

## The Supplement of the COMPARE Procedure

WeiQin Pan, Covance Inc., Princeton, NJ

### ABSTRACT

The SAS<sup>®</sup> COMPARE<sup>2</sup> procedure is very useful to help us identify discrepancies between the two data sets. It will show us disagreements in data set attributes, matching variables, attributes of matching variables and observations. It also compares the value of observations as well as variables; however, when character values are compared only the first 12 characters are displayed. This paper will demonstrate how an output data set and a numeric return code stored in the automatic macro variable &SYSINFO from this procedure can be used to generate a report for each pair of matching observations judged unequal at one or more variables. The report presents character variables of any length as well as numeric variables, but ignores values for any DIF or %DIF which are included in the table produced by PROC COMPARE.

### INTRODUCTION

If we use PROC COMPARE to examine two data sets, ADR1 and ADR2, (see List 1 for details):

```
proc compare base = mydir.adr1
             Compare = mydir.adr3
             transpose
             nosummary ;
             id prot inv pat bodytext preterm ;
run;
```

We will get the very helpful output shown in List 2. Notice however that the character variable whose length is more than 12 appears incomplete. If the discrepancy happens after the first 12 characters, we will not be able to easily identify the discrepancy between the two data sets. This paper introduces a SAS<sup>®</sup> macro MCOMPDS that supplements the PROC COMPARE procedure.

### ASSUMPTION

To compare them in terms of observations, the two data sets should have matching key variables, and they should appear in the ID statement. The variables to be compared should have the same name and the same type. If they are named differently, you can rename them when the parameters pass into the macro MCOMPDS.

### PARAMETERS

Macro positional parameters are:

```
DS1      - name of the base data set
DS2      - name of the comparing data set
KEYS     - key variable(s) used to match observations in the two
           data sets
```

```
NCOLWID - numeric variable column width for the final report
           (default = 30)
```

```
NTOTCOL - total column of numeric variables per line for the final
           report (default = 4)
```

Both DS1 and DS2 can be work data sets or permanent data sets with libname. For example:

```
libname mydir "/net-users/wpan/";
libname projdir "/proj/001/s001/";
%mCompDS ( DS1 = mydir.medhx,
           DS2 = projdir.medhx,
           KEYS = inv pat bodysysm n );
```

You can use any data set options like KEEP=, DROP=, RENAME= and WHERE= with the parameters DS1 or DS2. For example:

```
%mCompDS(
  DS1 = mydir.medhx( drop = rxdate ),
  DS2 = projdir.medhx( rename =( bodysys = bodysysm) ),
  KEYS = inv pat bodysysm n );
```

### PROCESSING STEPS

1. Sort the two data sets by the key variables, and then create corresponding data sets with only the key variables and a new variable to which the observation positions are assigned. This new variable will be used later to locate specific observations:

```
%do i = 1 %to 2;
  proc sort data = &&ds&i out = ds&i;
           by &keys;
  run;
  data dsB&i ( keep = &keys _obs&i );
           set ds&i;
           by &keys;
           _obs&i=_n_;
  run;
%end;
```

2. Use PROC COMPARE with OUTNOEQUAL option to suppress equal observations:

```
proc compare base = DS1
             compare = DS2
             out = compA
                 ( drop = _TYPE_ _OBS_ )
             noprint
             outnoequal ;
             id &keys;
run;
```

3. Assign the automatic macro variable &SYSINFO from PROC COMPARE to a new macro variable before the DATA step that uses bit-testing features to check for the existence of the discrepancies between the matching observations:

```
%let rc = &sysinfo;
data _null_;
  if &rc = '1.....'b then call symput( "notEQ", "1" );
  else call symput( "notEQ", "0" );
run;
```

4. If any pair of matching observations is judged unequal for one or more variables:

```
%if &notEQ = 1 %then %do;
  :: (see below for the details)
%end;
```

5. Call the macro WORDS to separate the macro variable &KEYS into individual words:

```
%word( string = &KEYS, root = keyvar, numw = keyn );

%macro words ( delim = %str( ),
             numw = ,
             root = ,
             string = )3;
  %local word count;
  %let count = 1;
  %let word = %qscan( &string, &count, &delim );
  %do %while( &word ^= );
    %global &root&count;
    %let &root&count = %unquote( &word );
    %let count = %eval( &count+1 );
    %let word = %qscan( &string, &count, &delim );
  %end;
  %if %length( &numw ) > 0 %then %do;
    %global &numw;
    %let &numw = %eval( &count-1 );
  %end;
%mend words;
```

