LETTER

Analysis of biochemical data by nonlinear regression: is it a waste of time?

Leatherbarrow's recent article[1] in *TIBS* described the basics and potential pitfalls of linear and nonlinear regression analysis of biochemical data. As he pointed out, for some analyses, such as where the Michaelis–Menten equation is fitted to enzyme-kinetic data, there are two alternatives: fit the data directly to this equation by nonlinear regression, or transform the data into one of the straight-line forms and analyse by linear regression. Although the use of nonlinear regression in enzyme kinetics has increased greatly as appropriate computer programs have been distributed and published[2–8], papers continue to appear in which the analysis is based on what is often considered to be the worst method of all: a line drawn by eye in a double-reciprocal plot.

For some years I have been an advocate of nonlinear regression methods, but here I wish to explore a different view*: most of the supposed advantages of nonlinear regression analysis are illusory and the results obtained from a hand-drawn line after linear transformation may be almost as useful.

The advantages of nonlinear regression are assumed to be: (1) it avoids the inherent subjectivity of graphical methods; (2) it gives 'best estimates' of the kinetic parameters; and (3) it provides a measure (the standard error) of the accuracy of the parameters. This list is not exhaustive; it merely represents a personal opinion of what are generally held to be the principal advantages of nonlinear regression. These will be discussed in turn.

First, if enzyme kinetic data are plotted in one of the linear transformations, drawing a line 'by eye' is necessarily subjective. While this subjectivity is usually regarded as an undesirable element, it should not be imagined that regression analysis is completely objective. Weighting factors are chosen more on the basis of intuition than experiment and this will influence the results of the analysis. And, as noted by Leatherbarrow[1], when the data are not weighted this is still a decision about weighting; in this case it is a choice to weight all data equally. Sometimes observations may be discarded because they are 'aberrant' or are 'exhibiting substrate inhibition'. Whatever may be the validity of these decisions, they are almost always subjective to some extent. It is the experience of the investigator that makes this subjectivity acceptable, just as we accept that the investigator has sufficient competence to do the experiment properly. Thus, both graphical and regression methods contain subjective elements and any distinction based on degrees of subjectivity must involve a subjective quantification of subjectivity!

Second, nonlinear regression analysis (when properly weighted) gives 'best estimates' of the kinetic parameters; or so it is believed. When the analysis is done using a computer (i.e. always) an answer consisting of several digits is produced. This creates an impression of precision that is totally unjustified, much more so than is probably realised. If, in an analysis of some data, the K_a of an enzyme for its substrate was determined to be 1.43084 mM, it would be wise to discard some of the digits and round the value to 1.431, 1.43 or even 1.4 mM. In fact, even the last of these implies an unwarranted confidence in the determination. It might be disappointing but it certainly would not be surprising to obtain a value anywhere from 1–2 mM if the experiment were to be repeated.

Table I shows some data that illustrate this point; a series of routine determinations of a Michaelis constant that were made over a period of several months. The majority of the determinations are reasonably close to the mean value of 569 μM although the smallest value is 16% lower and the largest is 39% higher. Such variation is not alarming, but it suggests that we would be unwise to accept more than one significant digit.

Table I. Variation of a Michaelis constant over a series of experiments

Experiment no.	Michaelis constant (μM)	Standard error (μM)	No. data points
1	523 (504)	46	8
2	694 (606)	48	5
3	557 (654)	60	10
4	617 (559)	88	5
5	790 (787)	40	5
6	479 (448)	41	14
7	513 (541)	22	7
8	654 (577)	52	7
9	534 (538)	42	14

The Michaelis constant of *E. coli* prephenate dehydratase was determined from nine experiments[9]. The first value in the column labelled 'Michaelis constant' and the standard error were obtained by nonlinear regression using the DNRP53 computer program[7]. The value in parenthesis is the Michaelis constant obtained by drawing a line by eye on a double-reciprocal plot of the data.

Table I also shows the values of the Michaelis constant obtained by drawing lines by eye on a double-reciprocal plot of the same data. The differences from the values obtained by nonlinear regression range from negligible (experiment 1) to considerable (experiment 3). However, the differences between the Michaelis constant obtained by nonlinear regression and by the graphical method of any one experiment are smaller than the day-to-day variation. Thus, it is difficult to imagine circumstances in which the supposedly more accurate values obtained by nonlinear regression could lead to different biochemical conclusions. The values obtained graphically would seem to be no less useful than the best fit values.

