defining differences between curves. In most physiological studies, the ultimate goal of the investigator is not to determine an absolute value for T^m and ED_{50} but rather to tell if two curves are different from each other. Furthermore, given that two curves are different, the investigator would usually like to decide if either the T^ms or the ED_{50}s or both are different. In the literature surveyed, most authors have attempted to perform this analysis by looking at individual values for tension at well-defined concentrations and evaluating whether an experimental group has a different tension from controls. Although this method is valid for carefully picked concentrations it clearly does not utilize all the information inherent in a well-performed dose-response curve. The previous section gives evidence that nonlinear regression analysis is an effective method of analysis for the type of data commonly encountered in physiological experiments. Next, we outline how this method lends itself to the construction of tests and confidence intervals for comparisons of parameter estimates.

Consider the case of two experimental groups that differ in T^m but have constant ED_{50}s. If all data points from these two experiments are grouped, we can fit the combined data to an equation containing either a single T^m or one containing two T^ms. Because the data contain two distinct T^ms the model that allows for this will fit significantly better than the one that doesn't. Consequently, the residual sum of squares will be far less for this model than for the one containing only one T^m. A more rigorous approach can be stated as follows.

Equation 1 describes the situation that exists when only one ED_{50} and T^m give rise to the observed data. Suppose two separate experiments (*A* and *B*) were performed and we wished to know if either the ED_{50} or T^m calculated for *experiment B* was different from that calculated for *experiment A*. This question can be approached in a simple manner. First, identify all pairs of T and C with the grouping variable δ that takes the value of zero for data from *experiment A* and the value of one for *experiment B*. Then fit the combined data to an equation that offers the choice of a second ED_{50} or T^m. With the appropriate use of δ this option can be limited to only one of the data sets as follows

$$T = \frac{(T^m + \alpha\delta)C}{ED_{50} + \beta\delta + C} \qquad (4)$$

Therefore, *Eq. 4* now allows all possibilities. The T^m for *experiment A* (where $\delta = 0$) becomes T^m and differs from the T^m assigned to *experiment B* by the factor α. Of course α can assume the value of zero, which would imply that a second T^m is not required to adequately fit the

data. The same holds for the relationship between the two ED_{50}s, the difference between these is determined by the value of β.

Because this equation covers all eventualities, it will fit the data better than a relationship that does not contain one or more of the parameters. Importantly, if removing one of these options produces a relationship that does not adequately describe the data, the fit will be significantly impaired. This will manifest itself as a significant increase in the weighted sum of squared residuals. To test whether two ED_{50}s or T^ms are required to fit the data, these choices are sequentially removed and the significance of this judged by assessing the difference in the weighted sum of squared residuals with an F test

$$F = \frac{SSR_2 - SSR_1}{\left(\frac{SSR_1}{df}\right)} \qquad (5)$$

In this case SSR_1 refers to the residuals generated by fitting to *Eq. 4*, which will always be the lowest because this represents the most general model. SSR_2 refers to the residuals from the second model, and df represents the available degrees of freedom. If removing a parameter significantly impairs the fit of the data SSR_2 will be much greater than SSR_1 and a significant F value will result. To examine the hypothesis that a second T^m is not required the data would be fitted to

$$T = \frac{(T^mC)}{[ED_{50} + (\beta\delta) + C]} \qquad (6)$$

and to test the hypothesis that the second ED_{50} is not required we would fit to

$$T = \frac{(T^m + \alpha\delta)C}{ED_{50} + C} \qquad (7)$$

How significant these parameters were to the overall fit is then judged by the appropriate F test. Next, the remaining parameter is removed by fitting the data to *Eq. 1*. Thus differences in both parameters can be tested for in a simple manner. The assumption underlying this testing process is that *Eq. 5* has an F distribution. Recent statistical work suggests that this assumption is correct (5) and although with all nonlinear models the distributional result is approximate, for this equation it does approximate an F distribution.