6. Use PROC CONTENTS to get the list of variables in the output data set (COMP A) and sort it in the order of variable types (numeric = 1 and character = 2.) In DATA \_NULL\_, assign the number of numeric variables to the macro variable &DIM1 and the number of character variables to the macro variable &DIM2. This step also assigns the variable names to macro variables &NUMn or &CHRn, depending on numeric or character variables and the sequence of the variables in the same type. The two character variables NUMKEY and CHRKEY are used to indicate the position(s) of numeric key variables and character key variables. They will look like "1,3,8".

```
proc contents data = compA noprint
    out = cont( keep = type varnum name );
run;

proc sort data = cont
    out = cont( rename =( type = _type ));
    by type varnum;
run;

%let type1 = 0;
%let type2 = 0;
%let dim1 = 0;
%let dim2 = 0;
%let numkey = 0;
%let chrkey = 0;
```

```
data _null_;
    set cont ;
    by _type varnum;

    length cVarN $ 2 ;

    array key[2] $60 NumKey ChrKey ;
    array V[2] $3 ( "Num" "Chr" );

    retain varN key;

    if first_&type then do;
        varN = 0;
        key[_type] = "0";
    end;

    varN+1;
    cVarN = trim( left( put( varN, 2. )));
    call symput( V[_type] || "V" || cVarN, name );

    if trim( left( name )) in ( %do keyi = 1 %to &keyN;
        "&keyVar&keyi"
        %end; ) then
        substr( key[_type], length( key[_type] ) + 1 ) = " " || cVarN ;

    if last_&type & varN > 0 then do;
        call symput ( "dim" || trim( left( put( _type, 1. ))), cVarN);
        call symput ( "type" || trim( left( put( _type, 1. ))), '1' );
        call symput ( V[_type] || "Key", trim( left( key[_type] )));
    end;
run;
```

7. You may already find the output data set from PROC COMPARE provides the result in the special way: .E for an equal value of a non-key numeric variable and "X" for each unequal character of a non-key character variable. The key variables still keep their own values:

| OBS | PROT | INV  | PAT  | BODYTEXT                | PREFTERM            |
|-----|------|------|------|-------------------------|---------------------|
| 1   | 101A | 9606 | 618  | CARDIOVASCULAR SYSTEM   | IRRITATION VENT     |
| 2   | 101A | 9626 | 1107 | CARDIOVASCULAR SYSTEM   | TACHYCARDIA VENT    |
| 3   | 102B | 9405 | 113  | BODY AS A WHOLE         | HEADACHE            |
| 4   | 102B | 9528 | 301  | SKIN & APPENDAGES       | APPLICAT SITE REACT |
| 5   | 201A | 9229 | 807  | CARDIOVASCULAR SYSTEM   | IRRITATION VENT     |
| 6   | 201A | 9229 | 814  | CARDIOVASCULAR SYSTEM   | HEM                 |
| 7   | 201A | 9531 | 1009 | METABOLIC & NUTRITIONAL | SYSTEM SGPT IN      |

OBS VERBATIM

```
1 .....
2 .....
3 .....
4 .....
5 .....
6 .....
7 .....
```

| OBS | CONT | SERIOUS | CAUSAL | CLINSIG | OUTCOME | RELDAY | DEGREE |
|-----|------|---------|--------|---------|---------|--------|--------|
| 1   | .    | E       | .      | E       | E       | E      | E      |
| 2   | .    | E       | E      | E       | E       | .      | E      |
| 3   | X    | .       | .      | .       | .       | E      | .      |
| 4   | X    | E       | E      | E       | E       | E      | E      |
| 5   | .    | E       | E      | E       | E       | E      | .      |
| 6   | .    | E       | E      | E       | E       | E      | E      |
| 7   | .    | E       | E      | E       | E       | E      | E      |

| OBS | RELATE | LABFLAG | D_TIME        | ONDRUG | AGEGROUP | ALLERGY |
|-----|--------|---------|---------------|--------|----------|---------|
| 1   | E      | E       | .....         | E      | E        | E       |
| 2   | E      | E       | .....         | E      | .        | .       |
| 3   | .      | E       | XXXXXXXXX.... | .      | E        | E       |
| 4   | E      | E       | XXXXXXXXX.... | E      | E        | .       |
| 5   | .      | E       | XXXXXXXXX.... | .      | E        | E       |
| 6   | E      | E       | XXXXXXXXX.... | E      | E        | E       |
| 7   | .      | .       | .....         | .      | E        | E       |

