A SAS® MACRO BASED SYSTEM FOR THE PREPARATION OF DATA FOR COMPARATIVE BIOAVAILABILITY ANALYSIS

Kelly K. Antony
Laurie T. Estun
Merrell Dow Research Institute

ABSTRACT

This paper presents an application of the SAS® system for the preparation of data in comparative bioavailability studies. These studies compare the rate and extent of absorption of a drug between two or more treatments. Examples of comparisons typically made include tablet vs. liquid, fed vs. fasted, and generic vs. trade. Three macros are described, and the end result of their execution is a permanent SAS data set consisting of key pharmacokinetic parameters such as area under the concentration-time curve, peak concentration, time to peak concentration and half-life.

INTRODUCTION

Comparative bioavailability studies typically utilize small numbers of subjects (less than 30) who are given a certain drug for a short period of time. Plasma and/or urine is collected over many time intervals and assayed for the concentration level of the drug. Crossover designs are usually employed in order to reduce between subject variability. Due to the similarity of the data and design of bioavailability studies, macros were created to reduce the amount of repetitive programming required to create a permanent SAS data set containing the necessary pharmacokinetic (PK) parameters. The macro system consists of 4 programs: the driver program which assigns values characteristic to the particular study to macro variables. Examples are titles, compound, and type of fluid being analyzed. An example of the driver program follows:

1. Organize raw data
2. Run driver to execute %SUBPLOT which creates the semi-log plots
3. Edit %HALFLIFE
4. Run driver to execute %PKCREATE thus creating the permanent data set

ORGANIZATION OF RAW DATA

The first step in the process is to manipulate the data into the format required to run the macros. This SAS data set must consist of subject (SUBJECT), treatment (TREAT), time of sample collection (TIME), and plasma or urine concentration (CONC) variables. Sequence (SEQ) and period (PERIOD) variables should also be present if the study is a crossover design.
The primary purpose of the semi-log plots is to determine the log-linear terminal elimination phase for the calculation of half-life (T_{1/2}), the time required for half of the amount of the drug introduced into the body to be eliminated. They are also used to set the interval for the calculation of area under the curve (AUC). The macro produces one plot for each subject treatment combination using PROC GPLOT. The lower limit of measurement (LLM) is denoted by a horizontal line on each plot and is determined by the sensitivity of the assay technique. Many times concentrations falling below this limit are not used in determining the time intervals for T_{1/2} and AUC. Each plot includes a regression line based on the last 3 concentrations which assayed above the LLM. This regression line extends from the time of peak concentration to the last point above the LLM. To base the regression line on more concentrations, specify the number in the driver program %SUBPLOT macro call (Num=). The regression line is created from output generated by the PROC REG procedure using predicted concentrations. The raw concentrations and predicted regression line are overlayed on each subject treatment plot. This facilitates the identification of the log-linear terminal elimination phase. The macro and an example of a resultant plot are shown below.

```plaintext
%SUBPLOT
%options driver=wp7350a nocaracters nocells vpos=100 bpos=100

create individual subject treatment semi-log plots

proc sort data=hhh out=hhh; by subject treat time;
proc plot data=hhh out=hhh; by subject treat;
```

```plaintext
DATA PRED; SET LASTN; BY SUBJECT TREAT;
OUTPUT;
IF LAST.TREAT THEN DO;
TIME = LAST.TREAT - LOGCONC;
OUTPUT;
END;
PROC SORT; BY SUBJECT TREAT;
PROC GPLOT DATA=PRED; BY SUBJECT TREAT;
MODEL LOGCONC=TIME;
OUTPUT OUT=REG PREDCON;
PROC SORT DATA=REG; BY SUBJECT TREAT TIME;
PROC SORT DATA=ALL; BY SUBJECT TREAT;
DATA ALL; MERGE ALL REG; BY SUBJECT TREAT;
PROC GPLOT DATA=REG; BY SUBJECT TREAT TIME;
LABEL=LOGCONC * TIME ;;
```

```plaintext
AXIS1
LABEL=(C..BLACK F..SIMPLEX J..JOIN L..19)
SYMBOL1 COLOR=BLACK P..SPECIAL H..3.5 V..JOIN L..11
SYMBOL2 COLOR=RED P..SPECIAL H..4.1 V..JOIN L..41
RUN;
```

```plaintext
END OF MACRO

END SUBPLOT;
```
HALF-LIFE PROGRAM

After the terminal elimination phase has been identified for each subject treatment combination, the half-life program must be edited. This program is included in the same program library as the driver program since it will be modified for each study. In this program, a data set is created consisting of subject, treatment and the lower and upper time points identified using the semi-log plots. If the lower and upper time points are the same for every subject treatment combination, then a simple DO LOOP will create the half-life data set. However, if the time intervals are different, an input statement will have to be used to enter the data. The result of this program is a sub macro (%HALFLIFE) which is called by %PKCREATE.

