Paper 041-2007

Graphical Representation of Mean Measurement over Time

Wei Cheng, Isis Pharmaceuticals, Inc., Carlsbad, CA

ABSTRACT

A graph of mean measurement over time is one of the most commonly used statistical graphs for data exploration and presentation. It can be as simple as a line plot, or can be a complex graph involving lots of annotation. This paper illustrates some useful techniques provided by SAS/GRAPH and ODS Statistical Graphics to generate graphical displays of mean measurement over time. Programs using these techniques are examined in detail. The examples start from a basic plot, gradually evolve to a complex plot.

INTRODUCTION

In many experiments, repeated measurements are taken at regular intervals. Data are collected over a period of time, a graph of mean or median values over time is helpful to compare the evaluations between different groups or find out the trends. For instance, in hypertension or obesity it is a simple matter to record each patient's blood pressure or body weight before a new treatment starts and at weekly intervals. In this paper, a clinical trial setting will be used. The subjects are required to return to the clinic for multiple visits during the trial with laboratory measurements made at each.

EXAMPLE DATA SET

There are four variables in the example data set EXAMPLE_DATA. Treatment arm ARMCD, subject identifier SUBJID, visit number VISITNUM, and laboratory test result RPTRESN are entered for illustrative purpose. The SAS code for creating this data set is shown below.
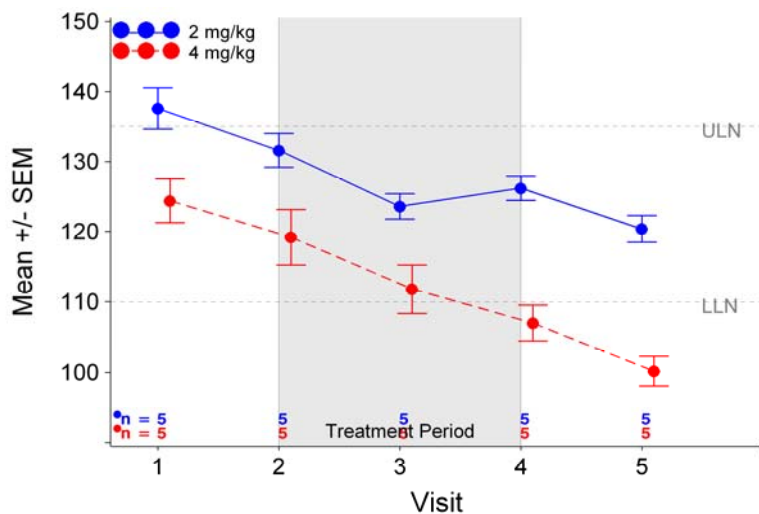
PROGRAM 1

```
data example_data;
    input armcd $ subjid $ visitnum rptresn;
    label visitnum = 'Visit' armcd = "Treatment";
    cards;
A 001 1 135
A 001 2 138
A 001 3 125
A 001 4 124
A 001 5 120
A 002 1 142
......
B 006 1 125
B 006 2 120
B 006 3 115
B 006 4 110
B 006 5 105
B 007 1 115
......
B 010 4 110
B 010 5 96
;
run;
```

CREATE GRAPHS FROM SAS/GRAPH

Figure 1 is a graph of mean laboratory test result measured on two groups of subjects over five clinical visits. There are lines connecting the mean values, there are numbers of subject for each group at each visit shown in the bottom of the graph, there are standard error bars marked for each mean value, there are laboratory normal ranges and treatment period shown on the graph.