8. Manipulate the output data set COMP A and create a new work data set with the key variables, a variable to indicate the number of lines the observation needs for the final report, and four new variables \_NUM, NDIFF, \_CHA and CDIFF. NDIFF and CDIFF are numeric variables that indicate the number of matching variables with unequal values for each type. \_NUM and \_CHA are character variables with the length of &DIM1 and &DIM2 respectively. These two variables have the values "0" or "1" which are used to indicate whether the original matching variable dose or dose not have equal values. The position of the character "0" or "1" within the variable \_NUM or \_CHAR is the same sequence of that original matching variable in the same type:

```
data varlist ( keep = &keys %if &type1 = 1 %then %do;
    _num ndiff
    %end;
    %if &type2 = 1 %then %do;
    _cha cdiff
    %end;
    nLine
);
set compA end = eof;

%if &type1 = 1 %then %do;
    array num _NUMERIC_;
%end;
%if &type2 = 1 %then %do;
    array chara _CHARACTER_;
%end;

length %if &type1 = 1 %then %do;
    _num $&dim1
%end;
%if &type2 = 1 %then %do;
    _cha $&dim2
%end;
;

Nline = 0;
```

```
%if &type1 = 1 %then %do;
  nDiff = 0;
  do over num;
    if num = .E then
      substr( _num, _i_, 1 ) = "0";
    else do;
      if _i_ in ( &NumKey ) then
        substr( _num, _i_, 1 ) = "0";
      else do;
        substr( _num, _i_, 1 ) = "1";
        nDiff + 1;
      end;
    end;
  end;
  nLine=ceil( nDiff/4 );
%end;

%if &type2 = 1 %then %do;
  cDiff=0;
  do over chara;
    if index( chara, 'X' ) = 0 then
      substr( _cha, _i_, 1 ) = "0";
    else do;
      if _i_ in ( &ChrKey ) then
        substr( _cha, _i_, 1 ) = "0";
      else do;
        substr( _cha, _i_, 1 ) = "1";
        cDiff + 1;
      end;
    end;
  end;
  nLine+cDiff;
%end;
run;
```

9. Merge the comparing data sets with the new work data set mentioned above by the key variables. Keep only the observations that found some discrepancies in the output data set COMPA:

```
proc sort data = varlist;
  by &keys;
run;
data varlist;
  merge dsB1 dsB2 varlist ( in = diff );
  by &keys;
  if diff;
run;
```

10. Use the POINT = option on the SET statement to point to the discrepant observations. Check the numeric variables with the variable \_NUM to pick up the discrepancies and do the same for the character variables with \_CHA. Write out the result to the report by using DATA \_NULL\_ and PUT statements. Each observation will have the key variable(s) and any variables with unequal values (see List 3 for the final report.)

```
data _null_;
  set varlist end = eof;

  retain pn 0 lmargin 12;

  file print ps=65 line=l ll=left col=c n=pagesize notitles;

  if l < 3 then link head;
```

```
%do i = 1 %to 2;
  set ds&i point = _OBS&i;

  if &i = 1 then do;
    if lleft < ( nLine*3 + 4 ) then do;
      put _page_;
      link head;
    end;

    put / @1 "##" @;
    %do keyi = 1 %to &keyN;
      put " &&keyVar&keyi = " &&keyVar&keyi @;
    %end;
    put " ##"
      // " * BASE * " @;
    firstL=;
  end;
  else if &i = 2 then put #(firstL+1) " * COMP * " @;
  %if &type1 = 1 %then %do;
    if ndiff > 0 then do;
      colN = lmargin;
      %do j = 1 %to &dim1;
        if substr( _num, &j, 1 ) = '1' then do;
          if colN > ( lmargin + &ncolwid *( &ntotcol - 1 ) )
            then do;
              put //;
              colN = lmargin;
            end;
          put @colN " &&NumV&j = " &&NumV&j @;

          /* &numcol = Numeric variable column width */

          colN + &ncolwid;
        end;
      %end;
      put //;
    end;
  %end;

  %if &type2 = 1 %then %do;
    if cDiff > 0 then do;
      %do j = 1 %to &dim2;
        if substr( _cha, &j, 1 ) = '1' then
          put @lmargin " &&ChrV&j = " &&ChrV&j // ;
      %end;
    end;
  %end;
%end;

if l > 62 then put _page_;
return;

head:
  pn+1;

  put #10 @1 "&datetime"
    @30 "Comparison of BASE Data Set "
      "(&DS1) with COMP Data Set "
      "(&DS2)"
    @120 "Page" pn 3.

  //;

  return;
run;
```

### CONCLUSION

The macro MCOMPDS is useful for comparing two data sets with long character strings. The output default is set for landscape printing, with four columns for numeric variables and one column for a character variable. The macro can be easily modified to meet the needs of the user.