To provide an example of this method and to demonstrate its power, two new data sets were generated. These are shown graphically in Fig. 2. Figure 2*A* represents a data set identical to that shown in Fig. 1*F* with the exception that for this case ED_{50} was increased to 0.024. Therefore, this data set differs from the original only in one value used to generate the original true data points. In addition it still has an error structure common to physiological experiments, the one we have denoted range = variable + outliers. Figure 2*B* is analogous; however, in this case the ED_{50} used to generate the true data points remained at 0.016 but the T^m was increased to 168 from 140.

These new data sets were analyzed by the nonlinear regression technique previously described and the results reported in Table 2 as *data sets 2* and *3*. *Data set 1* represents the original data reported in Fig. 1 and Table 1. It is apparent that once again this technique gave accurate and precise parameter estimates with confidence intervals that bracketed the known true values. The problem is whether or not the values reported for the T^m and ED_{50} for *data sets 2* and *3* are significantly different from those observed in *data set 1*. This question can be addressed by the technique just outlined.

The statistical analysis of this data is presented in complete form as Table 3. First, considering the comparison of *data set 2* with *data set 1*. Inspection of the results revealed that although the T^ms may be different, it appeared unlikely that the ED_{50}s were different. Therefore, the first step was to fit the combined data of *sets 1* and *2* to *Eq. 4*. As illustrated this resulted in the generation of four parameter estimates. Importantly, α, which represented the difference between the two T^ms, was a positive number, and the 95% confidence intervals for this estimate did not include zero. β, which represented the potential difference between ED_{50}s, was small and the confidence intervals constructed for this parameter included zero, implying that it was unlikely to be different from zero. This could be tested in two ways. Either the parameter value and estimated error could be used or the overall fit of the data to models with and without the parameter measured. Statisticians have recently emphasized that the latter technique is more appropriate, especially in nonlinear regression analysis (5). Because the sum of squared residuals is presented by most, if not all, nonlinear regression routines it is a simple matter to analyze the data in this manner. The weighted residual sum of squares for this analysis (SSR_2) was 375.35, as shown in Table 3. Because a difference in ED_{50} appeared unlikely from inspection of the data in Table 3, this was tested by fitting the data to *Eq. 7*. As shown, with removal of β, the parameter allowing a second ED_{50}, the weighted residual sum of squares (SSR_1) increased to only 375.44. From *Eq. 5* the F statistic was calculated as 0.085, suggesting that a second ED_{50} was not required to model these data ($P = 0.77$). Next, to evaluate the possibility that two T^ms were present we removed this option by fitting the data to *Eq. 1*. SSR_2 from this analysis was 478.17, the F statistic of 97.4 providing strong support for the presence of a second T^m ($P < 0.00001$).

Inspection of the calculated parameter estimates presented in Table 2 for *data sets 3* and *1* suggested that we would be unlikely to find a difference in T^ms but the ED_{50}s of these two data sets might be different. The combined data was fit to *Eq. 4* and the resultant parameter estimates with their constructed 95% confidence intervals presented in Table 3. These data supported our original hypothesis because it appeared that α would be unlikely to be different from zero. However, β also had confidence intervals close to zero, making it unclear whether we would be able to statistically distinguish two ED_{50}s. The weighted residual sum of squares (SSR_2) for this model was 365.84. Because a difference in T^m seemed so unlikely, the data was fit to *Eq. 6*, thus removing the choice of a second T^m. This fit was excellent, with SSR_1