FIGURE 1

The following programs will show you step by step how to create this graph.
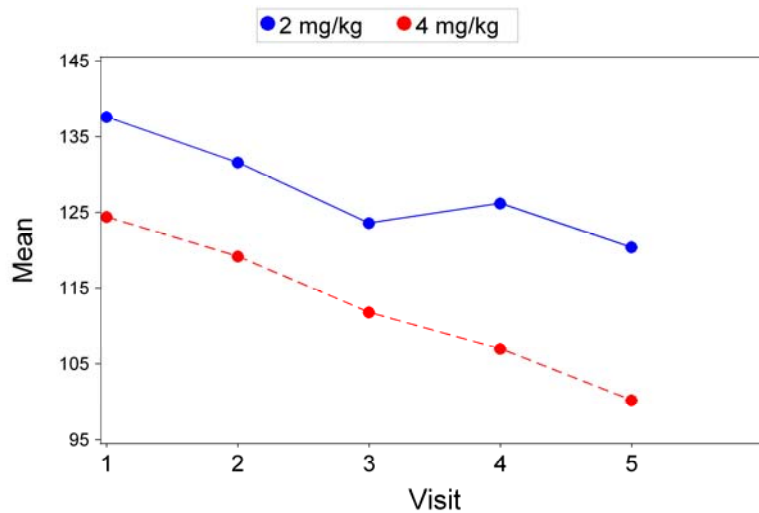
PROGRAM 2

```
proc sort data = example_data;
    by armcd visitnum;
run;
proc summary data = example_data;
    by armcd visitnum;
    var rptresn;
    output out = means (drop = _:) mean = mean n = n stderr = stderr;
run;                                                                          ❶
symbol1 i = j value = W font = marker c = blue h = 2.5 line = 1 width = 5;
symbol2 i = j value = W font = marker c = red h = 2.5 line = 3 width = 5;    ❷
proc gplot data = means;
    plot mean * visitnum = armcd;
    format armcd $armf.;
run; quit;                                      ❸
```

FIGURE 2

2

❶ PROC SORT and PROC SUMMARY are used to generate descriptive statistics MEAN, N, and STDERR for the graph.

❷ The simplest interpolation is JOIN which draws lines connecting the data points. Join plots are often used to show trends in the data. LINE = 1 will draw a solid line; other line types are different kinds of dotted lines.

❸ PROC GPLOT is used to draw the line plots in figure 2.

PROGRAM 3

```
data meanstd;
    set means;
    if armcd = "B" then visitnum = visitnum + 0.1;
    yvalue = mean; output;
    yvalue = mean + stderr; output;
    yvalue = mean - stderr; output;
run;                                                                           ❶

symbol1 i = hiloj value = W font = marker c = blue h = 2.5 line = 1 width = 5;
symbol2 i = hiloj value = W font = marker c = red h = 2.5 line = 3 width = 5; ❷

proc gplot data = meanstd;
    plot yvalue * visitnum = armcd / vref = 110 135 lvref = 2 href = 2 4 lhref = 1;
    format armcd $armf.;
run; quit;                                                                     ❸
```
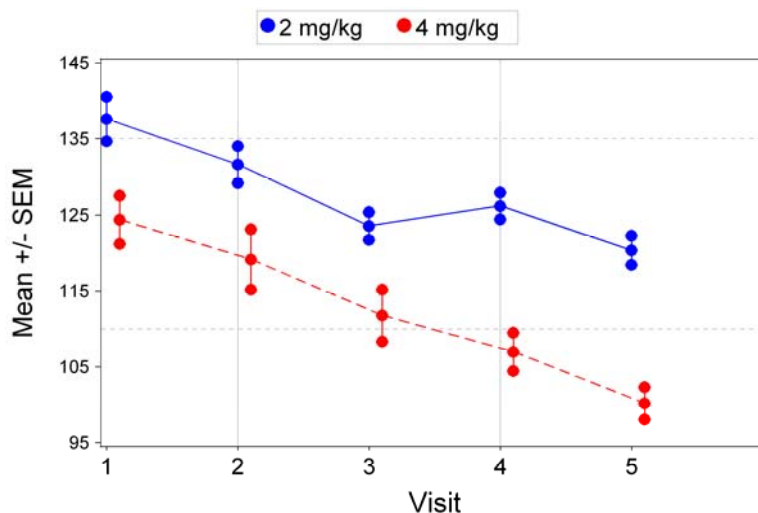
❶ The values for VISITNUM are slightly shifted between groups to avoid overlapping on the graph. There are two more observations been generated at each mean value for the standard error bars.

❷ I = HILO specifies that a solid vertical line connect the minimum and maximum y values for each x value. J specifies that the mean values are joined with a line.

❸ VREF and HREF specify the laboratory normal ranges and treatment period.