The third of the reputed advantages of nonlinear regression analysis is that it yields a measure of the reliability of the parameters in the form of confidence intervals or standard errors. Table I illustrates the type of result that is obtained frequently; the standard errors give little indication of the variation to be expected from one experiment to another. For example, the highest value (experiment 5) is over two standard errors from the next highest (experiment 2). Clearly, the calculated standard errors are not indicating in any absolute sense the likely range of the 'true' value. Perhaps the only thing the standard errors do indicate is whether the experiment is well designed and executed[2].

Although I have argued that the usual justifications for nonlinear regression have little merit, I do not wish to abandon such analyses.

The Michaelis–Menten equation is not intrinsically nonlinear in that it is possible to transform it into the equation for a straight line. However, as discussed by Leatherbarrow[1], many equations that are fitted to biochemical data cannot be manipulated in this way. Often, in these cases, plots involving combinations of experimental observations, or extrapolated curves, tangents and asymptotes have been proposed. A typical example is the Hill plot, which requires an estimate of the maximum velocity, which is obtained by extrapolating a curve.

*Copies of a more detailed exposition of these views can be obtained from the author.

The problem with intrinsically nonlinear models is not that they are impossible to analyse graphically but because it is cumbersome to do so. By contrast, direct fitting is usually quick and simple, especially when using a personal computer that can now be regarded as a standard fitting in most biochemical laboratories. Even for nonlinear models such as the Michaelis–Menten equation, which can be manipulated to give a linear form for graphical analysis, there seems to be little reason for doing so. It is usually far quicker to fit data than to transform and plot them. To answer the question posed in the title to this letter – no, nonlinear regression is not a waste of time. In fact, it saves time and that is probably the single most important reason for using nonlinear regression.

References

1 Leatherbarrow, R. J. (1990) *Trends Biochem. Sci.* 15, 455–458
2 Cleland, W. W. (1967) *Adv. Enzymol.* 29, 1–32
3 Cornish-Bowden, A. and Endrenyi, L. (1981) *Biochem. J.* 193, 1005–1008
4 Canela, E. I. (1984) *Int. J. Bio-Med. Comput.* 15, 121–130
5 Crabbe, M. J. C. (1982) *Comput. Biol. Med.* 12, 263–283
6 Duggleby, R. G. (1981) *Anal. Biochem.* 110, 9–18
7 Duggleby, R. G. (1984) *Comput. Biol. Med.* 14, 447–455
8 Siano, D. B., Zyskind, J. W. and Fromm, H. J. (1985) *Arch. Biochem. Biophys.* 170, 587–600
9 Duggleby, R. G. (1990) *Anal. Biochem.* 189, 84–87

RONALD G. DUGGLEBY

Department of Biochemistry, University of Queensland, Queensland 4072, Australia.

GRAPEVINE

Pfizer Awards 1990

The research division of the pharmaceutical company Pfizer present six awards annually to support young scientists who have carried out meritorious research at British Universities or equivalent Institutions. The Awards, which are £4000 in 1990, are spread across all scientific disciplines that have potential application in the search for human or animal health drugs.

The 1990 winners include: **Professor B. K. Park** (University of Liverpool), for his contributions to the evaluation of the chemical basis of some drug-induced toxicity (allergy and hypersensitivity). **Dr A. C. Dolphin** (St George's Hospital Medical School, London), for her contribution to the understanding of the factors governing agonist/antagonist properties of calcium channel ligands. **Dr S. P. Newman** (The Royal Free Hospital, London), for his significant contribution to the understanding of the critical biopharmaceutical principles involved in the design of inhalation drug delivery systems. **Dr C. A. Maltin** (Rowett Research Institute, Aberdeen), for her work towards elucidating the regulatory mechanisms of protein metabolism in skeletal muscle and, in particular, her studies on the mechanism of the anabolic action of clenbuterol.

Frank Allison Linville's R. H. Wright Award

Dr John Hildebrand of the University of Arizona is the winner of the 1990 Frank Allison Linville's R. H. Wright Award. This $25 000 prize is awarded annually to an individual who has made outstanding progress in research in olfaction. Dr Hildebrand's research has made a significant contribution to understanding of the neurological bases of olfaction.

Nominations for future recipients of this award should be sent to: Dr B. P. Clayman, Dean of Graduate Studies, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6.