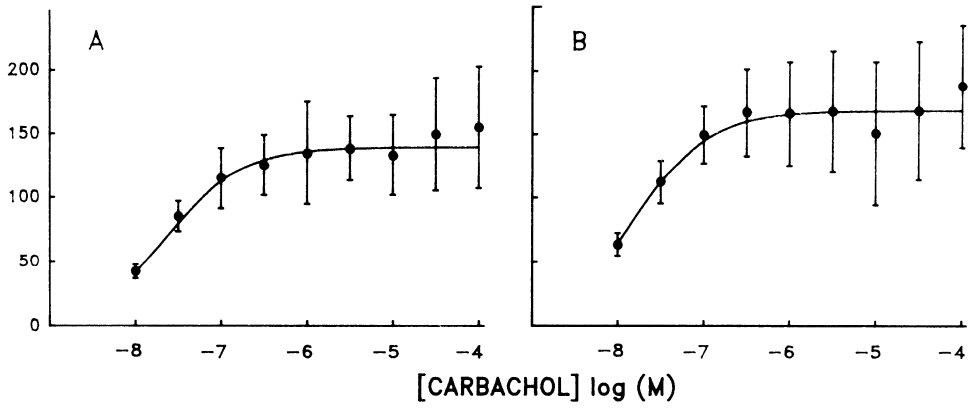

FIG. 2. Parameter estimation data. *A*: data set generated with the same constraints as Fig. 1*F* except that ED_{50} = 0.024 and T^m = 140. *B*: similar to *A* except that ED_{50} = 0.016 and T^m = 168.

TABLE 2. *Nonlinear regression parameter estimates*

Data Set		Known Parameters	Calculated Parameters
1	T^m	140	138.4 (131.2, 145.6)
	ED_{50}	0.016	0.017 (0.016, 0.018)
2	T^m	168	172.2 (167.6, 176.8)
	ED_{50}	0.016	0.017 (0.015, 0.019)
3	T^m	140	140.4 (134.6, 146.2)
	ED_{50}	0.024	0.023 (0.021, 0.025)

Data set 1 refers to the data presented in *panel F* of Fig. 1. The column labeled known parameters refers to the starting values of T^m and ED_{50}, respectively. Calculated parameters are the estimates provided by nonlinear regression analysis and are accompanied by the upper and lower 95% confidence intervals in parentheses. *Data sets 2* and *3* represent two new data sets depicted in Fig. 2. Significance of the data is discussed in the text.

equalling 366.07. The *F* statistic for this comparison was 0.22, providing no support for a second T^m ($P = 0.64$). To test for the presence of a second ED_{50} the data was finally fit to *Eq. 1* with a resultant SSR of 383.56 and the corresponding *F* value of 17. Therefore, the conclusion would be drawn that this data set was obtained under conditions different from those in effect when *data set 1* was measured; the entire difference residing in the ED_{50} ($P < 0.00001$).

DISCUSSION

In physiological studies numerous techniques yield data in which one variable varies with another in a sigmoidal fashion when plotted in a log-linear manner or as a hyperbolic function when plotted linearly. These relationships used to be particularly difficult to analyze efficiently; therefore, numerous methods evolved attempting to deal with the problem. Many of these methods were described before microcomputers and statistical software became widely available; thus, emphasis was placed upon ease of use, often with a hand calculator, and ready availability. Because virtually all investigators now have access to microcomputers and powerful statistical software, this emphasis can be changed to include methods that are accurate, precise, and relatively insensitive to error structures commonly encountered in experimental procedures. In practice it is often important for these techniques to establish both a maximal response (a T^m for the dose-response curve) and the concentration that produces a half-maximal response (the ED_{50}). Furthermore, because most experiments compare two or more dose-response relationships, a method should be readily available to statistically evaluate the curves and come to a reasonable conclusion regarding differences. Ideally, this method should distinguish between differences in either the T^m or the ED_{50} because this can have important physiological ramifications.

