Defining the linear portion of a sigmoid-shaped curve: Bend points

Article *in* Pharmaceutical Statistics · July 2003

DOI: 10.1002/pst.62

PHARMACEUTICAL STATISTICS
Pharmaceut. Statist. 2003; **2**: 167–174 (DOI:10.1002/pst.062)

Defining the linear portion of a sigmoid-shaped curve: bend points

MAIN PAPER

J. L. Sebaugh[1,*,†] and P. D. McCray[2]
[1] *Sebaugh's Information Services, Columbia, MO, USA*
[2] *Pharmacia Corporation, Skokie, IL, USA*

Formulae are provided that define the 'bend points', the beginning and end of the essentially linear dose–response region, for the four-parameter logistic model. The formulae are expressed in both response and dose units. The derivation of the formulae is shown in order to illustrate the general nature of the methodology. Examples are given that describe how the formulae may be used while planning and conducting bioassays. Copyright © 2003 John Wiley & Sons Ltd.

Keywords: *bend points; bioassays; four-parameter logistic model; linear region*

1. INTRODUCTION

The response pattern for most bioassays is typically described by an S- or sigmoid-shaped curve (when the concentration is portrayed on a logarithmic scale). If the response is low at low concentrations or if the response is expressed as percentage inhibition, the slope will be positive. There will be a lower plateau of minimal response at low concentrations that eventually bends upward into a log-linear response region. As the response nears its maximal level, the line bends into an upper plateau of maximal response. If the response is high at low concentrations or if the response is expressed as percentage control and the no-compound control has the maximal response, there will be a similar pattern with a negative slope.

The purpose of these bioassays often involves the estimation of the concentration that corresponds to the response level of a test sample or a specific level of response, such as 50%. Most commonly, interest is in estimating an EC50, IC50, or LC50 (effective, inhibitory, or lethal concentration corresponding to a 50% response). Without loss of generality, the term EC50 will be used in this paper.

The purpose of this paper is to provide specific formulae to express the beginning and end of the essentially linear concentration–response region, the 'bend points'. Formulae are derived and expressed in both response and concentration units. Examples are given that describe how these bend-point values may be used.

2. FOUR-PARAMETER LOGISTIC MODEL

The four-parameter logistic model has a long history of use in describing the sigmoid-shaped response pattern [1,2]. Whereas the probit and

*Correspondence to: J. L. Sebaugh, Sebaugh's Information Services, 3609 Holly Hills Ct., Columbia, MO 65203, USA.
†E-mail: sis.stat@centurytel.net

logit models are used to describe quantal responses (such as percentage survival based on reports of number dead and alive), the logistic model is used to describe quantitative measurements from bioassays, such as counts per minute (CPMs) or optical densities (ODs). The model assumes symmetry of response, that is, the rate of increase in the lower bend is the same as the rate of decrease in the upper bend.

The formula for the four-parameter logistic model may be expressed as

$$Y = \frac{a-d}{1+(X/c)^b} + d \tag{1}$$

where Y is the response, X is the concentration, a is the lower asymptote, d is the upper asymptote, c is the EC50 (in the same units as X), and b is a slope factor. The lower asymptote, a, is referred to as the minimum; the upper asymptote, d, is referred to as the maximum; b is referred to as the slope; and c is referred to as the 'model EC50' since it is the concentration corresponding to a response midway between the estimates of two of the model parameters, the minimum and the maximum.

It must be emphasized that the model EC50 is defined as the concentration corresponding to the response midway between the estimates for the minimum and the maximum (the lower and upper plateau). If the lower and upper plateau do not correspond to 0 and 100% effect, the model EC50 does not correspond to 50% effect. Unless the full concentration response is modelled using a wide range of concentrations, the estimated values of the minimum, maximum, and model EC50 have no intrinsic biological meaning. The estimate of the minimum will not necessarily be at the level of the minimum response possible, and the estimate of the maximum will not necessarily be at the level of the maximum response possible.

