USING SAS® SOFTWARE TO SIMPLIFY THE DISPLAY AND THE INTERPRETATION OF LABORATORY DATA

Kathleen H. Arthur, Sterling-Winthrop Research Institute

Abstract

This paper discusses the SAS® MACRO used to support the technique for data analysis proposed by Sogliero-Gilbert, Mosher, and Zubkoff. Their paper "A Procedure for the Simplification and Assessment of Lab Parameters in Clinical Trials", Drug Information Journal, Vol. 20 (1986), introduced two concepts to simplify the display and the interpretation of lab parameter data: 1) a normalization of each lab parameter value so that it can be interpreted without reference to its normal range and 2) a multivariate scoring system that combines lab parameters into functionally related groups. The SAS code used to carry out the normalization and multivariate scoring will be provided upon request.

Introduction

The monitoring of lab parameters is important to all clinical trials as an evaluation of the safety of a new drug by identifying possible abnormalities in any patient(s). When abnormalities are identified, a decision must be made whether these are due to the drug being tested or to other circumstances. Identifying abnormalities requires a knowledge of the lab normal range for each lab parameter as well as the laboratory used; often patients will have different lab normal ranges for the same parameter due to different laboratories. Thus, if every patient's lab value were studied along with the value's normal range, there would be a large amount of data to inspect in order to answer the drug safety question. Sogliero-Gilbert, Mosher, and Zubkoff[1] have proposed a way to simplify the analysis of lab parameter data, and this paper discusses their technique for analysis and provides SAS® code to implement their procedure.

The Simplification Technique

The following is a summary of the Sogliero-Gilbert, et. al. paper. Further details of their proposal can be found by consulting the original paper.

Two concepts used to simplify the assessment of lab parameters are as follows:

1) the normalization of lab parameters by dividing each lab value by its upper limit of the normal range. This range is established by the laboratory producing the measurement;

2) a multivariate scoring system that assigns to each patient a number produced from the normalized parameters over a functionally related group of parameters (e.g., hepatic group).

The transformation chosen for normalization (i.e. division of each value by its upper limit of normal range) satisfies the following:

1) All transformed values are positive.

2) All parameters have an upper normal limit of of 1 and all values are bounded below by 0.

3) Values outside the upper limit are easily found.

The introduction of normalization alone enables identification of abnormalities at the upper end by noting values greater than 1. However, identification of abnormalities at the lower end is not determined independent of the normal range lower limit; this situation is corrected by observing only the deviations from the normal range in the multivariate scoring system. Using the normalized values, percents above and below the normal range are easily calculated. Values within the normal range have a deviation of 0.

For the multivariate scoring system, Sogliero-Gilbert, et. al. have created the Genie Score, which uses the normalized deviations from the normal range of a combination of related parameters (e.g. hepatic parameters, renal parameters, etc.). The Genie Score is a global score for the group of functionally related parameters, indicating abnormalities for the whole group as opposed to an individual lab parameter. The Genie Score is defined as following:

Consider a group of N functionally related lab paramenters. Let

i = 1, ..., N parameters.

X_i = measurement associated with the i^{th} parameter.

$X_{i(UL)}$ = upper limit of normal range for i^{th} lab parameter.

$X_{i(LL)}$ = lower limit of normal range for i^{th} lab parameter.

Z_i = $X_i/X_{i(UL)}$, the normalized value of X_i.

$Z_{i(LL)}$ = $X_{i(LL)}/X_{i(UL)}$, the lower limit of the normalized range.

$Z_{i(UL)}$ = 1, the upper limit of the normalized range.

Then, let the normalized deviation for each i^{th} parameter be D_i, where

$$D_i = \begin{cases} Z_i - 1, & \text{if } Z_i > 1.0 \\ 0, & \text{if } Z_{i(LL)} \le Z_i \le Z_{i(UL)} \\ Z_i - Z_{i(LL)}, & \text{if } Z_i < Z_{i(LL)} \end{cases}$$

Therefore, if D_i is positive then X_i is above the upper limit of the normal range; if $D_i=0$ then X_i lies within the normal range; if D_i is negative then X_i is below the lower limit of the normal range. The deviations are expressed as percentages above or below the normal range.

The Genie Score (GS) for the group of functionally related parameters is then defined as follows:

$$GS = K * [\frac{1}{N} \sum_{i=1}^{N} S_i W_i |D_i|]$$

where

i = 1, . . ., N.

S_i = Stretch factor = $\begin{cases} 1, & \text{if } D_i \ge 0 \\ 2/Z_{i(LL)}, & \text{if } D_i < 0 \end{cases}$

W_i = weight assigned to each lab parameter.

