Br. J. clin. Pharmac. (1987), **23**, 199–205

The analysis of dose-response curves—a practical approach

S. H. D. JACKSON[1], M. J. JAMIESON[2], A. JOHNSTON[1] & A. M. M. SHEPHERD[3]
[1]Department of Clinical Pharmacology, St Bartholomew's Hospital, London EC1A 7BE, [2]Department of Therapeutics and Clinical Pharmacology, University of Aberdeen, Aberdeen, Scotland and [3]Division of Clinical Pharmacology, University of Texas Health Science Center, San Antonio, Texas, USA

1 The rationale for the objective assessment of dose-response curves (DRCs) is presented.
2 Using data derived from isoprenaline/heart rate responses studies, two new statistical methods of objectively defining the terminal linear segment of an incomplete DRC are presented.
3 Using data derived from phenylephrine/diastolic blood pressure response studies, the parallel shift quadratic model of Sumner *et al.* (1982) has been extended to include a measure of the suitability of the quadratic model for each individual data set using the Akaike information criterion.
4 A parallel shift E_{max} model is proposed for complete DRCs.

Keywords dose-response curves

Introduction

The technique of constructing dose-response curves (DRCs) is not new and its applications are widespread throughout scientific disciplines. Within the basic medical sciences and specifically within pharmacology comparisons of log DRCs—usually sigmoidal in shape—have centred around comparisons of the parameters E_{max}, ED_{50} and the slope of the simple linear regression line relating effect and log dose (or log concentration). The slope is calculated using only the approximately linear central portion of the sigmoid curve. These comparisons have usually been made using graphical methods. In the presence of a competitive antagonist, parallel shift of the DRC is expected. Mathematically the relationship can be expressed as:

$$E = \frac{E_{max} \cdot D}{ED_{50} + D} \quad (1)$$

where E = effect of any drug dose D
E_{max} = maximum effect
ED_{50} = dose producing half maximum effect
D = drug dose

This formula was originally described by Hill (1910) and has been used in many different situations including relating both drug concentration in a body compartment to effect and also the magnitude of an extraneous stimulus to effect. When log dose is plotted against response the relationship takes on a sigmoidal form. This mathematical relationship does not include terms to describe compensatory homeostatic effects induced by the administered drug.

These commonly used graphical methods of comparison are associated with a number of problems:

1. In human studies the top of the dose-response curve cannot usually be constructed hence E_{max} and ED_{50} cannot be calculated directly from the observed data. Such curves have been described as incomplete (Figure 1).
2. Unless modelling techniques are employed E_{max} and ED_{50} derived from graphical methods do not utilise all the data points.
3. When slopes of the 'linear' segments are compared all data points are rarely utilised. The initial points are disregarded usually on the basis

Figure 1 'Incomplete $\log_{10}$ dose-response curve. Theoretical plot of the function:

$$E = \frac{E_{max} \cdot dose}{ED_{50} + dose} \text{ where } E_{max} = 50, ED_{50} = 25$$

of their appearance, when effect is plotted against log dose or log concentration. When points above this 'linear' segment exist, they too are disregarded for this purpose.

Methods have been employed to 'straighten' the graphical representation of the relationship between dose and response. The most frequently used is the double reciprocal plot which plots 1/dose against 1/response (Figure 2). Such methods

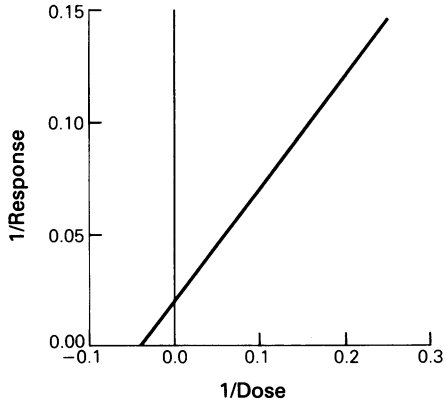

Figure 2 Double reciprocal plot (1/dose *vs* 1/response) of the function:

$$E = \frac{E_{max} \cdot dose}{ED_{50} + dose} \text{ where } E_{max} = 50, ED_{50} = 25$$

y axis intercept = $1/E_{max}$; x axis intercept = $-1/ED_{50}$

are of greatest potential use when only the lower part of the dose-response curve can be constructed and are of no additional benefit when the data for the complete curve are available. Unfortunately because of the distribution of the points on a double reciprocal plot, the precision of the line in the area of interest close to the axes is poor. In addition, the reciprocal transformations produce heteroscedastic (unequal variance) distributions along the line requiring weighted regression procedures. Other more complicated methods have also been described (Paton, 1961).

