Analysis of dose-response curves in molecular pharmacology by the SAS System

Paola Baiardi¹ and Stefano M. Candura²,³

¹ Medical Informatics and ² Toxicology Unit, Clinica del Lavoro Foundation, Pavia Medical Center; ³ Department of Preventive, Occupational and Community Medicine, University of Pavia. Italy.

Abstract

When different levels of a stimulus are administered in a biological system, it is likely that the induced effect, quantal or quantitative, will be a function of the level of the stimulus. This is known as dose-response relationship.

In pharmacological studies of hormones, drugs and neurotransmitters, the transfer function between an input (log dose) and the related output (quantitative response or its metamer), is typically monotonic and S-shaped.

Morphological characteristics of this sigmoidal curve may reflect mono- or multi- compartmental systems due to the involvement of one or more receptor subtypes.

In order to describe pharmacological dose-response relationship, a good compromise between the empirical and the theoretical approach is the four parameter logistic function. According to the principle of effects superposition, a summation of logistic terms could be used to identify multi-compartmental models.

We developed a general non-linear (optionally weighted) least-squares curve-fitting routine based on PROC NLIN with Gauss-Newton algorithm. Goodness of fit is valuable on the basis of the residual variance and non-randomness of the residuals (Runs Test); comparison among models (one or more receptor sites) is obtained by F-test. Plotting utilities (SAS GRAPH) complete the analysis.

This procedure has proved useful to analyze dose-response curves from neurochemical experiments, in which the effects of single concentrations of agonists were assayed in the presence of varying concentrations of antagonists for models of one or more non-interacting sites.

Introduction

Methods investigating the biological properties of a compound by the analysis of its effects on living matter are called biological assays.

The typical bioassay involves a stimulus (a drug, a vitamin) applied to a subject (an animal, a piece of animal tissue) yielding a change in a measurable characteristic of the subject. This is the response (quantitative or quantal) of the subject which is a function of the level of the stimulus.

The relation between the dose of the stimulus and its effect may be expressed diagrammatically on a graph or algebraically by an equation. The estimation and interpretation of this relationship is of fundamental importance in biological studies [1].

Actually the experimental relationship is not exact, because the dependence of the effects on living matter often introduces variability between apparently identical measurements (variation within subject or between replicate subjects); so a statistical analysis of the resulting curve is advisable to use experimental data for the best.

It's to keep in mind that a good rule for planning experiments is first to identify the statistical principles that underlie the assay methods, to construct the experimental design and then to perform the statistical analysis. Its primary aim, coinciding with the aim of the assay, is the estimation of certain interesting parameters and their fiducial limits in order to ensure validity and reliability to the assay.

Studying dose-response relationship means to identify a mathematical model describing the characteristic of the curve, and to estimate the parameters of the model; if an opportune approximation function is used, these will coincide with the biological parameters under study.

Dose-response relationships are often encountered in many physiological and pharmacological studies, such as radioimmunoassays and binding experiments. The generated curves are typically monotonic, symmetrical and S-shaped when the dose is portrayed on a logarithmic scale.

Computerization has allowed an extended automation of the processing of this kind of
experimental data, and many programs for dose-response curve analysis are commercially available; for example, binding experiments can refer to LIGAND [2] while simultaneous analysis of families of sigmoidal curves finds application in ALLFIT [3]. The former approximates the experimental dose-response relationship using the exact physical-chemical model of the ligand-binding system [4] and therefore is specific for such experiments, while the latter uses a more general semiempirical model in solving a particular, even if common, problem [5].

When the shape of the curve (shallow) gives indication of more complex experimental situation (involving, for example, one or more receptor subtypes) the basic models may fail to converge or may give non consistent parameter estimates, so one needs a more precise model.

The present paper is dealing with modeling dose-response relationship and curve fitting based on the four parameter logistic equation with particular regard to experimental pluri-compartment systems (corresponding to multiphasic curves) and to model comparisons (one vs more compartments).

A study case analyzing curves from neurochemical experiments in which the effects of single concentrations of agonists were assayed in the presence of varying concentrations of antagonists is also shown.

Experimental design

The theory of experimental design is wide [6] and its treatment lies outside this paper, nevertheless some general considerations about planning experiments are worth remembering.