D_i = normalized deviation.

k_1 = (1+0.2(NSP)), where NSP = number abnormal parameters.

k_2 = [1-0.1(N-NSP)].

K = $k_1 * k_2$.

The weight, W_i, if known, is selected to incorporate the relative clinical importance of the lab parameter in the functionally related group. A convenient set of weights would be one where the individual weights sum to 1; these weights should be determined by the person most familiar with lab parameter values. The stretch factor, S_i, is used to compensate for the fact that small deviations below the lower limit for some parameters might be as serious as much larger deviations above the upper limit for some other parameters.

A GS>0 indicates there was at least one abnormal lab parameter; a GS=0 therefore indicates there were not any abnormal lab parameters. Interpretation of a non-zero GS depends on the individual D_i's that make up the GS.

Program for Calculating Genie Score

The program for calculating the Genie Score (GS) uses %MACRO and gives as output for each patient with a non-zero GS a table containing their GS for each group of functionally related parameters at each time of observation along with the normalized deviation for each lab parameter that contributes to the GS. Two plots are available to describe the normalized deviations and GS: a plot of the actual values over time and a plot of the change from baseline over time. Additionally, bar charts of the normalized values and GS are available. The tables, plots, and charts can be outputted by an optional classification variable (e.g., treatment).

This program assumes that clinical trials have more than one observation time at which lab parameters are measured, and therefore requires an observation time variable.

The following is the MACRO call for the GS calculations and the definition of the MACRO variables:

```
%GENIE(DATSET=, GRPVAR=, LABVAR=,
LABVAL=, WGTVAR=, LOWER=, UPPER=,
IDVAR=, TIMEVAR=, CLASSVAR=, BYVARS=,
LABLBL=, IDLBL=, TIMELBL=, CLASSLBL=,
FMTVAR=, PLOT=NO, BASE=, CHARTS=NO);
```

where

DATSET = SAS data set name. Required.

GRPVAR = Name of variable (e.g. Group, Type, etc.) which contains all of the group names for the functionally related parameters (e.g. Hepatic, Renal, etc.). If this variable is not specified the program assumes that all of the lab parameters are in the same group. Optional.

LABVAR = Name of variable (e.g. Labs, Chemistry, ect.) containing the lab parameter names (e.g. BUN, SGOT, etc). Required.

LABVAL = Name of variable containing lab parameter values. Required.

WGTVAR = Name of variable containing the weights, W_i, for the parameters that are members of the group. The default value is for every parameter to have an equal weight where

$$\sum_{i=1}^{N} W_i = 1.$$

LOWER = Name of variable containing the lower limit of normal range. Required.

UPPER = Name of variable containing the upper limit of normal range. Required.

IDVAR = Name of variable identifying individual patients. Required.

TIMEVAR = Name of variable containing observation times. Required.

CLASSVAR = Name of classification variable (e.g. treatment). Optional.

BYVARS = Name(s) of BY variables to be used in BY statement. Optional.

LABLBL = Label for the LABVAR variable. Optional.

IDLBL = Label for the IDVAR variable. Optional.

TIMELBL = Label for the TIMEVAR variable. Optional.

CLASSLBL = Label for the CLASSVAR variable. Optional.

FMTVAR = Name of the format variable (from PROC FORMAT) for LABVAR. Optional.

PLOT = NO/YES. Choose YES if plots of each individual patient are desired. Gives plots of normalized deviations and GS for each time observation. The plots are generated from SAS/GRAPH®. NO is default.

BASE = If the input data has baseline values, and a plot of change from baseline vs. time for the Genie Score and the normalized deviations for each individual patient is desired, then define this variable as the name of the baseline time contained in the TIMEVAR variable. For example if the name of the baseline time is PRE, B, 0, etc., then BASE would be defined as PRE, B, 0, etc. The plots are generated from SAS/GRAPH®. Optional.

CHARTS = No/Yes. If a bar chart of the normalized deviations and a bar chart of GS broken down into its weighted normalized deviations is desired, choose Yes. The charts are generated from SAS/GRAPH®. No is default.

Example

In a clinical study with 112 patients, drugs A and B were compared. Lab parameter values were observed at a baseline (0 weeks), and at weeks 2, 4, 8, and 12. Although many lab parameters were measured, only hepatic parameters were studied for this example. The hepatic group included alkaline phosphatase, total bilirubin, CPK, LDH, SGOT, and SGPT. The parameters SGOT, SGPT, and LDH were determined clinically to be of more importance than the other parameters and were arbitrarily weighted 3 to 1 to the other hepatic parameters.