Despite these problems graphical methods continue to be employed to compare DRCs. We have developed a rational approach to the analysis of such data which attempts (usually successfully) to maximise the power of a comparison and provides a statistical approach to the exclusion of data points when necessary. We also describe methods which permit, whenever possible, utilisation of all data points.

Methods

1 Objective methods to eliminate early data points and define the 'linear' segment of the curve

Traditionally plots have been examined visually and early data points excluded on that basis. This method can produce, with some data sets, large intra- and inter-observer variation. We propose two methods which, between them, will objectively choose the points that most reasonably describe the 'linear' segment. They will not, however, accept or reject the linear model.

Once the points making up the linear segments of a pair of DRCs have been defined the dose ratio (DR) can be calculated from the difference in the x axis intercepts of the two lines if they are parallel. If not, the difference in the x coordinates of an arbitrary y value can be used. Using such plots, $DR = 10^{(x_1 - x_2)}$ where x_1 and x_2 are the x values of a given y value. For example, in the case of isoprenaline dose/heart rate response curves an ID_{20}, the isoprenaline dose which increases heart rate by an arbitrary 20 beats min^{-1}, is derived from placebo and β-adrenoceptor antagonist curves and the dose ratio calculated as ID_{20} (antagonist)/ID_{20} (placebo).

(a) *Quadratic check (QC) method* Empirically, incomplete DRCs are often described by a quadratic equation. This observation has already been described by Sumner *et al.* (1982) who fitted quadratic functions to contrived data derived from the Hill equation (equation 1) and to experimental data. Using experimental data we

have adopted a more rigorous approach applying the Akaike information criterion (AIC) (Akaike, 1976) to accept or reject the quadratic term (Table 1). As the early points are disregarded, the remaining points are less and less well fitted by a quadratic model and correspondingly better fitted by a straight line (Figure 3).

Initially, all the data is fitted to a simple linear function:

$$y = b_0 + b_1x$$

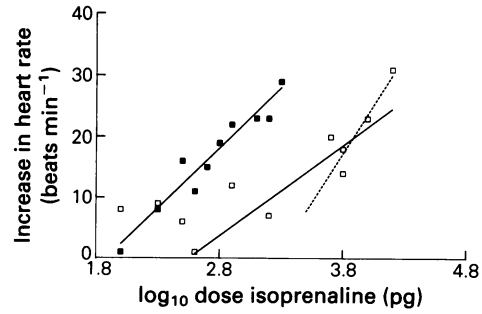

Figure 3 Isoprenaline dose-response curves obtained after i.v. pretreatment with saline placebo (■) and propranolol 5 mg (□) in a healthy volunteer. When applied to the placebo data the simple linear regression (SLR) and quadratic check (QC) methods choose all points; when applied to the propranolol data the SLR method chooses the last nine points (----) whilst QC method chooses the last six points (——).

Table 1(a) Residual sums of squares (RSS) and Akaike information criteria (AIC) calculated before (n = 10) and after serial data point deletion for both linear and quadratic models used to fit log isoprenaline dose/heart rate response data after pretreatment with placebo in a healthy volunteer. The quadratic check (QC) method chooses the maximum number of points (marked †) at which the quadratic model first fails to produce a lower AIC than the linear model. The raw data are plotted in Figure 3 with the placebo points marked (■).

	Model			
	Linear		*Quadratic*	
Points	*RSS*	*AIC*	*RSS*	*AIC*
10	47.3	42.6	45.0	44.1†
9	44.4	38.1	44.3	40.1
8	42.8	34.1	41.1	35.7
7	25.2	26.6	17.8	26.2
6	16.4	20.8	15.7	22.5
5	11.5	16.2	10.7	17.9
4	11.3	13.7	6.8	13.7
3	97	10.8	9.7	12.8

(b) Residual sums of squares (RSS) and Akaike information criteria (AIC) calculated before (n = 11) and after serial data point deletion for both linear and quadratic models used to fit log isoprenaline dose/heart rate response data after pretreatment with propranolol 5 mg iv in the same subjects as in Table 1a. The quadratic check (QC) method chooses the maximum number of points (marked †) at which the quadratic model first fails to produce a lower AIC than the linear model.