Within the last 10 years several studies have documented the shortcomings of traditional linearizing transformations in the field of membrane transport processes

TABLE 3. *Statistical comparison of data*

Data Sets Compared	Fitted to Equation	Parameter Estimate				SSR	*F*
		T^m	α	ED_{50}	β		
	4	138.7 (133.3, 144.1)	33.5 (27.1, 39.9)	0.017 (0.015, 0.019)	−0.0004 (−0.002, 0.002)	375.35	
2 vs. *1*	*7*	138.3 (133.0, 143.6)	34.4 (27.7, 41.1)	0.017 (0.015, 0.019)		375.44	0.085
	1	152.3 (145.3, 159.3)		0.017 (0.013, 0.021)		478.17	97.4
	4	140.4 (134.6, 146.3)	−2.0 (−14.9, 10.9)	0.023 (0.021, 0.025)	−0.005 (−0.009, −0.001)	365.84	
3 vs. *1*	*6*	139.4 (135.2, 143.6)		0.022 (0.020, 0.024)	−0.005 (−0.007, −0.003)	366.07	0.22
	1	139.5 (135.4, 143.6)		0.021 (0.019, 0.023)		383.56	17.0

Data sets compared refers to the data set number presented in Table 2. Equation numbers refer to those in the text, and the parameter estimates are those obtained directly from the nonlinear regression routine. Confidence intervals (95%) are in parentheses for each estimate. Finally, the minimized weighted sum of squared residuals for each analysis is presented as the SSR. The *F* test was calculated as outlined.

(2–4). In general these studies have emphasized the value of direct nonlinear regression analysis in constructing precise and accurate parameter estimates despite random error in the observations. The primary advantage of nonlinear regression is that it allows accurate weighting of the data points, in a manner inversely proportional to its variance. However, the recommendations of these studies are not widely used in the field of pharmacological dose-response analysis, as emphasized by the literature analysis we present. Over the last 3 years less than one study in five utilizes these techniques and close to three-fourths still rely on either visual analysis or some form of linearizing transformation. Although the reason for this observation is unclear, one major drawback of the nonlinear regression techniques is that, until now, no readily apparent means of analyzing statistical differences between parameter estimates was available. This is a serious limitation from the physiological perspective because most studies present two or more curves and conclusions are often based on knowledge of significant differences between these.

In this communication we have attempted to perform two major tasks. First, we have evaluated the commonly used analysis techniques to obtain parameter estimates from six data sets. These data were selected to mimic a common physiological experiment; observing muscle tension induced by graded doses of an agonist. After generating a "perfect" data set of tension over a concentration range commonly employed, six different types of random error were introduced into the data. Although not applicable to all physiological data these error structures were similar to those observed in many experimental studies. These data were then evaluated by techniques utilized in the literature and by nonlinear regression analysis. The results suggest, that with the exception of nonlinear regression, all techniques perform poorly in the presence of these types of random error. Despite relatively large data ranges and the presence of a significant number of outliers, nonlinear regression analysis consistently gave precise and accurate estimates of both parameters without requiring removal of data points. These conclusions are not radically different from those reached by Atkins and Gardner (2, 3); however, we have used a different type of data. The concentration range simulated in our study was much wider and the selected concentrations irregularly scattered compared with theirs. In many respects it would appear that the concentrations we chose were not optimal; however, this was by design because we expressly wished to simulate pharmacological dose-response studies in the literature. From the data we present it would appear that the majority of recently published studies have utilized data analysis techniques poorly suited for the task they were asked to perform.

Because part of the reticence to switch to nonlinear regression techniques may be the lack of easily available statistical analysis of parameter estimates, the second task we set out to accomplish was the definition of a simple, yet rigorous, statistical evaluation of these estimates. The method we present in the RESULTS section clearly demonstrates statistically significant differences that were known to be present in simulated data. These differences were modest and of an order commonly encountered in physiological experiments. Despite large, random error inherent in the data, nonlinear regression analysis clearly identified each parameter, and furthermore, the statistical method we present was powerful enough to convincingly demonstrate the differences. The technique, as outlined, is simple, straightforward, and represents an extension of the nonlinear regression technique itself. It does not require extensive mathematical knowledge and can be performed by essentially any investigator in this field with equipment that most, if not all, laboratories have routinely available.