Table 1 shows the GS and normalized deviation D_i for patient 1 for hepatic parameters at each week of observation and Figure 1 shows the values of GS and D_i graphically. For patient 1, GS peaks at 2 weeks with a value substantially higher than the other GS's, and then returns to near baseline. Inspection of the D_i for week two shows SGOT with a value of 6.2 (620% increase from upper limit) and SGPT with a value of 2.53 (253% increase from upper limit); this indicates that these two parameters contributed the most to the week two GS value. The D_i for the remaining weeks for SGPT sharply decrease and remain less than 1.0 (100% increase); SGOT, on the other hand, decreases then begins to increase again. Also, total bilirubin (BILT) is steadily increasing over the weeks, going from 1.0 (100% increase from upper limit) to 2.67 (267% increase). The remaining parameters also have non-zero D_i with the exception of CPK; the three most important parameters all have non-zero D_i. Therefore, for patient 1, there may be concern with respect to hepatic safety and further

evaluations may need to be made to determine if the possible abnormalities were due to drug received.

Table 2 and Figure 2 show the change from baseline for the values of GS and D_i over time for patient 1. The interpretation of these results are the same as for Table 1 and Figure 1.

Figure 3 shows a bar chart of the individual lab parameters' actual normalized deviations (before weighting) and the GS for each week. Figure 4 is a bar chart showing the weighted D_i for the lab parameters as their proportion of GS, which in turn illustrates which lab parameters most influence GS for each week. Both figures give similar information, and both show that SGOT had the greatest influence on the GS.

Discussion

Lab parameter data plays an important part in evaluating the safety of a new drug. Simplification of the display of the lab data is much desired, and the method presented in the paper by Sogliero-Gilbert, et. al. seems very promising. The individual patient tables, plots, and charts showing GS and D_i provide a concise summary of the deviations from the normal range.

This procedure should not be the only method used in the analysis of lab data, but it is a good first approach at answering the safety question of a new drug and at finding individual patients with serious abnormalities. It provides a preliminary summary for the clinician inspecting the data, and enables the clinician to ignore patients with no abnormalities (GS=0).

SAS®, and SAS/GRAPH® are registered trademarks of SAS Institute, Inc., Cary, NC USA.

References

1. Sogliero-Gilbert, Gene, Ken Mosher, and Lonni Zubkoff, 1986. A Procedure For The Simplification and Assessment of Lab Parameters in Clinical Trials. Drug Information Journal, Vol. 20, pp. 279-296.

Note:
To obtain copies of the macro please contact the author at the address listed below:

Kathleen H. Arthur
Department of Biometrics
Sterling-Winthrop Research Institute
81 Columbia Turnpike
Rensselaer, New York 12144
(518)445-8781

TABLE 1

GENIE SCORES AND NORMALIZED DEVIATIONS (D)

GROUP=HEPATIC DRUG =A PATIENT ID =1

		LAB PARAMETER									
		ALKALINE PHOS.		TOTAL BILIRUBIN		CPK		LDH		SGOT		SGPT	
TIME (WEEKS)	GENIE SCORE	NORMALIZED DEVIATION	D	NORMALIZED DEVIATION	D	NORMALIZED DEVIATION	D	NORMALIZED DEVIATION	D	NORMALIZED DEVIATION	D	NORMALIZED DEVIATION	D
0	0.2263	0.4918		1.0000		0.0000		0.2793		1.5750		0.3929	
2	0.7861	0.1391		1.3333		0.0000		0.3067		6.2000		2.5333	
4	0.1065	0.3652		1.6667		0.0000		0.2989		0.1500		0.1778	
8	0.3544	0.2951		2.6667		0.0000		0.4376		2.5000		0.3750	
12	0.3544	0.2951		2.6667		0.0000		0.4376		2.5000		0.3750	

TABLE 2

CHANGE FROM BASELINE FOR GENIE SCORE AND NORMALIZED DEVIATION (D)

GROUP=HEPATIC DRUG =A PATIENT ID =1

		LAB PARAMETER							
		ALKALINE PHOS.	TOTAL BILIRUBIN	CPK	LDH	SGOT	SGPT
TIME (WEEKS)	GENIE SCORE	CHANGE FROM BASELINE	CHANGE FROM BASELINE	CHANGE FROM BASELINE	CHANGE FROM BASELINE	CHANGE FROM BASELINE	CHANGE FROM BASELINE
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.5599	-0.3527	0.3333	0.0000	0.0273	4.6250	2.1405
4	-0.1197	-0.1266	0.6667	0.0000	0.0096	-1.4250	-0.2151
8	0.1282	-0.1967	1.6667	0.0000	0.1583	0.9250	-0.0179
12	0.1282	-0.1967	1.6667	0.0000	0.1583	0.9250	-0.0179