	Model			
	Linear		*Quadratic*	
Points	*RSS*	*AIC*	*RSS*	*AIC*
11	347.3	68.4	147.9	61.0
10	292.1	60.8	137.3	55.2
9	199.1	51.6	136.3	50.2
8	198.9	46.3	66.6	39.6
7	86.0	35.2	61.1	34.8
6	53.3	27.9	53.2	29.8†
5	46.1	23.2	17.1	20.2

and to a quadratic function:

$$y = b_0 + b_1x + b_2x^2 \quad (2)$$

using least squares simple linear and multiple linear regression analysis. The AIC for each function is defined as:

$$AIC = n.\log_e(RSS) + 2p \quad (3)$$

where n = number of points
RSS = residual sum of squares
p = number of parameters (linear function = 2, quadratic = 3)

If the quadratic AIC is lower than the linear AIC the quadratic term is regarded as contributing significantly to the reduction in residual sum of squares. The lowest point of the DRC is therefore eliminated and the procedure repeated until the quadratic AIC exceeds the linear AIC. At this point the quadratic term no longer contributes significantly. The remaining points are chosen as the best linear segment.

Previously, we have employed analysis of variance to estimate the significance of the quadratic term (Jamieson *et al.*, 1985). In doing so we used $P > 0.2$ as the rejection criterion for the quadratic term. We have found this method to give slopes for the linear segment and dose ratios which are not significantly different from those defined using the AIC. Although satisfactory in our experience for DRCs having a range of 6–18 points, the choice of $P > 0.2$ is arbitrary.

(b) *Simple linear regression (SLR) method*
Whereas the QC method identifies the number of points at which the quadratic term in the multiple regression equation no longer contri-

butes to the equation, the SLR method identifies the number of points which produce the variance ratio with the most significant P value derived from ANOVA.

SLR analysis is first performed on all points, then all but the lowest dose point and so on until only three points remain. The series of points yielding the lowest P value is taken as the 'linear' segment of the DRC. (Table 2, Figure 3).

2 *Nonlinear parallel shift models*

In concept these methods are identical to the methods of calculating DR once the 'linear' segment has been defined. In practice they are combined with an additional refinement that not only applies the nonlinear model to each data set (full model) but also applies a single best fit model to the combined data set solving for a parallel shift term (reduced model).

(a) Parallel quadratic functions Although not a nonlinear model the solution is most easily achieved using iterative nonlinear regression analysis software. This method was described by Sumner *et al.* (1982) who applied to each of two data sets a quadratic model of the form:

$$y = b_0 + b_1x + b_2x^2.$$

Making the assumption that parallel shift will occur in the presence of a competitive antagonist, a reduced model can then be constructed of the form:

$$y = b_0 + b_1 (x+\Delta) + b_2(x+\Delta)^2$$

where Δ is the parallel shift term.

DR can be calculated from:

$$DR = 10^{\Delta} \quad (4)$$

(if $\log_{10}$ dose transformations are used)

or $$DR = e^{\Delta} \quad (5)$$

(if $\log_e$ dose transformations are used)

Using the same test statistic as Sumner *et al.* (1982) the increase in the residual sum of squares associated with the reduced model can be used to obtain a P value:

$$F = \frac{(RSS\{full\} - RSS\{reduced\})/(P_1 - P_0)}{RSS\{reduced\}/(n-P_1)} \quad (df = 2, n-6)$$

where RSS {full} = residual sum of squares of full model (both curves)

RSS {reduced} = residual sum of squares of reduced model

P_1 = number of parameters in full model

P_0 = number of parameters in reduced model

n = number of data points

If the P value is below an arbitrary value of 0.05 the reduced model may be rejected as unsatisfactory compared to the full model. This statistic does not accept or reject the quadratic model, however.

