

Paper 175-26

**A SAS Macro Program for Star-plots**

Chuanchieh Hsu, AHP/Whitehall-Robins Healthcare, Madison, NJ  
 Zhongwei Zhou, Novartis Pharmaceuticals Co., East Hanover, NJ  
 J. Michael Hardin, University of Alabama at Birmingham, Birmingham, AL

**ABSTRACT**

In a variety of research settings including clinical trials, education, and psychology, it is very common to compare a composite score, consisting of several individual items, across the treatment groups. However, often the researcher also wishes to compare differences in the individual items comprising the composite score across the various treatment groups. A useful graphical technique to assist the researcher in visualizing and interpreting the results of such comparisons is the Star plot. The Star plot displays multivariate data by coding the values of the variables into the lengths of the rays emanating from the center of the plot. This paper illustrates how to use SAS/BASE®, SAS® MACRO, SAS/STAT®, and SAS/GRAPH® to generate a star-plot, which shows the differences of each individual item among the groups. In addition, the results of a pairwise comparison for each item are also presented in the graph.

**INTRODUCTION**

The primary use of graphics is to display either large quantities of information or complicated statistical results for quick and easy interpretation. Bar charts, scatter plots, pie charts, and line charts are the most commonly used graphs. However, some specially designed graphs are often helpful for specific analysis needs.

In clinical trials as well as many other areas such as education, psychometrics and psychology, researchers often wish to compare a composite score that consists of several individual items across the treatment groups. It is also often important to know if there is any difference between individual items among the groups. In this paper, a star-plot is presented that shows the differences of each of these individual item among the treatment groups, and shows the pairwise comparisons. Such a graphical representation provides an easy visualization for researchers to assist in their interpretations of the magnitude of differences for each of the items among the groups. A SAS program has been developed using SAS/BASE®, SAS® MACRO, SAS/STAT®, and SAS/GRAPH® to perform the plotting task.

**PREPARING THE DATA**

The example used in this paper is a fictitious clinical study with three treatment groups, Placebo, Treatment 1, and Treatment 2 under study. Each subject is assumed to complete a questionnaire containing six items. As per the protocol, the composite score of the six individual items will be analyzed via a generalized linear model and then followed by appropriate pairwise comparison tests between the treatment groups. Additionally, each individual item also will be examined by the same type of analysis. The star-plot presented in this paper will assist the investigator that wants to examine differences between the groups for each individual item

For the convenience, we use a DATA statement to import the least squared means of the groups by each of the individual items and the p-values of a pairwise comparison among the groups.

```
*****
* Import Least Square Means. *
* l1, l2, and l3 are the least squared *
* means for placebo, and treatment groups *
* 1 and 2, resp. *
*****
```

```
*****;
data LSMSE;
  infile "/data subdirectory/lsmse.dat";
  input item l1 l2 l3;
  cards;
1 1.01 1.046 1.155
2 ...
...
6 ...
run;
```

The data LSMSE, which are the least square means for each item by treatment group, are then transposed to the following format.

Item	Rx	M
1	1	1.01
2	1	0.91
3	1	1.277
...	...	...
1	2	1.046
2	2	1.115
3	2	1.267
...	...	...

The transpose can be easily done by the following code or by array statements.

```
data stat(drop=l1 - l3);
  rx = 1; m = l1; output;
  rx = 2; m = l2; output;
  rx = 3; m = l3; output;
run;
```

The data PAIR, for the p-values from the pairwise comparison analysis are as follows. The variables p12, p13, and p23, are the p-values of the comparisons for placebo vs. treatment 1, placebo vs. treatment 2, and treatment 1 vs. treatment 2, respectively.

