Paper 212-27

DYNAMIC DATA MATCHING IN CLINICAL TRIAL RESEARCH

Zaizai Lu, Kendle International, Inc.
David Shen, CTG Consulting

ABSTRACT

This paper introduces a macro we created. The purpose of the macro is to find dose-related safety parameters by dynamically matching safety parameters with dose information effectively. The macro enables the user to find adverse events (or other safety parameters) within a certain period of time following each dose. It can also find all adverse events after each dose, if no time frame is specified. The macro is easy to understand and use. The user only needs to supply with a variable for safety parameter time, a dose time variable, and a time window between the safety parameter and the previous dosing, and some other simple variables. Everything else is automatically resolved through the macro.

INTRODUCTION

In clinical trial projects, we often need to combine dose administration dataset with safety datasets, such as vital signs, adverse events, ECG, and lab results. One of the common tasks is to find dose-related safety parameters. For example, a patient is given several doses at certain time points. After each dose, his vital signs are taken, or some adverse events may occur. We need to know which adverse event or vital sign parameter is associated with which dose. In the perfect world, we would hope an adverse event to occur within 3 hours of a dose, or expect vital signs to be measured 2 hours after a dose. But in the real world, most data are irregular. Some data maybe missing, some may contain duplicates. We might have some control over vital sign measurement times, but we do not have any control over adverse event occurrences. There is high unpredictability as to when an adverse event may occur. This unpredictability makes it difficult to find dose-related safety parameters. This paper presents a technique that can dynamically match safety parameters with dose information effectively. One macro program is introduced here. This macro enables the user to find all adverse events (or other safety parameters) within a certain period of time following each dose. If no time frame is defined, then all adverse events after each dose will be presented. This macro is easy to understand and use, and has great applicability in real-life clinical trial projects, especially Phase-I trials.

This paper is composed of 6 parts. Part 1 includes the abstract and introduction. Part 2 describes the types of data file that can be analyzed using this macro. Part 3 presents some data files used in the paper to generate the sample output of the SAS® macro program. Part 4 explains the macro program in more details. Part 5 presents the source code of the macro. Part 6 exhibits some examples of the applications of the macro program and their output.

DATA TYPE

A variety of types of data may be applied with this method. The major types are:
1. Adverse Events
2. Vital Signs
3. ECG
4. Lab Results

DATA PROCESSING

Suppose there is a dose dataset and an adverse event dataset.

Dose Data

OBS	PATNO	DOSEDTTM	DOSEAMT
1	0001	30JAN1999:08:00:01	10
2	0001	06FEB1999:08:00:00	30
3	0001	20FEB1999:08:00:00	90
4	0001	27FEB1999:08:00:02	120
5	0002	30JAN1999:08:01:02	10
6	0002	06FEB1999:08:01:01	30
7	0002	20FEB1999:08:01:00	90
8	0002	06MAR1999:08:01:00	120
9	0003	30JAN1999:08:02:01	10
10	0003	06FEB1999:08:02:01	30
11	0003	13FEB1999:08:02:01	60
12	0003	27FEB1999:08:02:08	90
13	0003	06MAR1999:08:02:01	120

Adverse Event Data

OBS	PATNO	AE	AEDTTM
1	0001	NERVOUSNESS	30JAN1999:06:00:00
2	0001	TACHYCARDIA	30JAN1999:12:15:00
3	0001	NAUSEA	06FEB1999:16:20:00
4	0001	DIZZINESS	20FEB1999:09:20:00
5	0002	HEADACHE	06FEB1999:14:10:00
6	0002	NAUSEA	06FEB1999:17:40:00
7	0002	VASODILATATION	20FEB1999:14:30:00
8	0002	ASTHENIA	20FEB1999:18:20:00
9	0002	ANXIETY	27FEB1999:20:01:00
10	0002	CHILLS	28FEB1999:06:00:00
11	0002	CONSTIPATION	28FEB1999:22:00:00
12	0003	PALLOR	30JAN1999:10:17:00
13	0003	SWEATING	30JAN1999:12:17:00
14	0003	NAUSEA	06FEB1999:15:17:00
15	0003	DIZZINESS	06FEB1999:17:17:00
16	0003	HYPESTHESIA	13FEB1999:09:30:00
17	0003	PAIN	13FEB1999:12:30:00
18	0003	CHILLS	27FEB1999:16:45:00
19	0003	EUPHORIA	06MAR1999:10:15:00
20	0003	DIZZINESS	06MAR1999:18:46:00

Now, we need to create a new dataset to list all adverse events within a certain period of time (for instance, 6 or 12 hours following each dose administration).