We have tested the appropriateness of the quadratic model compared with the linear model using the AIC (Akaike, 1976). If the AIC for the quadratic model is lower than that for the linear model in the majority of subjects a quadratic model is acceptable. If not, a parallel quadratic function model is no more appropriate than a simple linear model.

(b) *Parallel E_{max} model* Using an E_{max} model of the form of equation 1 the full sigmoid function (log dose *vs* response) can be modelled rather than the lower part of the curve. Parallel shift can then be described using a reduced model of the form:

$$y = \frac{E_{max} \cdot dose}{(ED_{50} \cdot \Delta) + dose}$$

where Δ is the parallel shift term (in this case equal to the dose ratio).

Depending on the shape of the observed sigmoid DRC, a better fit may be obtained using an additional parameter, n, as originally described by Hill (1910) which can be applied to a reduced model of the form:

$$y = \frac{E_{max} \cdot dose^n}{(ED_{50} \cdot \Delta)^n + dose^n}$$

The E_{max} model requires data points close to if not at E_{max} in order to provide reasonably precise parameter estimates.

In addition to the commercially available statistical package SAS we have used a modification of our BASIC non linear regression analysis program SIMP (Johnston, 1985). A copy of the code can be obtained from the authors.

Results

1 SLR method

Using conventional techniques, i.v. isoprenaline bolus/HR DRCs were obtained in a previous study after pretreatment with both placebo and 5 mg of i.v. propranolol (Jamieson *et al.*, 1985). When all the placebo log dose-response data points were subjected to SLR a highly significant linear relationship was found (Table 2a). When serial deletion of data points starting with the lowest dose was performed, the significance of the relationship was reduced. Thus this method

Table 2(a) Analysis of variance before (n = 10) and after each successive deletion of log isoprenaline dose/heart rate response data points after pretreatment with placebo (same data as in Table 1a). The simple linear regression (SLR) method chooses the number of points with the lowest P value (marked †) to reflect the 'linear' segment.

Points	F	*d.f.*	P
10	99.4	1,8	0.000009†
9	49.6	1,7	0.0002
8	26.6	1,6	0.0002
7	37.8	1,5	0.002
6	23.2	1,4	0.009
5	10.9	1,3	0.05
4	3.1	1,2	0.2
3	1.1	1,1	0.5

(b) Analysis of variance before (n = 11) and after each successive deletion of log isoprenaline dose/heart rate response data after pretreatment with 5 mg iv propranolol (same data as in Table 1b). The simple linear regression (SLR) method chooses the number of points with the lowest P value (marked †) to reflect the 'linear' segment.

Points	F	*d.f.*	P
11	10.9	1,9	0.009
10	12.0	1,8	0.008
9	17.7	1,7	0.004†
8	8.9	1,6	0.02
7	22.5	1,5	0.005
6	22.7	1,4	0.009
5	7.6	1,3	0.07

chose all the data to represent the 'linear' segment (Figure 3). Using the propranolol data this approach deletes the first two data points (Table 2b, Figure 3).

2 QC method

Using the same data as in the SLR analysis all 10 points are chosen when applying the QC method to the placebo data (Table 1a, Figure 3). When applied to the propranolol data, however, this method chooses only the last six data points compared with the last nine points for the SLR method (Table 2b, Figure 3).

Thus, although the two methods produce identically sloped regression lines for the placebo data (regression coefficient = 19.8), the QC method produces a slope approximately twice that of the SLR method when the propranolol data are considered (regression coefficients of 33.3 and 15.2 respectively). Nevertheless, the difference in intercepts only translates to a DR of 9.57 for the QC method *vs* 9.59 for the SLR method, comparing doses needed to produce an increase in HR of 20 beats min^{-1}. When all 20 isoprenaline dose-response studies were considered, no consistent difference was seen between the two methods. Had the arbitrary choice of 20 beats min^{-1} not been used a different DR would have been obtained in this particular data set. It is important to use the same value for all data sets and to define the value before performing the data analysis as we did.

3 Parallel quadratic function

Using conventional techniques, i.v. infusions of phenylephrine/diastolic blood pressure DRCs were obtained in a previous study after pretreatment with both placebo and 30 mg of the α_1-adrenoceptor antagonist urapidil given i.v. (Jamieson *et al.*, 1986), (Figure 4).