It should be pointed out that the method we suggest is of evident utility in the comparison of two curves. However, it is often important to compare more than two groups, and this method has no clear solution for this problem. Standard methods of adjustment (such as the Bonferroni) might be used to handle multiple comparisons, but more appropriate methods would involve the use of additional indicator variables to identify the other factors under consideration. Further work is required to define a simple and direct method to accomplish this.

Finally, although our emphasis has been on the analysis of dose-response data, we hope the techniques outlined will have a much broader appeal. Other investigators have demonstrated that the inadequacies we describe are encountered in data analysis of intestinal transport (2, 3, 4, 10), enzyme kinetics, and radioligand binding (6). Although we have agreed with the recommendations that nonlinear regression analysis is optimal, we are now much stronger in our opinion because a reasonable statistical approach to the evaluation of parameter estimates is available. The techniques outlined here are widely applicable to many types of study that have in common the analysis of data similar to the classic dose-response relationship. We hope that more investigators make use of what appears to us to be an extremely simple, rapid, convenient yet powerful analytical tool for physiological studies.

The authors thank I. McGovern, without whom the preparation of this manuscript could not have been completed.

Funding for these studies came from generous grants donated by the Alberta Heritage Foundation for Medical Research and the Medical Research Council of Canada.

Address for reprint requests: J. Meddings, University of Calgary, Faculty of Medicine, Dept. of Medicine, Health Sciences Center, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Canada.

Received 17 February 1989; accepted in final form 17 July 1989.

REFERENCES

1. ARMITAGE, P., AND G. BERRY. *Statistical Methods in Medical Research* (2nd ed.). London: Blackwell, 1987, p. 365–366, 488.
2. ATKINS, G. L. A comparison of methods for estimating the kinetic parameters of two simple types of transport process. *Biochim. Biophys. Acta* 732: 455–463, 1983.
3. ATKINS, G. L., AND M. L. G. GARDNER.The computation of saturable and linear components of intestinal and other transport kinetics. *Biochim. Biophys. Acta* 465: 127–145, 1977.
4. ATKINS, G. L., AND I. A. NIMMO. Current trends in the estimation of Michaelis-Menten parameters. *Analyt. Biochem.* 104: 1–9, 1980.
5. BATES, D. M., AND D. G. WATTS. *Nonlinear Regression Analysis and Its Applications.* New York: Wiley, 1988.
6. DELEAN, A., P. J. MUNSON, AND D. RODBARD. Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose-response curves. *Am. J. Physiol.* 235 (*Endocrinol. Metab. Gastrointest. Physiol.* 2): E97–E102, 1978.

7. ENDRENGI, L., AND F. H. F. KWONG. *Analysis and Simulation of Biochemical Systems.* Amsterdam: North-Holland, 1972, p. 219–237.
8. FIELLER, E. C. Some problems in biological assay. *J. R. Statis. Soc.* B16: 175–184, 1954.
9. FINNEY, D. J. *Statistical Methods in Biological Assay* (3rd ed.). New York: MacMillan, 1978, section 4.12, p. 80–82.
10. GARDNER, M. L. G., AND G. L. ATKINS. Kinetic analysis of transport processes in the intestine and other tissues. *Clin. Sci.* 63: 405–414, 1982.
11. LITCHFIELD, J. T., JR., AND F. WILCOXON. A simplified method of evaluating dose-effect experiments. *J. Pharmacol. Exp. Ther.* 96: 99–113, 1949.
12. PECK, C. C., AND B. B. BARRETT. Nonlinear least-squares regression programs for microcomputers. *J. Pharmacokinet. Biopharm.* 7: 537–541, 1979.
13. TALLARIDA, R. J., AND L. S. JACOB. *The Dose-Response Relation in Pharmacology.* New York: Springer-Verlag, 1979.
14. TALLARIDA, R. J., AND R. B. MURRAY. *Manual of Pharmacologic Calculations with Computer Programs.* New York: Springer-Verlag, 1981.