SAS AUTOMATED IC50 PARAMETER ESTIMATION WITH REPEATED MEASURES DATA

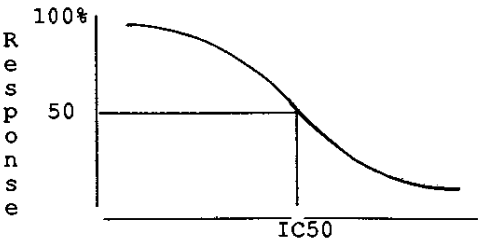
Dori Emmrich, Syntex
Linda Archer, Syntex
Chuck Villarrubia, Syntex
Greg Silva, GW Consulting

In researching new pharmaceutical compounds, the basic research group at Syntex conducts dose ranging experiments involving repeated measures sampling. Often these studies involve a small number of experimental units, with several doses of a compound administered to each unit. Parameters such as the 50% inhibitory concentration (IC50) or the 50% effective dose (ED50) are estimated using non-linear regression techniques. The IC50 is the concentration of a compound which will inhibit a response by 50%. The ED50 is the dose of a compound which elicits 50% of the maximum response. Due to the non-independence of the repeated measures, an adapted Seemingly Unrelated Nonlinear Regression (SUNR) method, developed by Keith Muller and Carolin (Malott) Frey, University of North Carolina, Chapel Hill, 1989, is used in estimating these parameters.

Due to the complexity of the statistical calculations performed for the SUNR method, as well as the large number of dose ranging studies being conducted, it was necessary to automate the process using SAS programming software. For ease of use by many clients in basic research, the system was automated further using the menu interface capabilities of SAS/AF* software and screen control language.

The Problem

The basic research experiments typically involve two or more compounds administered in a range of 5 to 12 doses, with each experimental unit (e.g., tissue, or animal) receiving each dose of the compound. The response is then measured at each dose, and a graph of the data will show the typical dose-response relationship. An IC50 is a standard measure of comparison among different compounds.



Due to the non-independence of the repeated measures, estimation of the standard error of the IC50 and the other model parameters must account for both the variation in the multiple responses within a subject as well as the variation in responses between subjects. A version of the seemingly unrelated nonlinear regression method is used for this purpose. This SUNR method assumes that the covariance matrix of the residuals is compound symmetric. That is, it assumes equal variances and equal correlations between residuals.

The System

The software system provides menu options allowing the user to enter, edit, and archive data, calculate means, perform the dose response model fitting and IC50 calculation, and plot dose-response curves. The user interface is a menu driven shell developed in SAS/AF.

Main Menu

1. Enter or edit data
2. Calculate means
3. Run SUNR
4. Exit

Enter and Edit Datasets

The user can enter compound, dose, and response information for each experimental unit or subject, directly into SAS datasets via the data entry screen in the format shown below:

Data Entry/Edit

COMPOUND	DOSE	ID1	ID2	ID3

Each dataset must have a unique identification number. After the data have been entered and verified, two data transformations are performed before fitting the nonlinear model. The first is the calculation of the $\log_{10}$ doses. The second is an orthogonal transformation of the responses, to eliminate the dependent structure of the data.

To perform an orthogonal transformation, the orthogonal polynomial coefficients corresponding to the number of doses in the experiment are obtained using the SAS IML subroutine ORPOL. This subroutine requires that one vector of coefficients be provided to it. The linear vector's coefficients are most easily obtained. Given an even number of doses, M, the vector will have the form: [-(M-1), -(M-3), ..., -1, 1, ..., (M-3), (M-1)].
If the number of doses is odd, the vector will have the form: [-(M-1)/2, -(M-3)/2, ..., -1, 0, 1, ..., (M-3)/2, (M-1)/2].