EXPLANATION OF THE MACRO

Our macro %match is able to do this job effectively and efficiently. What you need to do is to define the name of the dose data set (dosedata), the name of the safety parameter data set (safedata), the time window between the safety parameter and the previous dosing (hrs). Other macro variables used in this macro include the patient id variable (patno), the variable name of the dose date and time (dosetime), the variable name of the safety parameter date and time (safetime), and the output dataset name (outdata). Both dosetime and safetime are numeric. In theory, patients are supposed to take doses on schedule. However, some patients may fail to follow the schedule and miss several doses. So actually, patients may not have the same number of doses. The first step in this macro is to capture as a macro variable the largest number of doses among all the patients. The output of PROC FREQ contains one observation per PATNO, with the variable COUNT indicating the number of doses each patient receives. Because of the ORDER=FREQ option on the PROC FREQ statement, the variable COUNT is put in descending order, with the first observation having the largest value. Using a DATA *_NULL_* step and selecting _ N _ =1, we are able to create a macro variable to get the maximum number of doses in the whole dose dataset. The next step PROC TRANSPOSE creates a temporary dataset containing one observation for each patient with all the dosing data. After one-to-many merging, the time difference can be conveniently calculated by subtracting safety parameter time from the dosing time. To be easily understood, we can deal with one dose each time through a loop from 1 to the largest dosing number. The resulting data set will be created by comparing the calculated time difference with the defined time window. Each dose will generate a subset of datasets after the process. When the looping is done, all resulting datasets with the expected safety parameter data will then be merged together to generate the final output dataset with the name you wish to call.

SOURCE CODE OF THE MACRO PROGRAM

```
%macro match (dosedata=, safedata=, hrs=,
patno=, dosetime=, safetime=, outdata=);

proc sort data=&dosedata;
  by &patno &dosetime;
run;

proc freq data=&dosedata order=freq noprint;
  table &patno /out=temp1;
run;

data _null_;
  set temp1;
  if _n_=1 then call symput('maxdose',
              count);
run;

proc transpose data=&dosedata out=temp2
(drop=_name_) prefix=dosen;
  by &patno;
  var &dosetime;
run;

%local num;
  %do num=1 %to &maxdose;
     proc sort data=temp2
     out=dose&num(keep=&patno dosen&num);
     by &patno;
     run;

     data ndata&num ;
     merge &safedata(in=a) dose&num(in=b);
     by &patno;
     if a and b;
     dosen=&num;
     hrs=(&safetime - dosen&num)/3600;
     &dosetime=dosen&num;

  %if &hrs ne %then %do;  *ALL DATA WILL BE
COLLECTED FOLLOWING EACH DOSE IF NO HOURS IS
SET;
      if 0=<hrs<=&hrs then output ndata&num;
  %end;
```

```
  %else %if &hrs eq %then %do;  *ALL DATA
WILL BE COLLECTED FOLLOWING EACH DOSE IF NO
HOURS IS SET;

 if 0=<hrs then output ndata&num;
%end;
 drop dosen&num;
run;
%end;

data &outdata;
  set ndata1;
run;

%do num=2 %to &maxdose;
  data &outdata;
  set &outdata ndata&num;
  run;
%end;

proc sort data=&outdata;
  by &patno dosen &safetime;
run;

proc print;
  format &safetime  &dosetime datetime18.;
  title "&outdata";
run;

%mend match;
```

EXAMPLES OF THE MACRO

If we call this macro with different time windows,

```
%match (dosedata=dose, safedata=adverse,
hrs=6, patno=patno, dosetime=dosedttm,
safetime=aedttm, outdata=a);

%match (dosedata=dose, safedata=adverse,
hrs=12,  patno=patno,  dosetime=dosedttm,
safetime=aedttm, outdata=b);
```

the following 2 datasets would be generated.