Figure 3 is generated by Program 3:

FIGURE 3

PROGRAM 4

```
data numlabel;
    length function sytle text $ 8;
    retain function 'label' xsys ysys '2' hsys '3' style 'swissb' size 2 when 'a';
    set means;
```

```
     y = mean;
     text = compress(n);
     if armcd = 'A' then do;
           color = 'blue'; x = visitnum + 0.1;
     end;
     else if armcd = 'B' then do;
           color = 'red'; x = visitnum + 0.2;
     end;
     output;
run;                                                            ❶

proc gplot data = meanstd;
     plot yvalue * visitnum = armcd / vref = 110 135 lvref = 2
                                     href = 2 4 lhref = 1 annotate = numlabel; ❷
     format armcd $armf.;
run; quit;
```
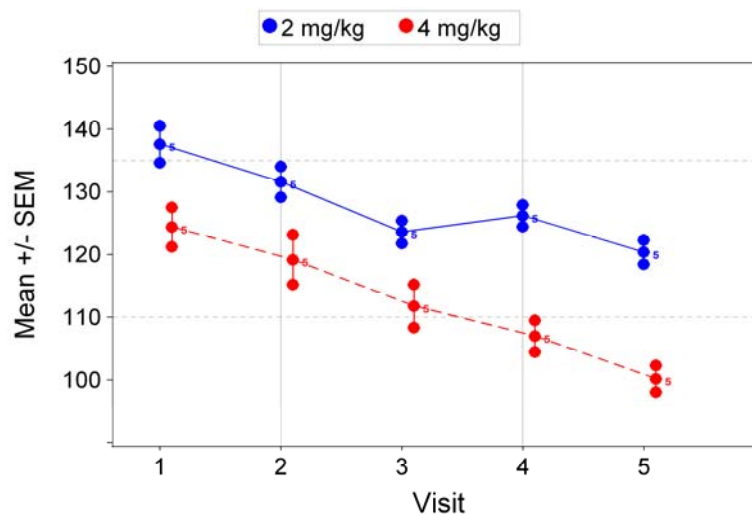
❶ NUMLABEL is an annotate data set to have numbers of subject at each visit be drawn next to the mean values, as shown in Figure 4.

❷ NUMLABEL is specified in the ANNOTATE option in the PLOT statement.

FIGURE 4

PROGRAM 5

```
data anno1;
     length function $8 text $8;
     set means;
     retain xsys ysys '2' hsys '1';
     if n > 0 then do;
           function = "Label"; x = visitnum; position = "5"; size = 3.5;
           text = put(n, 2.); style = "swissb"; angle = 0;
           if armcd = "A" then do;
                 y = 94; color = "blue"; output;
           end;
           else do;
                 y = 92; color = "red"; output;
           end;
     end;
run;                                                            ❶

data anno2;
```

4

```
    length function $8 text $8 style $6;
    retain xsys hsys "1" ysys "2" postion "5" angle 0 size 3.5;
    armcd = "A"; x = 1; y = 94; style = "none"; color = "blue"; text = 'dot';
    function = 'symbol'; output;
    armcd = "A"; x = 4; y = 94; style = "swissb"; color = "blue";      text = "n =";
    function = 'label'; output;
    armcd = "B"; x = 1; y = 92; style = "none"; color = "red"; text = 'dot';
    function = 'symbol'; output;
    armcd = "B"; x = 4; y = 92; style = "swissb"; color = "red"; text = "n =";
    function = 'label'; output;
run;                                                        ❷

data anno3;
    length function $8 text $8;
    retain xsys "3" ysys "2" function "label" size 3 color "gray" position "6" x 90;
    y = 135; text = "ULN"; output;
    y = 110; text = "LLN"; output;
run;                                                        ❸

data anno4;
    length xsys ysys hsys when position $1 function $8 style color $20 text $50;
    retain xsys "2" ysys "1" hsys "1";
    function = "move"; x = 2; y = 0; output;
    function = "bar"; color = "grayE8"; style = "solid"; line = 3; when = "b";
    x = 4; y = 100; output;
    function = "move"; x = 3; y = 2; output;
    function = "label"; text = "Treatment Period"; color = "black"; position = "b";
    style = "'Arial'"; size = 5; when = "a"; output;
run;                                                        ❹

data anno;
    set anno4 anno1 anno2 anno3 ;
run;                                                        ❺

symbol1 i = hiloctj c = blue line = 1 width = 5;
symbol2 i = hiloctj c = red line = 3 width = 5;             ❻
symbol3 i = none value = W font = marker c = blue h = 2.5;
symbol4 i = none value = W font = marker c = red h = 2.5;   ❼

axis1 label = (h = 6 a = 90 "Mean +/- SEM")
      minor = none value = (height = 4)
      major = (width = 3 c = black)
      width = 3 value = (t=1 " ")
      order = (90 to 150 by 10);
axis2 label = (h = 6 "Visit")
      minor = none value = (height = 4)
      major = (width = 3 c = black)
      value = (t=6 " ") width = 3
      order = (1 to 6 by 1)
      offset = (6,0);

axis3 label = none
      order = (90 to 150 by 10)
      value = none major = none
      minor = none style = 0;

legend1 label = none
        position = (top left inside)
        offset = (0.5, -0.25)
        across = 1
        value = none
```