3. DERIVATION OF THE BEND POINTS

The bend points occur at the points on the curve at which the slope of the essentially linear concentration response changes upon approaching the lower and upper plateau. The slope factor b 'corresponds to the slope of the logit–log plot, when X is portrayed in terms of natural logarithms' [1, footnote 2]. Taking the first derivative of the response variable, Y, with respect to the slope, b, provides a formula describing the rate of change of the response, Y, with respect to the slope, b:

$$\frac{\mathrm{d}Y}{\mathrm{d}b} = \frac{(d-a)(X/c)^b \log(X/c)}{(1+(X/c)^b)^2} \tag{2}$$

It is a function of the concentration, X. In Figure 1, the solid line is an example of a four-parameter logistic curve fit ($a=0.212$, $d=1.38$, $b=1.359$, $c=0.517$). The dashed line is the value of its first derivative with respect to b. This is zero on both the lower and upper plateaus. It has a region that appears to be essentially linear over the same concentrations as the full logistic model. This region starts at the bottom of a concave curve and ends at the top of a convex curve. The concentrations corresponding to the bottom and top are at the bend points.

Note that the rate of change at each of these points is zero (being negative on one side and positive on the other side). Differentiating $\mathrm{d}Y/\mathrm{d}b$ in (2) with respect to X, we obtain the mixed partial derivative of Y with respect to b and X:

$$\frac{\mathrm{d}}{\mathrm{d}X}\left(\frac{\mathrm{d}Y}{\mathrm{d}b}\right) = \frac{(d-a)(X/c)^b\left(1+(X/c)^b + b\log(X/c) - b(X/c)^b\log(X/c)\right)}{X\left(1+(X/c)^b\right)^3} \tag{3}$$

The dotted line in Figure 1 shows the values of this mixed partial derivative over concentrations. Note that this curve crosses the X axis where the rate of change in the slope is zero. To derive the values of X at these points, this equation can be set to zero and solved for X.

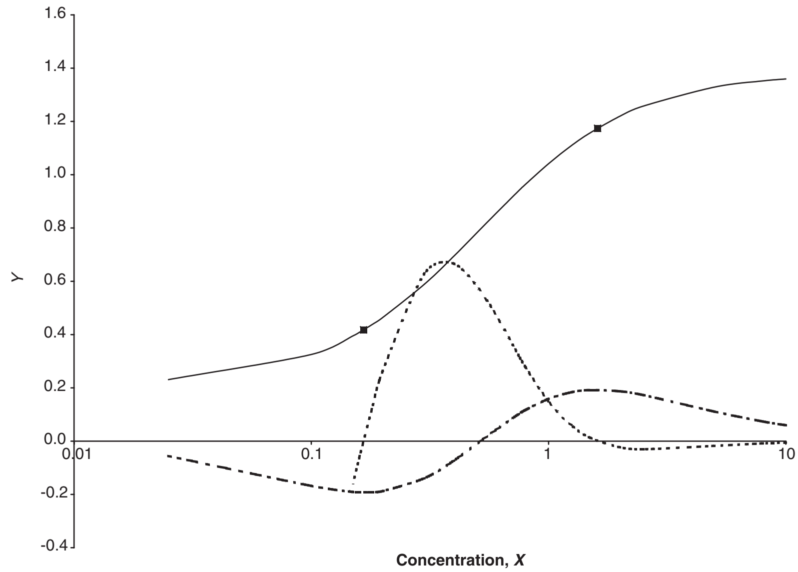


Figure 1. $Y = f(X)$ (solid curve), first derivative of Y with respect to b (dashed curve) and mixed partial derivative with respect to b and X (dotted line). Bend points are shown as filled squares.

If we let $z = (X/c)^b$, then $b\log(X/c) = \ln z$. We may rewrite the numerator of the expression in (3) as

$$\frac{\mathrm{d}}{\mathrm{d}X}\left(\frac{\mathrm{d}Y}{\mathrm{d}b}\right) = \frac{(d-a)z(1+z+\log z - z\log z)}{X \cdot z^3} \tag{4}$$

Since X and c are both concentrations that are constrained to be greater than zero, the second derivative is zero when the numerator is zero. Actually, the numerator is zero when the last term in the numerator, $1 + z + \log z - z\log z$, is zero. Therefore, to find the values of z for which the second derivative is zero, we may set this term to zero and solve for z. If $z > 0$ is a solution, it can be easily shown that $1/z$ is also a solution.