Item	P12	P13	P23
1	0.797	0.311	0.450
2	0.045	0.049	0.986
...	...	...	...

```
data pair;
  infile "/data subdirectory/pvalues.dat";
  input item p12 p13 p23;
  cards;
1 0.797 0.311 0.450
...
6 ...
run;
```

It is necessary to make the numeric variables, such as item and rx, clearer in the graph by using Proc Format.

```
proc format;
  value tp
    1='Item 1' 2='Item 2' 3='Item 3'
```

```

    4='Item 4' 5='Item 5' 6='Item 6';
value rx 1='Placebo' 2='Trtmnt 1'
        3='Trtmnt 2';
run;

```

Several macro variables need to be preset prior to running the star-plot macro program. For example,

```

/* Make &timept in the macro equals to
variable, item */
%let timept=item;

/* Format of treatment variables*/
%let rx=rx;

/* Number of treatment groups */
%let numrx = 3;

/* symbols to denote significance of pairwise
comparison */
%let psymbols=%str(* +);

/* the colors for each treatment group */
%let colors=black blue red;

/* line types for each treatment group */
%let lines=1 2 8;

/* Number of individual items in the graph
*/
%let s=6;

```

Since a star-plot depicts the improvements in each individual item among the groups, it presents six equal-length sticks, based on the example used in this paper. Each item is shown in the graph in clockwise order with the first item at the 12:00 position. In each treatment group, the scores for the items are connected, and statistical significance indicators versus placebo are indicated on the side for each item. To ensure there is adequate spacing between the end of the lines and the mean values, the length of the radiate lines is computed as the maximum value among all the means multiplies by 1.25. The value of the radiate lines of the plot is then merged with the data STAT.

```

%let incre=0.25;
proc sql feedback;
  create table stat1 as
  select *, max(m)*(1+&incre) as max
  from stat;
quit;

```

The next step is to generate the radiate lines. To do so, the angle between the lines must be calculated. For example, each angle will be  $360/6$  degrees if there are 6 individual items. The following is the predefined line and text sizes of a star-plot.

```

%let rlinesz=0.3; * width of radio;
  * text size for item description;
%let textsz=2;
  * length of radiate lines;
%let rlinelen=25;
  * length of ends of radiate lines;
%let endwidth=0.5;
  * item descriptions for the radiate lines;
%let tpfmt=tp;

```

Since the graph is usually landscaped, y-axis is shrunk and has different scale from the x-axis. By multiplying a constant, 11/8.5 (letter size), to y-axis coordinate makes it same scale as the x-axis.

```

%let ratio=%str(11/8.5);

```

The main part of the program is given below. It uses SAS/Graph<sup>®</sup> annotation for the graph, which calculates the angle of the radiate lines, draws the radiate lines and tip mark at the end of each line, adds text onto the end of each line, and a mark at the middle of the 12-o'clock position line, etc. These tasks are very tedious because all the marks and texts need to be adjusted by the angle of the lines.

```

%let pi=3.1415926;
data radio(keep=xsys ysys hsys text position
           function size x y);
length xsys ysys hsys $1 function $8
       text $200;
retain xsys ysys hsys '3' pi &pi size text
       position x y function;
set maxval end=eof;

halfmax=max*0.5;
function='label';
size=&textsz;
position='4'; x=49;
y=50+&rlinelen*&ratio; text=put(max, 8.1);
output;
  * to draw a mark at the mid-point on
  * the 12-o'clock axis;
y=50+&rlinelen*&ratio*0.5;
text=put(halfmax, 8.1);
output;

do i=1 to &s;
  *** draw Radiate Lines ***;
  size=&rlinesz;
  function='move'; x=50; y=50; output;
  function='draw';
  x=50+&rlinelen*cos(pi/2-2*pi*(i-1)/&s);
  y=50+&rlinelen*sin(pi/2-2*pi*(i-
    1)/&s)*&ratio; output;

  *** Draw line ends ***;
  function='move';
  x=50+&rlinelen*cos(pi/2-2*pi*(i-1)/&s);
  x=x-&endwidth*cos(pi-2*pi*(i-1)/&s);
  y=50+&rlinelen*sin(pi/2-2*pi*(i-
    1)/&s)*&ratio;
  y=y-&endwidth*sin(pi-2*pi*(i-
    1)/&s)*&ratio; output;
  function='draw';
  x=50+&rlinelen*cos(pi/2-2*pi*(i-1)/&s);
  x=x+&endwidth*cos(pi-2*pi*(i-1)/&s);
  y=50+&rlinelen*sin(pi/2-2*pi*(i-
    1)/&s)*&ratio;
  y=y+&endwidth*sin(pi-2*pi*(i-
    1)/&s)*&ratio; output;

  *** Draw a mark on middle of the line ***;
  function='move';
  x=50+0.5*&rlinelen*cos(pi/2-2*pi*(i-
    1)/&s);
  x=x-&endwidth*cos(pi-2*pi*(i-1)/&s);
  y=50+0.5*&rlinelen*sin(pi/2-2*pi*(i-
    1)/&s)*&ratio;
  y=y-&endwidth*sin(pi-2*pi*(i-
    1)/&s)*&ratio; output;
  function='draw';
  x=50+0.5*&rlinelen*cos(pi/2-2*pi*(i-
    1)/&s);
  x=x+&endwidth*cos(pi-2*pi*(i-1)/&s);
  y=50+0.5*&rlinelen*sin(pi/2-2*pi*(i-
    1)/&s)*&ratio;
  y=y+&endwidth*sin(pi-2*pi*(i-
    1)/&s)*&ratio; output;