**EMACED BALFLIFE:**

---

**SUB MACRO : CALLED BY PKCREATE**

---

**DATA BALFLIFE:**

---

**SUB MACRO : CALLED BY PKCREATE**

---

**PROC PRINT;**

---

**END.**

---

**END MACRO : BALFLIFE**

---

**END MACRO : BALFLIFE**

---
%PKCREATE

%PKCREATE will create a permanent data set consisting of subject, treatment, AUC, peak concentration (Cmax), time to peak concentration (Tmax), and half-life and period for a crossover study. %HALFLIFE is called and the half-life data set is merged with the raw data. Data outside the lower and upper time points are deleted and two criteria are tested. To insure accuracy and precision, at least 3 data points in the log-linear phase must be above the LLM. Also, there should be sufficient spread in the concentration values used to calculate half-life. Therefore, the ratio of the initial to final concentration value in the log-linear phase must be greater than or equal to 1.4. This figure is somewhat arbitrary, however it insures that half-life is measured over the time span corresponding to at least 50% of its value. For failures of these criteria, a half-life will not be calculated. PROC REG is executed to determine the slope of the log-linear terminal elimination phase. An output data set containing these parameter estimates and precision, at least 3 data points over the time span corresponding to at least 50% of its value is created. The absolute value of the slope is the elimination rate constant (KE) and Thalf is calculated from this estimate. The next data set calculates other PK variables based on the raw data such as AUC, Cmax, and Tmax. AUC is calculated using the linear trapezoidal rule. The last data set combines the half-life data set and the resulting PK data set. Two additional PK parameters are calculated here, area under the curve to infinity (AUCINF) and area under the curve adjusted by the rate constant (AUCKE). A permanent data set is created and ready for analysis. The macro follows:

**MACRO PKCREATE (PK_TIME):**

```
***BEGINNING OF MACRO***

HALFLIFE

DATA RAW; SET MCH.PKRAW;
PROC SORT; BY SUBJECT TREATMENT TIME;

MERGE RAW DATA WITH HALFLIFE DATA SET, KEEP ONLY TIME INTERVAL FOR CALCULATION OF HALF-LIFE

DATA HALF; MERGE RAW HALFLIFE; BY SUBJECT TREATMENT;
IF TIME OR LAG (AND THE TIME IS UNBALANCED) OR CONC GT 0 THEN TSP (LOGCONC = LOG (CONC));
LABEL LOGCONC = 'LN OF CONCENTRATION';

TEST CRITERIA: 1) MUST BE AT LEAST 3 DATA POINTS ABOVE LLM
2) RATIO OF INITIAL TO FINAL CONC MUST BE >= 1.4
DELETE FAILURES FROM HALF-LIFE CALCULATION.

DATA CRITIA; SET HALF; BY SUBJECT TREAT;
KEEP SUBJECT TREATMENT;
IF FIRST.TREAT THEN DO;
(SOME CONC = 0);
END;
END;
ELSE;
IF TIME < (UP坞CONC = CONC LT 1.4) OR (LOGLT 3) THEN DO;
DEL = 1;
OUTPUT;
END;
DATA ALLHALF; MERGE HALF CRITIA; BY SUBJECT TREAT;
IF DEL = 1 THEN DELETE;
PROC SORT; BY TREAT SUBJECT;
```

**DETERMINING SLOPE OF THE TERMINAL ELIMINATION PHASE**

```
PROC REG DATA-ALLHALF OUTEST-ALLHALF; BY TREATMENT SUBJECT;
**-------------------------------------------------------------------**
DATA HALF; SET ALLHALF;
KEEP SUBJECT TREATMENT TIME;
KE=LOGCONC/TIME;
THALF=LOG(2)/KE;
/* CALCULATE HALF-LIFE */
LABEL KE='RATE CONSTANT FOR CONC.';
THALF='HALF LIFE FOR CONC.';
PROC SORT; BY TREATMENT SUBJECT;
**-------------------------------------------------------------------**
DATA PRED TEMP; SET RAW; BY SUBJECT TREATMENT;
KEEP AUC LASTTIME CMAX AUCINF;
IF FIRST.TREAT THEN DO;
CMAX=0;
LASTTIME=0;
AUC=0;
END;
IF CONC = 0 AND AUC = 0 THEN DO;
LASTTIME=0;
AUC=0;
END;
IF CONC > 0 THEN DO;
LASTTIME=0;
AUC=0;
END;
IF CONC < 0 THEN DO;
LASTTIME=0;
AUC=0;
END;
IF CONC = 1 THEN DO;
LASTTIME=0;
AUC=0;
END;
LASTTIME=0;
AUC=0;
END;
OUTPUT TEMP;
IF TIME_LT TIME THEN OUTPUT TEMP;
PROC PRINT DATA=PRED;
*** TEMP CAN BE USED TO CHECK THE PK CALCULATIONS ***
**-------------------------------------------------------------------**
DATA CRITIA; SET ALLHALF;
KEEP AUC INF LASTTIME CMAX AUCINF;
AUC = AUC FOR THE CONCENTRATION OUT TO INFINITY;
AUCINF = AUC FOR THE CONCENTRATION ADJUSTED BY EX;
LABEL AUC = 'CONCENTRATION CURVE TO LT TIME (UNITS BB)';
AUCINF = 'AUC UNDER THE CURVE TO INFINITY';
LASTTIME = 'TIME UNDER THE CURVE ADJUSTED BY EX';
CMAX = 'PEAK CONCENTRATION (UNITS)';
AUCINF = 'SUM OF PEAK RESPONSE';
PROC SORT; BY SUBJECT TREAT;
**-------------------------------------------------------------------**
PROC CONTENTS;
**-------------------------------------------------------------------**
TITLE 'PK PARAMETERS FILE';
TITLE 'PK PARAMETERS FILE';
**-------------------------------------------------------------------**
PROC PRINT;
RUN;
**-------------------------------------------------------------------**
```

```
SUMMARY

The process of preparing data from comparative bioavailability studies for analysis requires the calculation of several pharmacokinetic parameters. This paper illustrates a macro based system to accomplish this while eliminating repetitive programming. Although this system was written for use in the MVS operating environment, it can be easily modified to run on the PC.

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

The authors may be contacted at:

Merrell Dow Research Institute
Department of Biostatistics
2110 E. Galbraith Road
Cincinnati, OH 45215
(513) 948-6008/7521