If $1 + z + \log z - z\log z = 0$, we may solve for log z to obtain $\log z = (z+1)/(z-1)$. Taking the exponential of both sides of this equation results in $z = \exp((z+1)/(z-1))$. The exponential function is not defined for $z = 1$. However, since $z = 1$ is not a zero solution to the expression in (4), we have not modified the solution set by considering the transformed equation.

Figure 2 is a graph of the function $f(z) = \exp((z+1)/(z-1)) - z$ for values of z greater than 1. It suggests that $f(z)$ is a monotonically decreasing function that crosses the z axis only once. One can confirm this by noting that over the region of interest the derivative $f'(z)$ is negative and that the function $f(z)$ takes on both positive and negative values. Figure 2 also shows that the value of z satisfying $f(z) = 0$ is between 4 and 5. Therefore, by the intermediate value theorem, the only solution greater than 1 must be between 4 and 5.

The Mathematica® [3] computer algebra system can be used to solve for the value of z between 4 and 5 that satisfies $f(z) = 0$. There is not an exact formula for z, but Mathematica® derives the value iteratively (to 16 decimal places) as $z = 4.680\,498\,579\,882\,905\,6$. It can be rounded to 4.6805. We will denote this number by k.

Substituting the constant k for $z = (X/c)^b$ in the equation for the logistic model (1) results in one of the bend points being defined by

$$Y = \frac{a-d}{1+k} + d \tag{5}$$

Pharmaceutical
STATISTICS

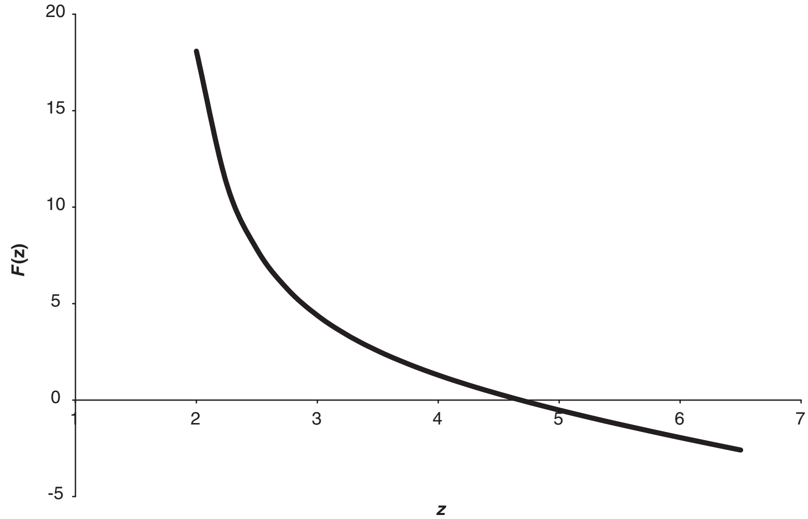


Figure 2. Graph of $\exp((z+1)/(z-1))-z$ for $z>1$.

Thus, the bend point formulae use a constant (k) derived from stationary point determination on a formula derived using differential calculus, namely $k=4.6805$. The bend points in (5) can be stated both in terms of the response units (%s, ODs, CPMs) or the concentrations (nM, ppm). The bend points in terms of the response units will be called '*Y*bend lower' and '*Y*bend higher'. The bend points in terms of the concentrations will be called '*X*bend lower' and '*X*bend higher'. The *Y*bend values depend only on the estimates of the lower and upper plateau, a and d:

$$\begin{aligned} Y\text{bend lower} &= \frac{a-d}{1+1/k}+d, \\ Y\text{bend higher} &= \frac{a-d}{1+k}+d \end{aligned} \tag{6}$$

Values in (6) can then be substituted in the formula for the four-parameter logistic model in order to calculate the corresponding concentrations as follows:

$$X\text{ bend} = c\left(\frac{a-Y\text{bend}}{Y\text{bend}-d}\right)^{1/b}, \tag{7}$$

where *Y*bend is either *Y*bend lower or *Y*bend higher. *X*bend also depends on the value of the EC50 (c) and slope (b).