It's clear that a good experimental design together with an adequate statistical analysis leads to optimal assay results, but no statistics can detect inconveniences altering or invalidating the response. The assay has to be designed taking into account the occurrence of probable sources of disturbances.

Moreover, the experimenter says he is not able to predict all what he will obtain from an assay, and hence difficulties may arise in supplying the statistician with whole the information needed for a precise design of the analysis. On the other hand, the statistician himself often fails to explain his point of view with sufficient clearness.

Only through a good and continuous collaboration between the two, statistical science may supply added value to biological or chemical technique, getting useful and significative results.

Statistical analysis could be an expense for an assay, and it must be considered when programming the experiment, either this cost is measured in monetary units or on any other scale (expenditure of time, labour). The experimenter often tries to minimize the analysis cost, but if the assay is carried out as a research project, whose primary aims are precision and reliability of the results, no consideration of design complexity should be made.

Finally, the experimenter should have a friendly, versatile computerized program at his disposal, so he will be able to analyze his own data quickly and easily, when the assay has been well designed.

In general, the design of an assay includes all the matters relating to the structure of the experiment: the choice of dose levels, the allocation of subjects to doses, the order in which measurements are made and other experimental constraints.

Its cost is very high in respect to the cost of data processing and usually the experimenter resources are limited. Past experience or pilot studies should be used to get indications about some choices in the design (dose levels, number of subjects at each dose, magnitude of the parameters to be estimated). A compromise has to be achieved between marginal savings, aimed at control of expenditure, and design perfection, in order to get the best results, remembering that sometimes a theoretically but complex design can give worse results than a simpler one, because of the risk of frequent errors (due to complexity).

In any case, the aim of designing is to ensure validity to the assay, that is, consistent estimates of the involved parameters. The estimates must not be only the results of some formulas applied to the observations, but they must also be correctly interpretable as a property of the examined substance.

The choice of subjects and the kind of response to be measured play an important part in reducing the risk of assay invalidity, too.

The stimulus may have a general effect on some characteristics of the subject and, sometimes, a gross characteristic may be more easily measured than an other more accurate one, but involving, for instance, the subject destruction. Moreover a species of subject can be more "responsive" than an other. Also in this case, balancing costs and precision is needed.

At last, the choice of metamers to investigate (usually the logarithm of the dose and the measured response or a combination of more than
Modeling

Modeling dose-response relationship in bioassay is an age-old problem and a lot has been written on this matter [7,8,9,10,11,12,13].

In pharmacological studies of hormones, drugs, and neurotransmitters the curve is smooth, symmetrical and sigmoidal when a log transformation of the dose axis is considered.

The characteristics of this dose-response relationship can be investigated by both a theoretical and an empirical approach [14].

Theoretical models are descriptive of the exact physical-chemical reaction involved. Usually they are based on the first order mass action law and they give information about parameters concerning the drug-receptor interaction (affinity constant, receptor concentration); however, they often involve many parameters, thus requiring many experimental data to provide a good curve fitting.

Moreover they are perfect for an ideal situation (equilibrium achievement) which may not be fulfilled in an experimental system [15].

Starting with the experimental data, the problem is approached in an empirical way. Several methods are used to define the relationship.

Curve fitting by spline or polynomial functions may give good results, since this methods honourably approximate every kind of dose-response, without forcing to a specific shaped curve. On the other side, these functions discard previous knowledge coming from other experiments or models, and do not allow a qualitative experimental data evaluation: these are the main disadvantages of such methods.

The simplest, and most common, empirical method is linearization of the sigmoidal relationship, that can be obtained by an opportune coordinate transformation (probit, logit, arc-sine). Linearization is advantageous since it involves only two parameters, permits easy tests of random or systematic deviations from the model, and allows computational facility. However, a linear relationship does not provide information on the extremities of the curve, and is inadequate to characterize complex experimental situations.

Combining the two approaches on the bases of mathematical principles, a good model for the curve is defined: the problem of curve approximation becomes a problem of parameter identification of a transfer function between the input (dose or log dose) and the output (response or its metamer).