LIST 1

Data Set MYDIR.ADR1

```

P I B P V
R N P O R E
O V A D E R
T E T Y F B
O S I T T A
O C T E E T
B O I N X R I
S L G T T M M

::
28 101A 9606 618 CARDIOVASCULAR SYSTEM IRRITATION VENT VENTRICULAR IRRITATION SECONDARY TO THE CATHETER
::
30 101A 9626 1107 CARDIOVASCULAR SYSTEM TACHYCARDIA VENT VENT TACHYCARDIA W/ INSERT OF PIGTAIL IN LV CEASING WHEN CATHETER MANIPULATED
::
68 102B 9405 113 BODY AS A WHOLE HEADACHE SINUS HEADACHE-PATIENT WITH HISTORY OF SINUS INFECTION
::
87 102B 9528 301 SKIN & APPENDAGES APPLICAT SITE REACT URTICARIC RASH AT AREA FO ELASTOPLAST APPLICATION
::
109 201A 9229 807 CARDIOVASCULAR SYSTEM IRRITATION VENT VENT IRRITATION SECONDARY TO THE CATHETER
::
118 201A 9229 814 CARDIOVASCULAR SYSTEM HEM RT GROIN HEMATOMA ATTEMPT TO CATHETERIZE (PROCEDURE WAS DONE ON LEFT GROIN)
::
130 201A 9531 1009 METABOLIC & NUTRITIONAL SYSTEM SGPT INC INCREASED SGPT
::

```

```

S C O D A
E C L U U R D I D R L A D O G A
R A I T G E D S E E B _ N G L A
C I U N C C L E C G L R F T D R E L
O O O S S O O D A _ R A A L I R O R A
B N U A I M D A T A E T E C A M U U G C
S T S L G E E Y H E E E X E G E G P Y T

::
28 2 . 2 1 3 . 2 2 1 2 1 1 . <1 min 2 4 1 00
::
30 2 6 2 1 3 2 2 2 1 2 2 1 . <1 min 1 . . 00
::
68 . . . . 1 1 2 2 . . 2 1 . 0:00:00 . 4 2 00
::
87 C 2 4 2 1 2 . 2 2 1 2 . 3 . --:--: 1 5 . 01
::
109 2 5 2 1 2 . 2 2 . . 1 1 . --:--: 1 5 2 00
::
118 2 4 1 . 1 . 2 2 1 2 1 1 . 0:00:00 1 5 2 00
::
130 . 3 . 3 1 . 2 2 . . 2 1 . --:--: . 5 2 00
::

```

Data Set MYDIR.ADR2

```

P I B P V
R N P O R E
O V A D E R
T E T Y F B
O S I T T A
O C T E E T
B O I N X R I
S L G T T M M

::
29 101A 9606 618 CARDIOVASCULAR SYSTEM IRRITATION VENT VENTRICULAR IRRITATION SECONDARY TO THE CATHETER
::
31 101A 9626 1107 CARDIOVASCULAR SYSTEM TACHYCARDIA VENT VENT. TACHYCARDIA W/ INSERT. OF FIGTAIL IN LV CEASING WHEN CATHETER MANIPULATED
::
70 102B 9405 113 BODY AS A WHOLE HEADACHE SINUS HEADACHE-PATIENT WITH HISTORY OF SINUS INFECTION
::
89 102B 9528 301 SKIN & APPENDAGES APPLICAT SITE REACT URTICARIC RASH AT AREA OF ELASTOPLAST APPLICATION
::
111 201A 9229 807 CARDIOVASCULAR SYSTEM IRRITATION VENT VENT IRRITATION SECONDARY TO CATH MANIPULATION
::
120 201A 9229 814 CARDIOVASCULAR SYSTEM HEM RT GROIN HEMATOMA ATTEMPT TO CATHETERIZE (PROCEDURE WAS DONE ON LEFT GROIN)
::
131 201A 9531 1009 METABOLIC & NUTRITIONAL SYSTEM SGPT INC INCREASED SGPT
::