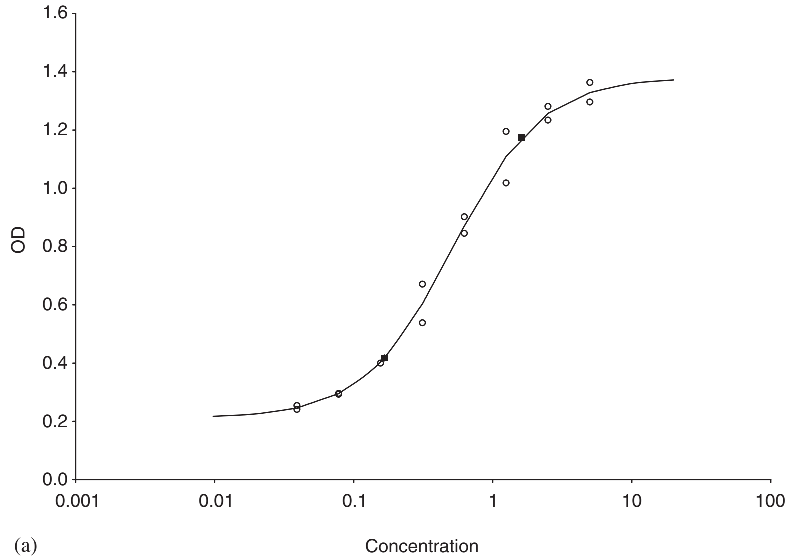
If the slope of the curve is negative, then the concentration corresponding to *Y*bend higher will be the lower concentration, and the concentration corresponding to *Y*bend lower will be the higher concentration. If the slope of the curve is positive, the reverse is true.

If the estimate of the lower plateau (a) were 0% and the estimate of the upper plateau (d) were 100%, solving for *Y*bend lower and *Y*bend higher would yield values of 17.6% and 82.4%. Thus, the portion of the sigmoid-shaped curve between the bend points consists of about the middle 64% of the response range (between the estimates of the lower and upper asymptotes).

Figure 3 shows the source data for Figure 1, Figure 3(a) having a positive slope and Figure 3(b) having a negative slope. The essentially linear portion of the curve is between the response, OD units of 0.42 and 1.17, or the concentrations of 0.166 and 1.609.

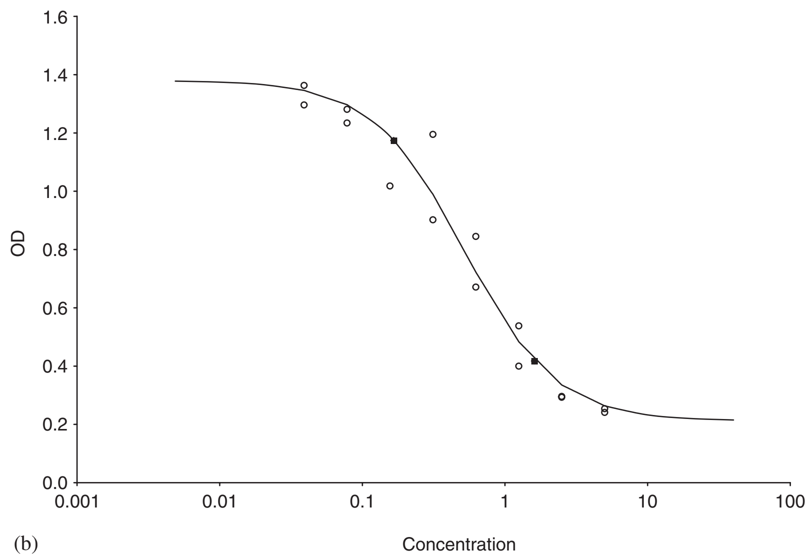
4. EXAMPLES OF USE

During assay development, it is useful to study the concentration–response patterns of the samples