A comparison of AIC values for linear and quadratic models showed only two out of the 16 data sets were not better modelled by quadratic functions (Table 3). These two were the largest data sets. When the reduced model was employed the increase in the RSS only reached statistical significance for one data set (Table 3). The mean dose ratio was 5.48 with a standard error of 0.66. The raw data used to compile Tables 1 to 3 can be obtained from the authors.

Discussion

The choice of methods to evaluate DRC data must be made on the basis of a particular set of data rather than a fixed approach. A number of factors are important in the decision making process:

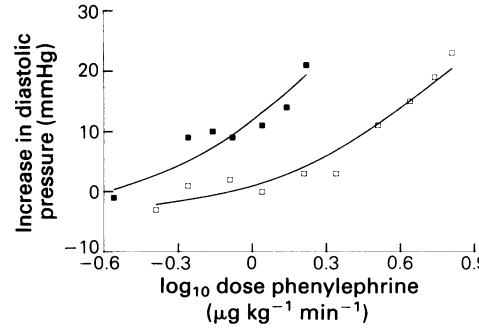

Figure 4 Phenylephrine dose-response curves obtained after i.v. pretreatment with saline placebo (■) and urapidil 15 mg (□) in a healthy volunteer. The best fit single quadratic function with parallel shift (reduced model) is shown.

Table 3 Difference in Akaike information criteria (AIC) calculated for both linear (AIC_l) and quadratic (AICq) models together with total number of data points (n), F statistic (see text) and the significance of this statistic (*P*) after both placebo and active drug (urapidil 30 mg i.v.) pretreatment in eight healthy volunteers. Each of the 16 dose response curves was constructed using log phenylephrine dose/diastolic blood pressure response data. Where the AIC difference is negative (marked *), the quadratic model is not superior to the linear model. A single dose-response curve (marked —) was a perfect quadratic curve and hence the error sum of squares was zero making an AIC difference impossible to calculate. The F statistic just reaches statistical significance for only one subject (marked †) and hence the reduced model is used to calculate dose ratios for all subjects.

Subject	*Study*	AIC_l-*AICq*	n	F	P	*DR*
1	Placebo	2.07				
	Active	1.61	14	4.52	0.049†	4.06
2	Placebo	−1.84*				
	Active	2.20	15	0.32	0.73	3.76
3	Placebo	—				
	Active	8.87	10	4.13	0.11	7.71
4	Placebo	−1.71*				
	Active	3.04	17	0.54	0.59	6.85
5	Placebo	0.89				
	Active	3.25	13	0.11	0.9	3.48
6	Placebo	2.77				
	Active	10.96	11	1.80	0.26	8.07
7	Placebo	10.06				
	Active	1.92	13	0.64	0.56	5.81
8	Placebo	2.49				
	Active	11.56	13	2.60	0.14	4.06

1 Number of data points

If, for example, an isoprenaline/rise in heart rate DRC is constructed with 15–20 points between 0 and 30 beats min^{-1} the loss of data points by using a linear model would not have much impact on the power of a comparison with a second similar data set. On the other hand, a tyramine/rise in systolic blood pressure DRC constructed with four or at most five data points could be reduced to an unacceptable two points if such a method was used.

2 Shape of the observed data points

A complete sigmoid curve obtained from *in vitro* work will best be described using the E_{max} model, regardless of the number of data points. A phenylephrine/rise in BP DRC reaching a maximum change in BP of 20 mmHg might be almost perfectly described by a quadratic function or, particularly if a higher maximum change in BP was achieved, might produce four or five points almost on a straight line. In this latter case, a linear model would be most appropriate, the more so if few points were obtained on the initial part of the curve.

3 Availability of nonlinear regression analysis (NRA) software

Simple and multiple linear regression analysis is now widely available in user friendly packages on most desk top microprocessors, but non linear regression analysis requires the use of more complex programs which, although commercially available in many larger statistical packages, are not widely used for this purpose.

Taking these factors, into consideration, a decision to use a linear or non linear method must be made after assessing the data with simple graphical and data fitting methods. First the data should be graphed using the software that will be used for the next step. The question normally being asked of the data in the setting of clinical pharmacology is does a drug shift the DRC? An impression can be obtained at this stage. If the two DRCs for most subjects are superimposable the choice of method is not going to influence the conclusions.