```

```

S C O D A
E C L U U R D I D R L A D O G A
R A I T G E D S E E B _ N G L A
C I U N C C L E C G L R F T D R E L
O O O S S O O D A _ R A A L I R O R A
B N U A I M D A T A E T E C A M U U G C
S T S L G E E Y H E E E E G E G P Y T

::
29 2 7 2 1 3 . 2 2 1 2 1 . <1 min 2 4 1 00
::
31 2 6 2 1 3 . 2 2 1 2 1 . <1 min 1 5 2 00
::
70 C 2 4 2 2 1 1 2 2 1 2 1 . --:--: 1 4 2 00
::
89 2 4 2 1 2 . 2 2 1 2 3 . 0:22:45 1 5 2 01
::
111 2 5 2 1 2 . 2 2 1 2 1 . <1 min 2 5 2 00
::
120 2 4 1 . 1 . 2 2 1 2 1 . --:--: 1 5 2 00
::
131 . 3 . 3 1 . 2 2 . 1 1 1 . --:--: 1 5 2 00
::

```

LIST 2

COMPARE Procedure  
 Comparison of MYDIR.ADR1 with MYDIR.ADR2  
 (Method=EXACT)

Comparison Results for Observations

PROT=101A INV=9606 PAT=618 BODYTEXT=CARDIOVASCULAR SYSTEM PREFTERM=IRRITATION VENT:  
 Variable Base Value Compare Diff. % Diff  
 CAUSAL . 7.000000 . .

PROT=101A INV=9626 PAT=1107 BODYTEXT=CARDIOVASCULAR SYSTEM PREFTERM=TACHYCARDIA VENT:  
 Variable Base Value Compare Diff. % Diff  
 RELDAY 2.000000 . .  
 AGEGROUP . 5.000000 . .  
 ALLERGY . 2.000000 . .  
 VERBATIM VENT TACHYCA VENT. TACHYC

PROT=102B INV=9405 PAT=113 BODYTEXT=BODY AS A WHOLE PREFTERM=HEADACHE:  
 Variable Base Value Compare Diff. % Diff  
 SERIOUS . 2.000000 . .  
 CAUSAL . 4.000000 . .  
 CLINSIG . 2.000000 . .  
 OUTCOME . 2.000000 . .  
 DEGREE . 1.000000 . .  
 RELATE . 2.000000 . .  
 ONDRUG . 1.000000 . .  
 CONT C  
 D\_TIME 0:00:00 -: -: -

PROT=102B INV=9528 PAT=301 BODYTEXT=SKIN & APPENDAGES PREFTERM=APPLICAT SITE REACT:  
 Variable Base Value Compare Diff. % Diff  
 ALLERGY . 2.000000 . .  
 VERBATIM URTICARIC RA URTICARIC RA  
 CONT C  
 D\_TIME -: -: - 0:22:45

PROT=201A INV=9229 PAT=807 BODYTEXT=CARDIOVASCULAR SYSTEM PREFTERM=IRRITATION VENT:  
 Variable Base Value Compare Diff. % Diff  
 DEGREE . 1.000000 . .  
 RELATE . 2.000000 . .  
 ONDRUG 1.000000 2.000000 1.000000 100.000000  
 VERBATIM VENT IRRITATI VENT IRRITAT  
 D\_TIME -: -: - <1 min

PROT=201A INV=9229 PAT=814 BODYTEXT=CARDIOVASCULAR SYSTEM PREFTERM=HEM:  
 Variable Base Value Compare  
 VERBATIM RT GROIN HEM RT GROIN HEM  
 D\_TIME 0:00:00 -: -: -

PROT=201A INV=9531 PAT=1009 BODYTEXT=METABOLIC & NUTRITIONAL SYSTEM PREFTERM=SGPT INC:  
 Variable Base Value Compare Diff. % Diff  
 RELATE . 1.000000 . .  
 LABFLAG . 1.000000 . .  
 ONDRUG . 1.000000 . .