```

```

*** Write text for each item ***;
function='label'; size=&textsz;
text=put(i, &tpfmt..);
x=50+&rlinelen*cos(pi/2-2*pi*(i-1)/&s);
y=50+&rlinelen*sin(pi/2-2*pi*(i-
1)/&s)*&ratio;
if abs(x-50)<1 then
do;
if y<50 then y=y-3;
else y=y+3;
x=50; position='5';
end;
else if x>50 then do;
x=x+1;
position='6';
end;
else do;
x=x-1; position='4';
end;
output;
end;
run;

```

Now, the radiate lines are ready for the plot. The next step is to place the least squares means or simply the means, (or whatever the analyst uses for the pairwise comparison), on the lines, and then connect the lines within each treatment group.

```

%let plinesz=10; * width of connected lines;

* sort data by trtment and item #;
proc sort data=stat out=stat;
by &rx &timept;
run;

data poly(keep=xsys ysys hsys color line
function size x y);
length xsys ysys hsys $1 function color
$8;
retain xsys ysys hsys '3' size &plinesz
pi &pi firstm x y function color
line;
set stat end=eof;
by &rx &timept;
if first.&rx then firstm=m;
if &timept=1 then function='poly';
else function='polycont';
x=50+cos(pi/2-2*pi*(&timept-
1)/&s)*&rlinelen*m/max;
y=50+sin(pi/2-2*pi*(&timept-
1)/&s)*&ratio*&rlinelen*m/max;
%do i=1 %to &numrx;
if &rx=&i then do;
color="&&color&i";
line="&&line&i";
end;
%end;
run;

```

Since pairwise comparisons are conducted for each item, it is necessary to determine which pair's comparisons are significant, so that an indicator can be added to show the significance level at either 0.05, 0.01, or 0.001. To show statistical significance between the treatments for a particular item, significance symbols are added on the top of the item description, which is located at the end of the radiate line for that item. Different symbols can be applied by setting the macro variable, PSYMBOLS. Different levels of significance for each symbol are then added to the graph based on the p-values in the data set, PAIR. For example, a \* may represent a significant difference between the placebo and Treatment 1 groups at the level of 0.05, and a ++ may denote a significant improvement of

Treatment 2 compared to the placebo at level of 0.01.

```

data symbol(keep=xsys ysys hsys text position
function size x y);
length xsys ysys hsys $1 function $8 text
$200;
retain xsys ysys hsys '3' pi &pi size
&textsz function 'label' text
position x y k;
set pair end=eof;
k=0;
%do i=&numrx %to 2 %by %str(-1);
%let j=%eval(&numrx-&i+1);
if p1&i<=0.05 then do;
k=k+1;
if p1&i<0.001 then text =
"&&psymbol&j.&&psymbol&j.&&psymbol&j";
else if p1&i<0.01 then text =
"&&psymbol&j.&&psymbol&j";
else if p1&i<=0.05 then text =
"&&psymbol&j";
x=50+&rlinelen*cos(pi/2-
2*pi*(&timept-1)/&s);
y=50+&rlinelen*sin(pi/2-
2*pi*(&timept-1)/&s)*&ratio-0.5;
if abs(x-50)<1 then do;
x=51; position='6';
if y>50 then y=y-k*(size+0.2);
else y=y+k*(size+0.2);
end;
else if x>50 then do;
x=x+0.5*length(put(&timept,
&tpfmt..));
y=y-k*(size+0.2); position='5';
end;
else do;
x=x-0.5*length(put(&timept,
&tpfmt..));
y=y+k*(size+0.2); position='5';
end;
output;
end;
%end;
run;

```

A legend at top-right corner of the plot area indicating the line types, line colors, and the thickness of the lines is also helpful to identify treatment groups.