Conc.	OD
5	1.296
5	1.363
2.5	1.234
2.5	1.281
1.25	1.018
1.25	1.195
0.625	0.902
0.625	0.845
0.3125	0.671
0.3125	0.538
0.15625	0.400
0.078125	0.296
0.078125	0.293
0.0390625	0.254
0.0390625	0.241

(a)



Conc.	OD
0.0390625	1.296
0.0390625	1.363
0.078125	1.234
0.078125	1.281
0.15625	1.018
0.3125	1.195
0.3125	0.902
0.625	0.845
0.625	0.671
1.25	0.538
1.25	0.400
2.5	0.296
2.5	0.293
5	0.254
5	0.241

(b)

Figure 3. (a) Example of data fitted to a four-parameter logistic model having a positive slope. (b) Example of data fitted to a four-parameter logistic model having a negative slope. Bend points are shown.

for which the four-parameter logistic model will be used. One of the key decisions made during assay development is the range of concentrations to use. The range is determined by the highest concentration, the number of concentrations, and the dilution factor (ratio of adjacent concentrations). During assay development, assays can be conducted using extra long dilution series for a set of compounds that represent the range of expected potencies in the assay. The four-parameter logistic model can be estimated for each compound, the bend points calculated, and the distribution of concentrations corresponding to the lower and upper bend points examined.

Table I shows the bend points for four samples that were run on each of four plates. The range of concentration bend points is from 0.037 μM to 1.082 μM. The samples to be screened were known to be much less potent than the samples used to generate these bend points. The number of samples to be screened was very large, limiting the number of concentrations per sample to six. Table II shows four scenarios with dilution factors of 4 and 5 and the highest concentrations of 10 μM and 100 μM. A researcher decided to use the higher concentration since the initial samples were not potent. The dilution factor of 4 did not allow the concentration to even reach the lower concentration bend point when the concentration series started at 100 μM. Thus, the researcher selected set 3, starting at 100 μM with a dilution factor of 5, as the initial plate set-up for this screening assay.

To balance the number of concentrations in the linear region and in the plateau, it is helpful to know what factors can affect the number of concentrations in the linear region. Figures 4 and 5 illustrate how the slope and dilution factor have a great impact on the distribution of concentrations in these regions.

Sometimes the slope is a function of the assay protocol developed during the assay development period. For a given protocol, the slope often has an assay-specific value. Using the formulae for the bend points, it is easy to demonstrate the effect of different slopes on the number of concentration in the essentially linear region. For the same dilution factor, a steeper slope results in fewer concentra-

Table II. Four scenarios for assay concentrations.

Set:	1	2	3	4
Dilution factor:	5	4	5	4
Highest conc. (μM):	10	10	100	100
Next 5 conc.:	2	2.5	20	25
	0.4	0.625	4	6.25
	0.08	0.1563	0.8	1.5625
	0.016	0.0391	0.16	0.3906
	0.0032	0.0098	0.032	0.0977

Table I. Using bend points to determine assay concentrations.

		Bend points, response units (% Control)		Bend points, concentration units (μM)	
Plate	Compound	Lower	Upper	Lower	Upper
1	1	13.4	67.5	0.0685	0.7388
	2	15.0	77.9	0.0370	0.7236
	3	16.6	77.5	0.0685	0.6299
	4	16.2	76.3	0.0696	0.6470
2	1	16.0	81.0	0.0489	1.0820
	2	16.8	83.9	0.0408	0.9180
	3	21.4	95.2	0.0405	0.5597
	4	21.0	93.1	0.0399	0.5562
3	1	15.6	79.7	0.0515	1.0193
	2	16.3	82.5	0.0429	0.9002
	3	20.3	94.3	0.0409	0.5593
	4	20.3	93.9	0.0377	0.5500
4	1	13.8	68.6	0.0649	0.7363
	2	15.5	78.2	0.0390	0.6783
	3	16.7	78.0	0.0670	0.6381
	4	16.3	76.5	0.0689	0.6324
	Min.	13.4	67.5	0.0370	0.5500
	Max.	21.4	95.2	0.0696	1.0820
	Median	16.3	79.0	0.0459	0.6627
	Mean	17.0	81.5	0.0517	0.7231