NOTE: Data set MYDIR.ADR1 contains 1 observations not in MYDIR.ADR2.  
 NOTE: Data set MYDIR.ADR2 contains 3 observations not in MYDIR.ADR1.  
 NOTE: Values of the following 14 variables compare unequal: VERBATIM CONT SERIOUS CAUSAL CLINSIG OUTCOME  
 RELDAY DEGREE RELATE LABFLAG D\_TIME ONDRUG AGEGROUP ALLERGY

**LIST 3**

21SEP98 08:41 Comparison of BASE Data Set (MYDIR.ADR1) with COMP Data Set (MYDIR.ADR2) Page 1

## PROT = 101A INV = 9606 PAT = 618 BODYTEXT = CARDIOVASCULAR SYSTEM PREFTERM = IRRITATION VENT ##

\* BASE \* CAUSAL = .  
\* COMP \* CAUSAL = 7

## PROT = 101A INV = 9626 PAT = 1107 BODYTEXT = CARDIOVASCULAR SYSTEM PREFTERM = TACHYCARDIA VENT ##

\* BASE \* RELDAY = 2 AGEGROUP = . ALLERGY = .  
\* COMP \* RELDAY = . AGEGROUP = 5 ALLERGY = 2

VERBATIM = VENT TACHYCARDIA W/ INSERT OF PIGTAIL IN LV CEASING WHEN CATHETER MANIPULATED  
VERBATIM = VENT. TACHYCARDIA W/ INSERT. OF PIGTAIL IN LV CEASING WHEN CATHETER MANIPULATED

## PROT = 102B INV = 9405 PAT = 113 BODYTEXT = BODY AS A WHOLE PREFTERM = HEADACHE ##

\* BASE \* SERIOUS = . CAUSAL = . CLINSIG = . OUTCOME = .  
\* COMP \* SERIOUS = 2 CAUSAL = 4 CLINSIG = 2 OUTCOME = 2

DEGREE = . RELATE = . ONDRUG = .  
DEGREE = 1 RELATE = 2 ONDRUG = 1

CONT =  
CONT = C

D\_TIME = 0:00:00  
D\_TIME = -: -: -

## PROT = 102B INV = 9528 PAT = 301 BODYTEXT = SKIN & APPENDAGES PREFTERM = APPLICAT SITE REACT ##

\* BASE \* ALLERGY = .  
\* COMP \* ALLERGY = 2

VERBATIM = URTICARIC RASH AT AREA FO ELASTOPLAST APPLICATION  
VERBATIM = URTICARIC RASH AT AREA OF ELASTOPLAST APPLICATION

CONT = C  
CONT =

D\_TIME = -: -: -  
D\_TIME = 0:22:45

## PROT = 201A INV = 9229 PAT = 807 BODYTEXT = CARDIOVASCULAR SYSTEM PREFTERM = IRRITATION VENT ##

\* BASE \* DEGREE = . RELATE = . ONDRUG = 1  
\* COMP \* DEGREE = 1 RELATE = 2 ONDRUG = 2

VERBATIM = VENT IRRITATION SECONDARY TO THE CATHETER  
VERBATIM = VENT IRRITATION SECONDARY TO CATH MANIPULATION

D\_TIME = -: -: -  
D\_TIME = <1 min

## PROT = 201A INV = 9229 PAT = 814 BODYTEXT = CARDIOVASCULAR SYSTEM PREFTERM = HEM ##

\* BASE \* VERBATIM = RT GROIN HEMATOMA ATTEMPT TO CATHETERIZE (PROCEDURE WAS DONE ON LEFT GROIN)  
\* COMP \* VERBATIM = RT GROIN HEMATOMA ATTEMPT TO CATHETERIZE (PROCEDURE WAS DONE ON LEFT GROIN)

D\_TIME = 0:00:00  
D\_TIME = -: -: -

## PROT = 201A INV = 9531 PAT = 1009 BODYTEXT = METABOLIC & NUTRITIONAL SYSTEM PREFTERM = SGPT INC ##

\* BASE \* RELATE = . LABFLAG = . ONDRUG = .  
\* COMP \* RELATE = 1 LABFLAG = 1 ONDRUG = 1

**REFERENCES**

1. SAS® is a registered trademark of the SAS Institute, Cary, NC.
2. SAS Institute Inc., *SAS® Procedures Guide, Version 6, Third Edition*, Cary, NC, 1990
3. SAS Institute Inc., *SAS® Guide to Macro Processing, Version 6, Second Edition*, Cary, NC, 1990, P. 256.

**ACKNOWLEDGMENTS**

My heartfelt appreciation goes to Michael Walega of Covance, Inc. Princeton, NJ, who provided an extensive review of this paper and gave me valuable advise.

**CONTACT INFORMATION**

Your comments and questions are valued and encouraged. Contact the author at:

WeiQin Pan  
Covance, Inc.  
210 Carnegie Center  
Princeton, NJ 08540 - 6233  
Phone: (609) 452 - 4751  
Fax: (609) 732 - 9119  
Email: weiqin.pan@covance.com