A four parameter logistic equation [12,16,17,18] is an optimal model for this purpose, since it contemporary represents an evolution of the empirical methods and an approximation of the theoretical ones. The parameters involved have a biological meaning and can be directly used for the assay characterization.

This represents a significant improvement over the three parameter version [19] since the base-line level becomes a parameter to be estimated, and it is not a priori known constant.

The logistic model reduces some of the above mentioned method disadvantages having a limited number of parameters, while maintaining the agreement to the real situation and properly characterizing the minimum and maximum dose regions.

Curve fitting involves non linear regression procedures which give more consistent parameter estimates in a not necessarily more complex and time consuming way.

Moreover, the model appropriately describes a great number of experimental situations [10,11].

If the presence of complex situations (such as cooperativity or allosteric effects) is suspected, the logistic model can be generalized by adding some parameters describing a hypothetical degree of asymmetry of the curve. A summation of logistics can be used in case of sigmoid relationships with steps and bumps reflecting the presence of multiple sites [9].

Weighting

In order that the relationship between the dose and the response can constitute the basis of the assay, some assumptions about the characteristics of the regression function must be valid.

Among these, a condition of homoscedasticity (variance of the response independent of the expected response) must be objectively verified. Barlett's test (properly adjusted, if necessary) is useful for this purpose, even if it could be inefficient, when few observations, and hence low degrees of freedom, are available.

Some transformations of responses (scedasticity transformation) may yield the regression homoscedastic.

In any least-squares method to perform regression analysis, when assumptions regarding uniformity of variance fail, an appropriate weighting function is advisable to compensate for the lack of variance uniformity [20].

If the dose-response relationship is analyzed by linearization (logit-log method), logit transformation introduces severe non uniformity of
variance. Performing an unweighted regression, in this case, could produce erratic results, not only as regards parameter estimates, but also for the shape of confidence limits around the curve. Usually logit-log underestimates the errors at the extreme of the curve and overestimates the errors at its centre. Curve truncation to the central region (between 20% and 80% response) constitutes a simple form of weighting.

The principle of weighting is the assignment of an opportune weight to each observation, according to the information it contains; for this purpose, weights inversely proportional to the variance of the observations make more precise points (whose variance is smaller) get larger weights. The problem is to identify a model describing the variance of the response variable as a function of the response level.

Since, usually, measurements are taken in duplicate or triplicate (a higher degree of replication increases the reliability of the response and the precision of the sample standard deviation), a plot of the response variance (var(y)) versus the mean response (\(\bar{y}\)) may draw to the identification of a systematic relationship between these quantities. If such relationship exists, non-uniformity of variance of the observations can be attested.

A linear, quadratic (generally polynomial) or a power function well describes the relationship in most experimental situations: a linear model can be used for Poisson-like variables, for which the variance is directly proportional to the expectation of the response, while a quadratic model can fit the case of a constant variation coefficient.

After testing different variance models, the one with the smallest residual variance and with positive predicted variance values must be chosen.

This form of weighting is based on the predicted, rather than observed, variance of the observations and it is advisable when few replicates and hence low degree of freedom are available.

An iteratively reweighted regression is needed for the original dose-response curve fitting.

Variances are calculated from the model on the basis of the response variable, therefore weights must be recalculated, after each iteration, according to the response level predicted, for a given dose, in the previous iteration. In this way, the weights become an implicit function of dose level, and this ensures that all replicates for the same dose will receive the same weight. This is the principle of "iterative reweighting".

Moreover, if no replicates are available, the analysis of the residuals from an unweighted dose-response fit can give indication on the necessity of weighting and on the nature of the relationship between var(y) and \(\bar{y}\).

Finally, when non-uniformity of variance is introduced by some variable transformation, theoretical formulas for the approximations of the predicted variance and corresponding weights are available.

Weighting should be used in all methods of least squares curve fitting: if non-uniformity of variance has been detected, using an unweighted four parameter logistic regression, can result, for instance, in an underestimation of the error in the low dose region, and in an overestimation in the high dose one.

In general failure in using weights does not result in serious errors in the estimate of dose-response curve parameters but of their confidence limits.

**Goodness of fit**

In the previous paragraphs a general theory of modeling and curve fitting has been presented, because of the extreme adaptability of these general principles to various experimental situations and their basic importance for a correct interpretation of the assay results.