5

```
         shape = symbol(8,3.5)
         mode = share;

legend2 label = none
         position = (top left inside)
         offset = (0.5)
         across = 1
         shape = symbol(8,3.5)
         mode = share
         value = (h =4 t= 1 "2 mg/kg" t = 2 "4 mg/kg");          ❽

proc gplot data = meanstd;
    plot yvalue * visitnum = armcd / vaxis = axis1 haxis = axis2 legend = legend1
                                     vref = 110 135 lvref = 2
                                     href = 2 4 lhref = 1 annotate = anno;

    plot2 mean * visitnum = armcd / vaxis = axis3 haxis = axis2 legend = legend2;
    format armcd $armf.;
run; quit;                                                                 ❾
```

❶ ANNO1 is an annotate data set to label the numbers of subject at each visit.

❷ ANNO2 is an annotate data set to label the symbols and "n=" in front of the numbers of subject at each visit.

❸ ANNO3 is an annotate data set to label the laboratory normal ranges.

❹ ANNO4 is an annotate data set to label the treatment period.

❺ Final annotate data set ANNO is created by concatenating the four annotate data sets together.

❻ Interpolation C draws tick marks at the close value instead of at the mean value. TJ adds tops and bottoms to each line, and joins the mean values.

❼ I = none produce plots with unconnected scatter plots.

❽ AXIS3 and LEGEND2 statements define y-axis and legend for PLOT2 statement in the PROC GPLOT procedure.

❾ PLOT statement draws the standard error bars and line segments connecting the mean points. PLOT2 statement draws the symbols for the mean values.

The resulting graph of mean laboratory test result over time, after submitting the preceding code in Program 5, is shown in Figure 1.

CREATE GRAPHS FROM ODS STATISTICAL GRAPHICS
In SAS 9, a number of SAS/STAT procedures can use the ODS to create statistical graphics automatically. You can use PROC TEMPLATE to create your statistical graphics templates and generate statistical graphics from DATA step. The following programs will show you how to generate the same graph of mean laboratory test result over time by ODS Statistical Graphics.

<div align="center">

PROGRAM 6

</div>

```
data means;
    set means;
    label mean = "Mean +/- SEM";
    elower = mean - stderr;
    eupper = mean + stderr;                                          ❶
run;
proc template;
 define statgraph mygraphs.meanover;
   layout gridded;
```

```
        entrytitle "Mean Measurement Over Time" ;
        layout overlay;
            scatterplot x = visitnum y = mean
                        / group = armcd name = "treat"
                          yerrorlower = elower
                          yerrorupper = eupper
                          markersize = 12px;
            seriesplot  x = visitnum y = mean
                        / group = armcd;
         endlayout;
         discretelegend "treat";
    endlayout;
 end;
run;                                                         ❷

ods html;
ods graphics on / reset  imagename='meanover' imagefmt = png;    ❸

data _null_;
  set  means;
  file print ods = (template='mygraphs.meanover');
  put _ods_;
run;                                                             ❹

ods graphics off;
ods html close;
```
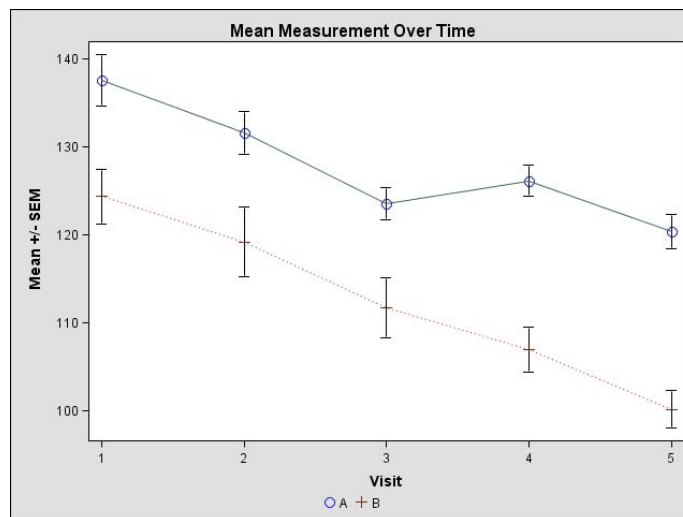
❶ Data values for mean and standard error bars are generated from DATA step.