FIGURE 1

GENIE SCORE AND NORMALIZED VALUES VS. TIME (WEEKS)

GROUP=HEPATIC DRUG =A PATIENT ID =1

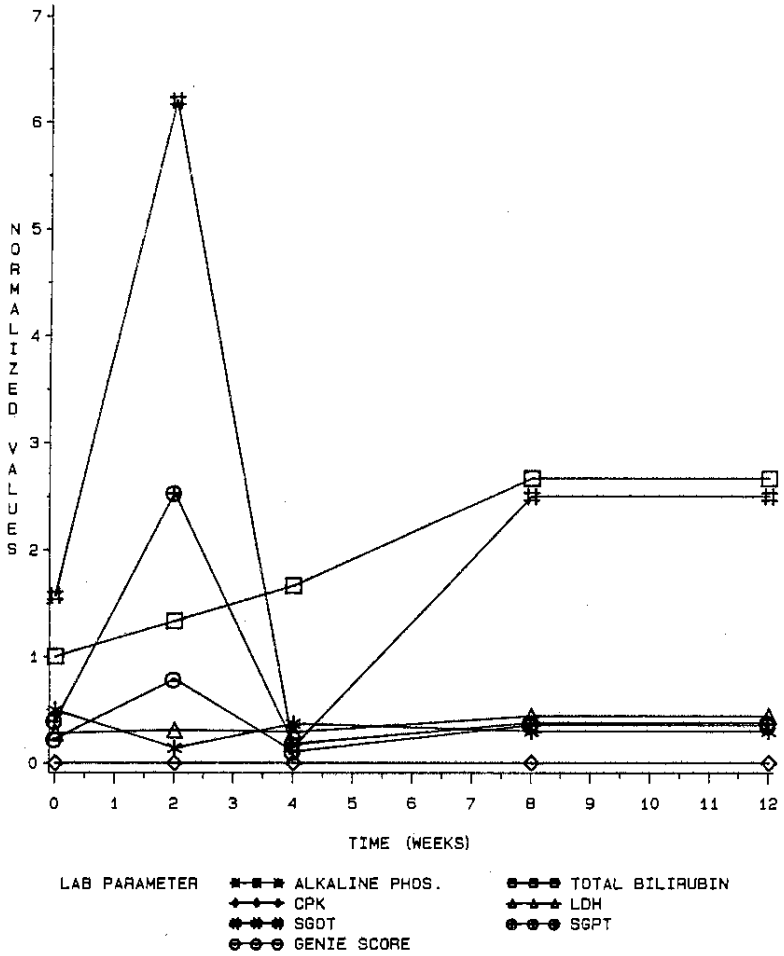

FIGURE 2

CHANGE FROM BASELINE FOR GENIE SCORE AND NORMALIZEDVALUES VS. TIME (WEEKS)