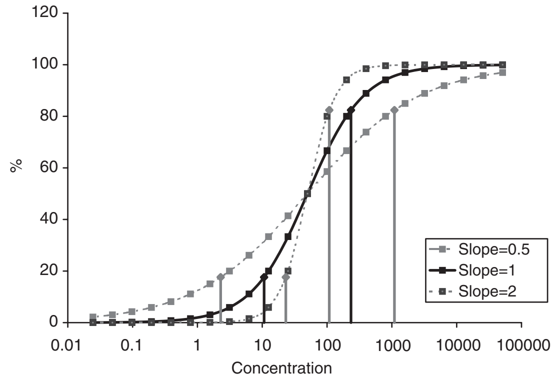


Figure 4. Effect of slope on size of linear region.

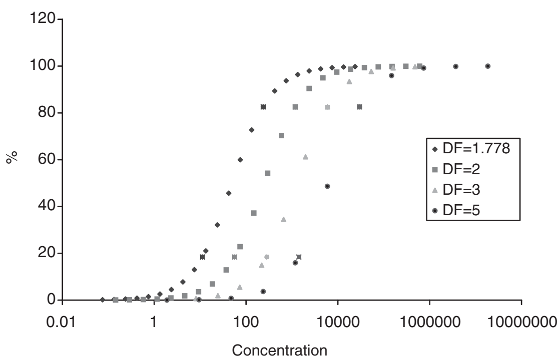


Figure 5. Effect of dilution factor (DF) on number of points in linear region.

tions in the linear region. Figure 4 gives an example for three curves having slopes of 0.5, 1, and 2, respectively. The lower and upper plateau, model EC50, and concentration-dilution factor are all held constant. The number of concentrations between the bend points varies from nine for the lowest slope of 0.5 to three for the highest slope of 2.

The dilution factor used also has an effect on the number of concentrations in the essentially linear region. Figure 5 shows four curves having a slope of 1, lower plateau of 0 and upper plateau of 100%. The model EC50 values were varied to allow simultaneous plotting. The number of concentrations between the bend points varied from six for the smallest dilution factor of 0.5 to only two concentrations for the highest dilution factor of 5.

The bend points may also be used to evaluate the adequacy of the data to make accurate estimates of the lower and upper plateau. It is the responses to the sample concentrations outside the bend points that contain the most information about the plateau. Depending on the slope and dilution factor, the number of concentrations outside this region that are required to obtain a reasonable estimate for the plateau can be determined. Using the bend-point formulae, this number of concentrations can be calculated and the adequacy of the data evaluated.

Relative potency can be estimated by using Finney's formulae [2] or by dividing the EC50 of the standard by the EC50 of the test sample. The bend points may be used to evaluate the acceptability of the results for a reference standard or assay release sample. An EC50 is a concentration. An estimated EC50, particularly for the standard, that does not fall within the range of concentrations corresponding to the essentially linear portion of the four-parameter logistic model would not be considered acceptable. This requirement can be restated in response units by saying that the 50% response value must fall within the lower and upper bend points (stated in response units).

The four-parameter logistic model is also used in standard curve assays. In the past, a great deal of effort has been made to determine the strictly linear range of concentrations, and only those concentrations within the strictly linear range were used. By using the four-parameter logistic model the full concentration–response curve may be used. Nevertheless, one can still restrict the estimation to the most stable region between the bend points. In the latter situation, the lower and upper bend points may be used to estimate the limits of quantitation.

5. SUMMARY

Formulae are developed that explicitly delineate the essentially linear portion of the curve for the four-parameter logistic model. The methodology is described using formulae and graphics to facilitate

Pharmaceutical STATISTICS

an understanding of the conceptual meaning of bend points. Examples were given of how the formulae for the bend points provided in this paper can be used in the development and use of bioassays.

REFERENCES

1. De Lean A, Munson PJ, Rodbard D. Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose–response curves. *American Journal of Physiology* 1978; **235**(2):E97–E102.
2. Finney DJ. *Statistical method in biological assay* 3rd edition. Griffin: London, 1978.
3. Wolfram S. *The Mathematica*® Book 4th edition. Cambridge University Press: Cambridge, 1999.