As regards the goodness of fit, some appropriate objective criteria for both a test of a significative deviation of the observations from the fitted curve, and the identification of the best model fitting the experimental data are given [14,4,5], according to what applied in the study case.

Several criteria are available to evaluate goodness of fit. The quantities coming into play are the variance of the deviations of the points around the curve (residual variance) and the root mean square error (RMS error) corresponding to the standard deviation of the observations around the regression.

The correlation coefficient, in this case, is not a discriminant figure since, in such a situation, the proportion of variance explained by the regression is very high (reflecting correlation coefficient usually greater than 0.9): the discriminant parameter is how much variability is unexplained by the fit, that is the RMS error.

A classical analysis of variance comparing the "between-dose" and the "within-dose" (error) variance can be performed, and a significant
source of variation of dose means can be established via the F test.

Non-randomness of the residuals (deviations of predicted from observed responses) need also to be investigated.

The runs test can be performed to compare the number of sign changes in the regression to the expected distribution in a random sequence of residuals. Data points are expected to be randomly distributed around the fitted curve. A number of runs less than expected is a signal of nonrandomness and hence of an inappropriate fit.

When the observation number is small, the mean square successive difference test (comparing the variability of successive points with the total one) can be used in order to evaluate data dispersion around the fitted curve.

Comparison between models applied to the same experimental situation is performed on the basis of the smallest residual variance.

When models involve different numbers of parameters, comparison is made via an approximate F test based on the "extra sum of squares principle" in order to detect a significative increase of the goodness of fit for the model with more parameters.

However, when further parameters are added to a model, the goodness of fit will improve (decreasing residual sum of squares) simply because of the increased flexibility of the model. As an extreme case, fitting a model with a number of parameters equal to the number of observations will give the exact fit.

The F test of the "extra sum of squares principle" takes into account both the obvious reduction of residual sum of squares and the difference in degrees of freedom (number of points minus number of parameters).

A one-tail F test can be performed, since the differences between the residual sum of squares from both fittings are always positive.

Measurement and estimate combination

Any assay involves four kinds of parameters [1]:
1) biological parameters (i.e. potency, affinity constant)
2) parameters of the dose-response mathematical model
3) parameters defining the metametric transformation
4) parameters describing the relationship between the variance of the response and the response itself.

Usually, data from an assay are not sufficient to provide consistent estimates of all of these: the problem of pooling information from assays arises.

If several curves within an experiment are available (for instance, a standard against several test preparations), a simultaneous analysis of all curves can be performed. In this way, information on the examined dose range may be pooled, and the advantages of the constrained fit can be exploited. Forcing the curves to share some parameters (upper or lower plateau, slope factor), provided that constraints are legitimate, a substantial reduction in the number of estimated parameters is obtained.

Besides, combining information from many curves can be useful to investigate the nature of the variance model and to identify an appropriate weighting function.

Pooling data from multiple experiments is also possible. The aim is to improve the definition of a structural model for experimental data, since conclusions drawn from a single new type of assay cannot be reliable enough.

Particular care is required in controlling the "between-experiment" variation in order to reduce it to the "within-experiment" variation level. Another potential problem is the non-independence of the observations [21]. However, if these variability sources are controlled, the precision of the parameter estimates is certainly improved.

Alternatively, it is possible to perform a series of separate assays and combine the results. If the assays are well designed, a good agreement between several estimates should be found and a composite figure is required.

In order to award more importance to the most precise and reliable individual values, a weighted mean is recommendable, taking weights inversely proportional to the parameter variance.

Finally, a test for the homogeneity of the single estimates can be performed (Chi-square distribution).

Anyway, the assay validity, in terms of consistent parameter estimates, must be pursued.

It may happen that the same data set gives slightly different estimates of both parameters and fiducial limits, in conformity with the chosen estimate procedures and metametric transformations.

The analysis and hence the result objectivity should not be questioned, if fiducial limits are interpreted as guides to the reliability of the estimates, and not in the strict mathematical meaning of the inference theory.