If NRA is not available, a linear model must be used. The data should then be fitted to a quadratic model (equation 2) and the resulting curve compared with the observed data. If the

two are very close, particularly in the upper part of the curve, the quadratic check method will not be useful and the SLR method must be used to exclude the early points. If not, either method can be used.

Both methods of defining the linear segment require a cut off to be made at an arbitrary level either using AIC values or *P* values depending on the method. Random error in the data in the area of the cut off may give undue weight to a given data point. However, this will not lead to bias in the dose ratios, merely to larger confidence intervals around the DR.

Once the linear segment has been defined, the least squares regression line for each data set should be plotted for each subject. Assuming competitive antagonism a parallel shift would be expected. Significant deviation from parallelism in a data set derived from a single individual can best be tested for using analysis of covariance which uses all the data points making up the two lines. Alternatively, the slope (regression coefficient) of each line for each subject can be tested using a paired *t*-test on the group data. Depending on whether the lines are parallel, a DR can be calculated at any y value (parallel) or a specific y value (non parallel). If statistically significant deviation from parallelism occurs in more than one subject it would be wise to reassess the laboratory and animal data on the test drug (looking for evidence of partial agonism of non-competitive receptor blockade) and the clinical experimental methodology.

Conceptually it would not be unreasonable to propose a parallel shift linear model after excluding initial points but this would again require more complex software which, if available, would be better used in applying one of the nonlinear models. When the number of data points is sufficiently small that initial data point exclusion cannot be afforded or the initial graphing does not suggest a 'linear' portion and NRA software is available, a parallel shift quadratic model should be considered.

If AIC values in the majority of DRCs in a particular study suggest the quadratic term should be used, the next step would be to test the reduced model against the full model as described by Sumner *et al.* (1982). We prefer the AIC to Boxenbaum's F statistic (Boxenbaum *et al.*, 1976) to test the value of the quadratic term as the latter provides too conservative an approach in this situation. If the full model is significantly better than the reduced model in the majority of comparisons this suggests nonparallel shift and reevaluation of existing data on the drug and the application of an alternative technique.

In some clinical situations (e.g. atropine dose *vs* rise in heart rate, β-adrenoceptor blocker dose *vs* suppression of exercise induced tachycardia) a full DRC can be constructed. Under these circumstances a parallel quadratic shift model would not be appropriate. A parallel shift E_{max} model would be the obvious choice if NRA software was available. Failing that, a linear model could be used after deleting both initial and terminal points. This could be done by adapting the SLR method or merely by choosing points from 20% to 80% of E_{max}. If sufficient points are present, a more conservative range of 25% to 75% might be used.

We are grateful for the statistical and computing advice of Dr Tom J. Prihoda, Robert C. Wood and Dr William Stewart.

References

Akaike, H. (1976). An information criterion (AIC). *Math Sci.*, **14**, 5–7.

Boxenbaum, H. G., Riegelman, S. & Elashoff, R. M. (1976). Statistical estimations in pharmacokinetics. *J. Pharmacokin. Biopharm.*, **2**, 123–148.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J. Physiol.*, **40**, iv–vii.

Jamieson, M. J., Patel, S. S., Shepherd, A. M. M., Jackson, S. H. D., Flanagan, P. H. & Galbraith, H. (1985). Is urapidil also a beta blocker? *Pharmacologist*, **27**, 259.

Jamieson, M. J., Patel, S. S., Shepherd, A. M. M., Jackson, S. H. D., Flanagan, P. H. & Galbraith, H. (1986). Urapidil's alpha blocking activity in humans. *Clin. Pharmac. Ther.*, **39**, 200.

Johnston, A. (1985). SIMP: A computer program in BASIC for non-linear curve fitting. *J. pharmac. Methods*, **14**, 323–329.

Paton, W. D. M. (1961). Rate theory of drug action. *Proc. Roy. Soc.* Series B, **154**, 21–69.

Sumner, D. J., Elliott, H. L. & Reid, J. L. (1982). Analysis of the pressor dose response. *Clin. Pharmac. Ther.*, **52**, 450–458.

(Received 7 August 1986, accepted 15 August 1986)