a) Adverse event dataset with time difference of 6 hours

PATNO	AE	AEDTTM	DOSEDTTM	DOSEN	HRS
0001	TACHYCARDIA	30JAN1999:12:15:00	30JAN1999:08:00:01	1	4.24972
0001	DIZZINESS	20FEB1999:09:20:00	20FEB1999:08:00:00	3	1.33333
0003	PALLOR	30JAN1999:10:17:00	30JAN1999:08:02:01	1	2.24972
0003	SWEATING	30JAN1999:12:17:00	30JAN1999:08:02:01	1	4.24972
0003	HYPESTHESIA	13FEB1999:09:30:00	13FEB1999:08:02:01	3	1.46639
0003	PAIN	13FEB1999:12:30:00	13FEB1999:08:02:01	3	4.46639
0003	EUPHORIA	06MAR1999:10:15:00	06MAR1999:08:02:01	5	2.21639

b) Adverse event dataset with time difference of 12 hours

PATNO	AE	AEDTTM	DOSEDTTM	DOSEN	HRS
0001	TACHYCARDIA	30JAN1999:12:15:00	30JAN1999:08:00:01	1	4.2497
0001	NAUSEA	06FEB1999:16:20:00	06FEB1999:08:00:00	2	8.3333
0001	DIZZINESS	20FEB1999:09:20:00	20FEB1999:08:00:00	3	1.3333
0002	HEADACHE	06FEB1999:14:10:00	06FEB1999:08:01:01	2	6.1497
0002	NAUSEA	06FEB1999:17:40:00	06FEB1999:08:01:01	2	9.6497
0002	ASTHENIA	20FEB1999:18:20:00	20FEB1999:08:01:00	3	10.3167
0003	PALLOR	30JAN1999:10:17:00	30JAN1999:08:02:01	1	2.2497
0003	SWEATING	30JAN1999:12:17:00	30JAN1999:08:02:01	1	4.2497
0003	NAUSEA	06FEB1999:15:17:00	06FEB1999:08:02:01	2	7.2497
0003	DIZZINESS	06FEB1999:17:17:00	06FEB1999:08:02:01	2	9.2497
0003	HYPESTHESIA	13FEB1999:09:30:00	13FEB1999:08:02:01	3	1.4664
0003	PAIN	13FEB1999:12:30:00	13FEB1999:08:02:01	3	4.4664
0003	CHILLS	27FEB1999:16:45:00	27FEB1999:08:02:08	4	8.7144
0003	EUPHORIA	06MAR1999:10:15:00	06MAR1999:08:02:01	5	2.2164
0003	DIZZINESS	06MAR1999:18:46:00	06MAR1999:08:02:01	5	10.7331

If we leave the time window missing, then the dataset will look like this, with all the adverse events following each dose:

```
%match(dosedata=dose,safedata=adverse,
hrs=,  patno=patno,  dosetime=dosedttm,
safetime=aedttm, outdata=c);
```