The curve fitting procedure

The basic principle of the statistical analysis of biological assays, above discussed as a general
treatment, have been applied to carry out a curve fitting routine for the study of sigmoidal dose-response relationship. This type of S-shaped curves are typical of the pharmacological field [10,11]: an application in molecular pharmacology is reported in the next session as an example.

The procedure is based on non linear regression analysis (PROC NLIN), performed by the Gauss-Newton iterative method [22].

The four parameter logistic equation was chosen as model for sigmoidal curves. Its form is:

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

(1)

where X is the arithmetic dose, Y the response and the parameters respectively represent:

- \(a\) the maximal response
- \(b\) the slope factor
- \(c\) the dose resulting in between a and d halfway response
- \(d\) the minimal response

The slope factor \(b\) can assume negative or positive values in relation to the increasing or decreasing trend of the response for increasing doses.

Generalizations of the four parameter logistic model [9] were made to properly describe experimental situations entailing a certain degree of asymmetry or shallowness of the curves.

Shallow curves are often due to the presence of multicompartmental systems, for which a summation of logistic terms is a more suitable model. In this case, a proper generalization of (1) is the following:

\[ Y = \sum_{i=1}^{n} \left( \frac{a_i - d_i}{1 + (X / c_i)^b_i} + d_i \right) \]  

(2)

This model has been implemented in the curve fitting routine up to \(n = 3\), accordingly with the needs of our experimental situations, but no limit is intrinsically present.

In order to widen the range of convergence, and to prevent some parameters from becoming negative during the iterations, the model has been reparameterized by [13]:

\[ c_i = \ln (c_i) \]

When \(n > 1\), 4xn parameters should be estimated, and the number of required observations increases. The model may be constrained (BOUNDS option) on the basis of consideration about the biological phenomena under study, obtaining a reduction in the number of parameters.

Typical constraints are:

\[ \sum_{i=1}^{n} a_i = a \]
\[ b_i = b \forall i \]
\[ d_i = d \forall i \]

Data entry can be made directly into SAS environment or through conversion from external (ASCII, DBF) file to SAS data set, considering \(Y\) expressed as raw data or as percentage of the maximum response level (\(a = 100\%\)).

Each single observation or the replicate means can be used to fit the sigmoidal curve. The choice (replicates or means) is based on the comparison between the "within-dose" and "between-dose" variances. Since the former is significantly smaller than the latter, means of replicates are advised as input for the regression, provided that sufficient data are available to ensure convergence. If the two variances are comparable, each observation should be entered.

In case of non uniformity of the variance of the observations, a weighted regression analysis (WEIGHT option) based on a linear, quadratic or power function as the variance model is allowed.

The routine lets residuals and residual sum of squares be stored to perform the runs test and the F test of the extra sum of squares principle to obtain an objective evaluation of the single regression and the comparison between different complexity models for the same experimental data. The latter test has been largely applied in the study case to verify the existence of a second class of receptor subtypes.

Two corollary simple programs complete the analysis procedure. The first carries out the estimation of initial parameters (b,c) for non linear regression via:

1- The fit of a regression line in the plane \(\log(X),Y\) performed in the portion of curve between 20% and 80% of the response. This is for monocompartment situations.

2- The fit of a segmented model (two straight-lines) in the same coordinate system of point 1- . This is for multicompartment cases.

The second program, still in progress, is concerning to the existence of a non homoscedasticity in variance, and to the comparison between models.

Final parameter estimates from the logistic model are memorized to be supplied to a graphic routine (PROC GPLOT), which completes the curve fitting procedure. In case of multicompartment systems, a composite plot
global model and each compartment curve are particularly effective since it allows the experimenter to identify the activation regions of each receptor site. Moreover, when a certain degree of uncertainty characterizes the fitting, this kind of plot may give information about supposed probable presence of pluri-compartment system, whose objectiveness is not yet attained, due, for example, to the scarcity of experimental points.

The representations are accurate and features enhancing the plot quality are available, since often graphic outputs are used not only during the study stage, but also to be published.

At present the procedure runs in MS-DOS environment on a PC 386xx, but it will be soon ported to Windows system according to the standards of our Institution.

Study case: analysis of curves from neurochemical experiments

Transmission of messages between cells in the nervous system involves the release of a chemical neurotransmitter and its subsequent recognition by a second cell. The interaction with a specific receptor will generate a transmission signal which will ultimately lead to a physiological response.