GROUP=HEPATIC DRUG =A PATIENT ID =1

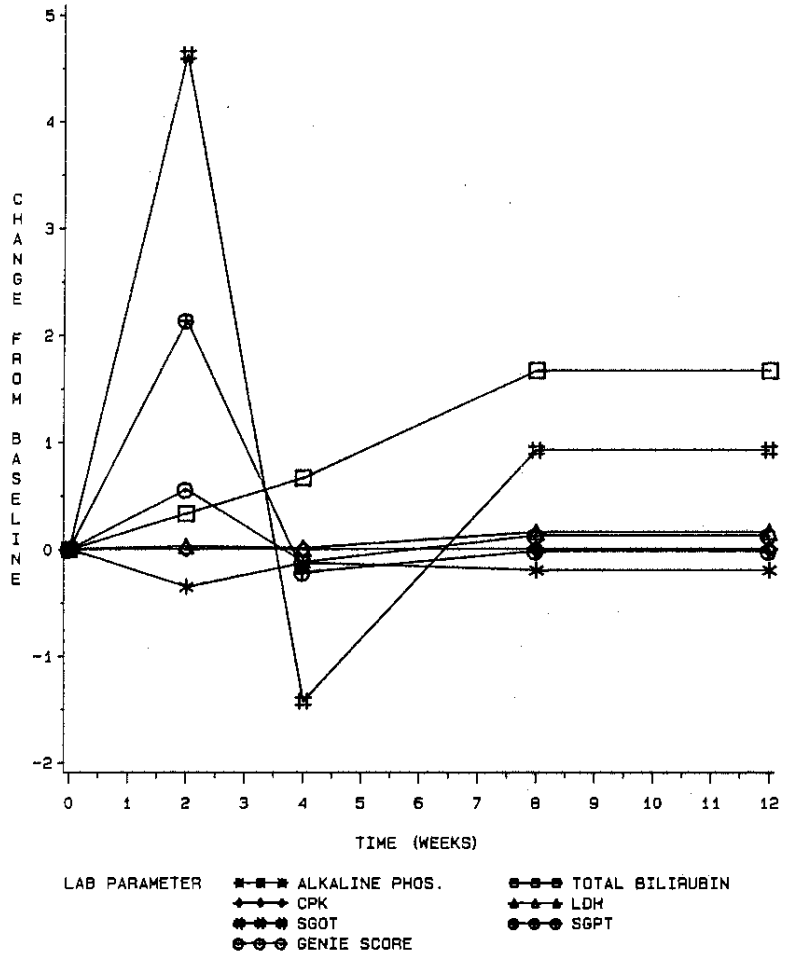

FIGURE 3
GENIE SCORE AND NORMALIZED VALUES OVER TIME (WEEKS)
GROUP=HEPATIC DRUG =A PATIENT ID =1

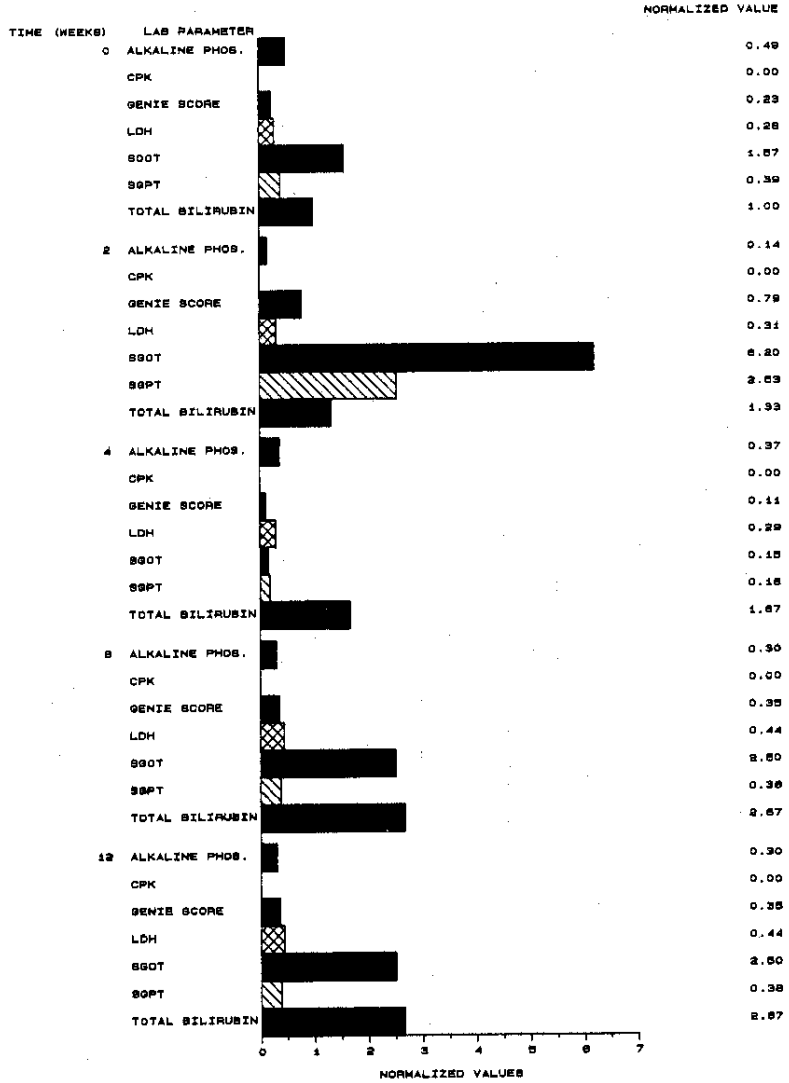

FIGURE 4
GENIE SCORE BROKEN DOWN BY NORMALIZED VALUES OVER TIME (WEEKS)
GROUP=HEPATIC DRUG =A PATIENT ID =1