c) Adverse event dataset with no time difference defined

PATNO	AE	AEDTTM	DOSEN	HRS	DOSEDTTM
0001	TACHYCARDIA	30JAN99:12:15:00	1	4.250	30JAN99:08:00:01
0001	NAUSEA	06FEB99:16:20:00	1	176.333	30JAN99:08:00:01
0001	DIZZINESS	20FEB99:09:20:00	1	505.333	30JAN99:08:00:01
0001	NAUSEA	06FEB99:16:20:00	2	8.333	06FEB99:08:00:00
0001	DIZZINESS	20FEB99:09:20:00	2	337.333	06FEB99:08:00:00
0001	DIZZINESS	20FEB99:09:20:00	3	1.333	20FEB99:08:00:00
0002	HEADACHE	06FEB99:14:10:00	1	174.149	30JAN99:08:01:02
0002	NAUSEA	06FEB99:17:40:00	1	177.649	30JAN99:08:01:02
0002	VASODILATATION	20FEB99:14:30:00	1	510.483	30JAN99:08:01:02
0002	ASTHENIA	20FEB99:18:20:00	1	514.316	30JAN99:08:01:02
0002	ANXIETY	27FEB99:20:01:00	1	683.999	30JAN99:08:01:02
0002	CHILLS	28FEB99:06:00:00	1	693.983	30JAN99:08:01:02
0002	CONSTIPATION	28FEB99:22:00:00	1	709.983	30JAN99:08:01:02
0002	HEADACHE	06FEB99:14:10:00	2	6.150	06FEB99:08:01:01
0002	NAUSEA	06FEB99:17:40:00	2	9.650	06FEB99:08:01:01
0002	VASODILATATION	20FEB99:14:30:00	2	342.483	06FEB99:08:01:01
0002	ASTHENIA	20FEB99:18:20:00	2	346.316	06FEB99:08:01:01
0002	ANXIETY	27FEB99:20:01:00	2	516.000	06FEB99:08:01:01
0002	CHILLS	28FEB99:06:00:00	2	525.983	06FEB99:08:01:01
0002	CONSTIPATION	28FEB99:22:00:00	2	541.983	06FEB99:08:01:01
0002	VASODILATATION	20FEB99:14:30:00	3	6.483	20FEB99:08:01:00
0002	ASTHENIA	20FEB99:18:20:00	3	10.317	20FEB99:08:01:00
0002	ANXIETY	27FEB99:20:01:00	3	180.000	20FEB99:08:01:00
0002	CHILLS	28FEB99:06:00:00	3	189.983	20FEB99:08:01:00
0002	CONSTIPATION	28FEB99:22:00:00	3	205.983	20FEB99:08:01:00
0003	PALLOR	30JAN99:10:17:00	1	2.250	30JAN99:08:02:01
0003	SWEATING	30JAN99:12:17:00	1	4.250	30JAN99:08:02:01
0003	NAUSEA	06FEB99:15:17:00	1	175.250	30JAN99:08:02:01
0003	DIZZINESS	06FEB99:17:17:00	1	177.250	30JAN99:08:02:01
0003	HYPESTHESIA	13FEB99:09:30:00	1	337.466	30JAN99:08:02:01
0003	PAIN	13FEB99:12:30:00	1	340.466	30JAN99:08:02:01
0003	CHILLS	27FEB99:16:45:00	1	680.716	30JAN99:08:02:01
0003	EUPHORIA	06MAR99:10:15:00	1	842.216	30JAN99:08:02:01
0003	DIZZINESS	06MAR99:18:46:00	1	850.733	30JAN99:08:02:01
0003	NAUSEA	06FEB99:15:17:00	2	7.250	06FEB99:08:02:01
0003	DIZZINESS	06FEB99:17:17:00	2	9.250	06FEB99:08:02:01
0003	HYPESTHESIA	13FEB99:09:30:00	2	169.466	06FEB99:08:02:01
0003	PAIN	13FEB99:12:30:00	2	172.466	06FEB99:08:02:01
0003	CHILLS	27FEB99:16:45:00	2	512.716	06FEB99:08:02:01
0003	EUPHORIA	06MAR99:10:15:00	2	674.216	06FEB99:08:02:01
0003	DIZZINESS	06MAR99:18:46:00	2	682.733	06FEB99:08:02:01
0003	HYPESTHESIA	13FEB99:09:30:00	3	1.466	13FEB99:08:02:01
0003	PAIN	13FEB99:12:30:00	3	4.466	13FEB99:08:02:01
0003	CHILLS	27FEB99:16:45:00	3	344.716	13FEB99:08:02:01
0003	EUPHORIA	06MAR99:10:15:00	3	506.216	13FEB99:08:02:01
0003	DIZZINESS	06MAR99:18:46:00	3	514.733	13FEB99:08:02:01
0003	CHILLS	27FEB99:16:45:00	4	8.714	27FEB99:08:02:08
0003	EUPHORIA	06MAR99:10:15:00	4	170.214	27FEB99:08:02:08
0003	DIZZINESS	06MAR99:18:46:00	4	178.731	27FEB99:08:02:08
0003	EUPHORIA	06MAR99:10:15:00	5	2.216	06MAR99:08:02:01
0003	DIZZINESS	06MAR99:18:46:00	5	10.733	06MAR99:08:02:01

REFERENCES

SAS Institute Inc. (1990) SAS/STAT® User's Guide, Version 6, Fourth Edition, Cary, NC: SAS Institute Inc.

SAS Institute Inc. (1990) SAS® Guide to Macro Processing, Version 6, Fourth Edition, Cary, NC: SAS Institute Inc.

TRADEMARKS

CONTACT INFORMATION

For additional information, contact the author at:

Zaizai Lu
Kendle International, Inc.
1200 Carew Tower
441 Vine Street
Cincinnati, OH 45202
lu.zaizai@kendle.com