The transformed data are then calculated using the following equations. First, denote the vectors of orthogonal coefficients as Vi1, Vi2, ..., ViM for i=1,...,M. Specifically, Vi1 is the vector of 1's (divided by the appropriate constant), Vi2 contains the linear coefficients, Vi3 contains the quadratic coefficients, etc. Also let Di be an indicator variable denoting which orthogonal component (linear, quadratic,...) has been calculated. Then, letting Y1, Y2, etc. represent the responses for a given individual, set Y* = V1j*Y1 + V2j*Y2 + ... + VMj*YM and Dj=1 and Dk=0 for all k<>j.

As with the input data set, the transformed data set has M observations for each subject. But now each observation contains one transformed response (Y*), M indicator variables (Di) and all M $\log_{10}$ doses.

Mean Calculations

When the user chooses this option, a table is generated of means and standard deviations presented by treatment and dose. Obtaining means is useful for determining starting values for the parameters in the dose response model.

SUNR Analysis

Analysis Menu

1. Run SUNR
2. View output
3. Print output
4. Plot Data

Dose Response Model

When the user chooses the Run SUNR option, a screen appears which allows the user to indicate which parameters will be included in the model. The software then generates starting values for these parameters and fits the model in a two-step procedure described below.

The nonlinear model is fit using PROC NLIN in SAS, using the DUD method. Given parameter starting values, the predicted values are calculated at each $\log_{10}$ dose using the following equation: F_i = MIN + (MAX - MIN)/(1 + exp(-(x_i - IC50)/SIGMA), where MIN and MAX represent the minimum and maximum responses, SIGMA is a curvature parameter and x_i denotes the ith $\log_{10}$ dose.

Parameter starting values are estimated as follows: MIN and MAX are estimated by the average response at the minimum and maximum doses, respectively. The starting value for SIGMA is 1 for an increasing trend, and -1 for a decreasing trend. The IC50 is estimated by the concentration corresponding to an average response closest to (MAX-MIN)/2.

The model with Y* as the response is then:

```
Y* = D1*(V11*F1 + V21*F2 + ... +
VM1*FM) +
     D2*(V12*F1 + V22*F2 + ... +
VM2*FM) +
                    ...          +
     DN*(V1N*F1 + V2N*F2 + ... +
VMM*FM)
```

After fitting the model, the following calculations are performed: 1) the residuals are calculated as the observed response minus the predicted response: RESIDi = Yi - PREDi. 2) The covariance matrix on these residuals is obtained using PROC CORR with options COV and NOPROB. 3) From the elements in the covariance matrix, a correlation coefficient (RHOHAT) is calculated as follows:

```
RHOHAT = [(sum(all elements) -
TRACE)/(M*(M-1))] / S2,
```

where "all elements" refers to all elements of the covariance matrix, TRACE is the trace of the matrix (=sum of diagonal elements), M is the number of doses and S2 is a variance estimate equal to TRACE/M. 4) A weight for each observation which equals 1/VAROBS where VAROBS is defined as:

```
VAROBS = D1*(1+(M-1)*RHOHAT) + (sum(of
D2-DM))*(1-RHOHAT).
```

After completion of these calculations, the model described above is fit again but this time as a weighted analysis where the weight for each observation is 1/VAROBS. The parameter estimates should change only minimally, but the standard errors of these estimates may change substantially since they now account for both the within and between variation in the data.

The software will indicate whether or not the model converged. If convergence is not achieved, the user may choose to provide new starting values and try refitting the model.

Output Specification

A plot displaying the predicted dose response curves is output, as well as a table displaying the estimates, standard errors and 95% confidence intervals of all parameters in the final dose response model.

Archive Data Sets

This menu option allows the user to save datasets to floppy disks, according to experiment ID. A table directory of all datasets is maintained by the system.

Conclusion

SAS software was utilized to automate the tedious and complicated mathematical calculations necessary to perform the statistical analysis of dose-response data. A menu-driven system was developed for use on a routine basis by clients in basic research.

* SAS/AF is a registered trademark of SAS Institute Inc., Cary, NC, USA.