Drugs that act on a physiological receptor mimicking the effects of the endogenous neurotransmitter are called agonists. Other compounds may bind to the receptor but have no intrinsic activity; the result of such binding may be interference with the effect of an agonist. Compounds that are themselves devoid of intrinsic activity but cause effects by inhibition of the action of an agonist (e.g. by competition for agonist binding sites) are termed antagonists.

Acetylcholine (ACh) is a major neurotransmitter in the mammalian brain. It may bind to different kinds of receptors (nicotinic or muscarinic). Muscarinic receptors are heterogeneous and five genes have been identified in the mammalian genome encoding for them.

The binding of ACh to muscarinic receptors induces the hydrolysis of cell membrane phosphoinositides leading to the intracellular formation of inositol phosphates, which exert second messenger roles.

The data of this study case were generated by neurochemical assays in which 3H-prelabelled cerebral cortex slices from 7-day-old rats were incubated with a stimulating concentration of a muscarinic agonist (carbachol or betanechol), in the absence and in the presence of increasing concentrations of the antagonist pirenzepine. The experiments were carried out in triplicate. Data were expressed as percentage of basal (unstimulated) 3H-inositol phosphates formation. This was the measured response.

In the following, only data from two representative experiments are shown. The assays were repeated twice with similar results.

These experiments are part of a research project, which is examining the effects of muscarinic agonists and antagonists on phosphoinositide metabolism in the developing brain, in an attempt to evaluate the role of muscarinic receptor subtypes in this process.

Pirenzepine inhibited muscarinic agonist-stimulated 3H-inositol phosphates formation in a concentration-dependent manner and a dose-response relationship was identified.

Data were previously analyzed by a classical approach calculating the Hill coefficients (nH) [25] for pirenzepine inhibition of agonist-stimulated 3H-inositol phosphates formation. This coefficient gives an indication of the slope of the curve; when it is less than unity suggests the presence of an heterogeneous receptor population.

When betanechol was used, the resulting pirenzepine curve was steep with an Hill coefficient close to unity. On the other hand, pirenzepine inhibition curve of the response to carbachol was shallow, with nH<1. Thus, it is likely that more than one muscarinic receptor subtype is involved in the activation of phosphoinositide hydrolysis in the immature rat brain.

The curve fitting procedure has proved useful in verifying this hypothesis. Accordingly to small "within-dose" variance, means of triplicates were used as input data (11 concentrations). Barlett's test for homoscedasticity was not significative in all the examined experiments, and unweigthed regression was performed.

Experimental results were fitted to both single and pluricompartment systems, in order to identify the model which best described the data.

A four parameter logistic model significantly described experimental data for betanechol, while a summation of logistic terms resulted a better model than the simple one for carbachol confirming the existence of a pluri-compartment situation.

In this latter case, the interest was in analyzing the contribution of each receptor to the measured response by decomposition of the dose-response relationship into a summation of two compartments with unitary slope.

Really slope factor b of the logistic function (1) cannot be interpreted in the same way as the Hill
Fig. 1: Example of bicompartent system splitted into its two significative components.

Fig. 2: Example of monocompartent system fitted by a six parameter logistic equation. Existence of a non significative component.
coefficient, but it gives an indication of the slope of the curve too.

Thus, experimental data were fitted to a constrained six parameters model (2) with $b_1=b_2=1$ and $d_1=d_2=d$. Constraints didn't worsen the fit (F test >0.05) and then, for each term of the coefficient, but it gives an indication of the slope of the curve too.

Plot utilities were used to graphically illustrate the experimental situation (Fig.1).

Fitting a two compartment model for bethanechol (Fig.2) resulted, also from a graphical point of view, in a lack of heterogeneity explained by the constancy of a component in the summation.

Conclusions

The procedure has been use to analyze many molecular pharmacology experiments. Even if the program was designed to be easy to use, often experimenters prefer the statistician runs it, owing to their traditional embarrassment with all the matters concerning statistical science.

As the procedure is home-made, we suppose it will be a helpful tool in other fields such as RIA or quantal response studies, since the differences this kinds of problems show can be worked out with a small adapting effort.

References