The last step of the procedure is to put all the above annotations together to generate the star-plot. Example code to accomplish this is shown below:

```

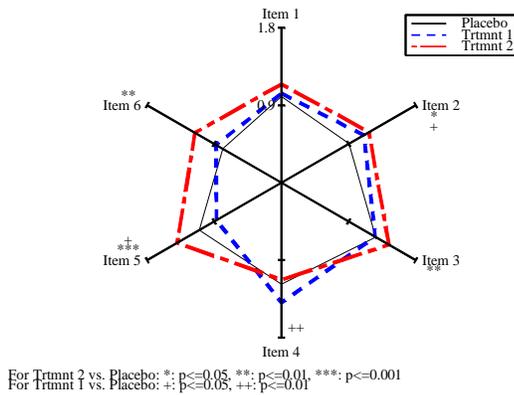
data anno;
set legend radio poly symbol;
run;

proc gslide gout=sugi26 annotate=anno;
run;
quit;

```

A graph of the generated star-plot is presented in Figure 1.

Figure 1

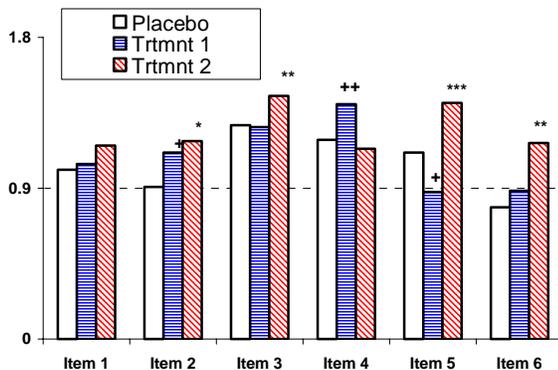


**A COMMONLY USED GRAPH**

A bar chart is commonly used to demonstrate the differences between the treatment groups. Figure 2 depicts the mean values of each treatment group for all six items using a bar chart. The problem with such a graph are that

1. there will numerous bars in one graph if there are several individual items and several groups; for example, 40 bars for a study with 10 items and 4 treatment groups.
2. it is difficult to ascertain if there is a trend showing that the Treatment 2 group is better than the Treatment 1 group, and/or the Treatment 1 group is superior to the placebo.

Figure 2



**CONCLUSION**

A star-plot provides a clear and easy visual presentation for investigators to visualize the differences among groups, and to compare individual items.

The codes and explanations presented in this paper are only one of many ways the graph could have been produced. Modifications to the graph, such as titles, footnotes, fonts, and graph output format, can easily be made by SAS programmers. SAS/GRAPH® software offers the programmers many creative ways to meet these particular needs.

**REFERENCES**

SAS Institute Inc., SAS/GRAPH software, Volumes 1 & 2: Reference Version 6, First Edition, Cary NC: SAS Institute Inc.,

1990.

SAS Institute Inc., SAS/STAT User's Guide, Volumes 1 & 2, Version 6, Fourth Edition, Cary NC: SAS Institute Inc., 1990.

SAS Institute Inc., SAS Guide to Macro Processing, Version 6, Second Edition, Cary NC: SAS Institute Inc., 1990.

SAS Institute Inc., SAS/BASE User's Guide, Version 6, Fourth Edition, Cary NC: SAS Institute Inc., 1990.

**CONTACT INFORMATION**

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

Chuanchieh Hsu, Ph.D. Statistician  
 American Home Products/Whitehall-Robins Healthcare  
 Phone: (215) 574-3174  
 Email: [jhsu60@hotmail.com](mailto:jhsu60@hotmail.com)

Zhongwei Zhou, M.S. Scientific Analyst, Novartis Pharmaceuticals Co.  
 Phone: (973) 781-5652  
 Email: [Zhongwei.Zhou@pharma.novartis.com](mailto:Zhongwei.Zhou@pharma.novartis.com)

J. Michael Hardin, Ph.D. Professor, University of Alabama at Birmingham.  
 Phone: (205) 975-9895  
 Email: [mhardin@uab.edu](mailto:mhardin@uab.edu)