❷ Starting in SAS9.1, a new template type STATGRAPH, which describes the layout and appearance of a graph, was added to the ODS template type family. You can use PROC TEMPLATE to define the graphic template for the graph of mean measurement over time.

❸ ODS GRAPHICS statement will turn on the ODS Statistical Graphics facility. It is integrated with ODS and supports popular image formats such as PNG, JPEG, etc.

❹ Finally a DATA step is used to feed the data to the graphics template via FILE and PUT statement.

Figure 5 is generated by Program 6:

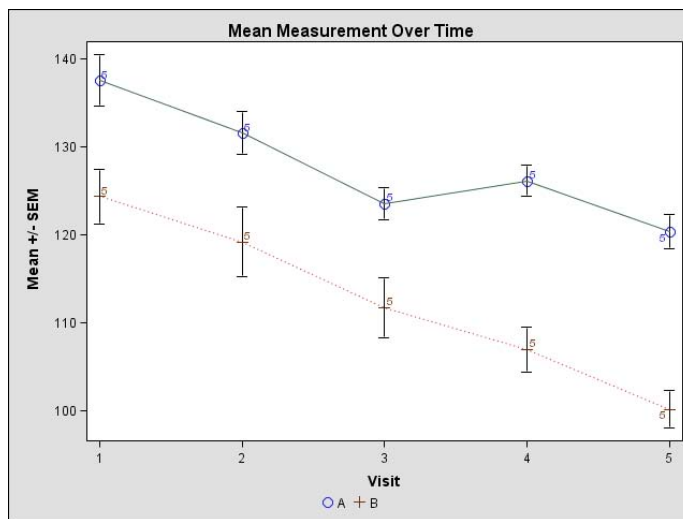**FIGURE 5**

PROGRAM 7

```
proc template;
 define statgraph mygraphs.meanover;
    layout gridded;
        entrytitle "Mean Measurement Over Time" ;
        layout overlay;
            scatterplot x = visitnum y = mean
                    / group = armcd name = "treat"
                      yerrorlower = elower
                      yerrorupper = eupper
                      markersize = 12px
                      datalabel = n ;                    ❶
            seriesplot  x = visitnum y = mean
                    / group = armcd;
         endlayout;
         discretelegend "treat";
    endlayout;
 end;
run;
```

❶ DATALABEL option can be used to add the numbers of subject at each visit, which is shown in Figure 6:

FIGURE 6

CONCLUSION

SAS/GRAPH global statements, GOPTIONs, options in the GPLOT procedure, and the annotate facility are keys to generate graphs of mean measurement over time. ODS Statistical Graphics greatly enhances the ability of graphical representations, offers maximum flexibility and maintainability. Clinical trail data are used as an example is this paper, it can be extended to any other repeated measures over time.

REFERENCES

1. SAS Institute, Inc. (2002-2004), *SAS/Graph Reference, SAS OnlineDoc® 9.1.3,* Cary, NC: SAS Institute, Inc.

2. SAS Institute, Inc. (2005), "TEMPLATE Procedure: Creating ODS Statistical Graphics Output (Experimental)"
< http://support.sas.com/rnd/base/topics/statgraph/proctemplate/> (March 4, 2007).

3. SAS Institute, Inc. (2005), "ODS Statistical Graphics (Experimental) and ODS Styles: Usage and Reference"
< http://support.sas.com/rnd/base/topics/statgraph/v91StatGraphStyles.htm> (March 4, 2007).

5. Cheng, Wei (2006), "ODS Statistical Graphics for Clinical Research" *Proceedings of the 31th Annual SAS Users Group International (SUGI) Conference.*

ACKNOWLEDGMENTS

CONTACT INFORMATION

I welcome and appreciate your comments and questions. Contact the author at:

Wei Cheng,
Isis Pharmaceuticals, Inc.,
1896 Rutherford Rd., Carlsbad, CA 92008
(760) 603-3807
Email: wcheng@isisph.com