

Paper 175-31

Creating Bar Charts and Scatter Plots on the Same Page, Using PROC GPLOT or PROC GBARLINE in SAS® ?

Rita Tsang, Averion Inc., Framingham, Massachusetts

ABSTRACT

The SAS GPLOT and GCHART procedures are powerful tools for generating graphical displays of clinical data, as they are usually employed to create scatter plots and bar charts respectively. However, we often need to display both a scatter plot and a bar chart within the same graph. The GCHART procedure cannot generate scatter plots and it does not have the ability to superimpose graphical displays. How do we get around this limitation?

Previously, my approach to the problem was to use the SAS DATA step (Tsang, 2005) and SAS ANNOTATE (Tsang, 2005) in GPLOT to accomplish the task.

In this paper, I will further explore the functionality of the new SAS version 9 PROC GBARLINE, which, was also created to tackle the limitation of GCHART (SAS Institute, 2005). A comparison of the two procedures, GPLOT and GBARLINE, will be given.

INTRODUCTION

This paper will delineate two main approaches to create bar charts and scatter plots on the same page in GPLOT, namely using DATA MANIPULATION and SAS ANNOTATE. The new GBARLINE procedure, now available in SAS version 9, will also be explored.

THE DATA

Our task at hand is to display the urine output data (Urine__) in a vertical bar chart and plasma creatinine data (Labval) in a scatter plot by assessment date (Asdt) for each patient. Here is a sample display of the Urine__ data.

Site	Randno__	Asdt	Labval	Urine__
3	1101	14MAR2004:20:00:00		27
3	1101	15MAR2004:04:00:00		595
3	1101	15MAR2004:04:20:00	1.4	
3	1101	15MAR2004:12:00:00		329
3	1101	15MAR2004:14:45:00		29
3	1101	15MAR2004:16:45:00		24
3	1101	15MAR2004:18:45:00		6
3	1101	15MAR2004:20:45:00	1.1	48
3	1101	15MAR2004:22:45:00		53
3	1101	16MAR2004:00:45:00		53
3	1101	19MAR2004:03:00:00	1.2	

1. DATA MANIPULATION

The following DATA step is a simple manipulation of the clinical data to create additional records of zero value for urine output (Urine1) and missing value for plasma creatinine (Labval) at each assessment time point. By creating these extra records, we can create vertical bars for the urine output data at each assessment time point by linking the extreme values, zero and the original values, using the INTERPOL=BOX00F option in the GPLOT SYMBOL statement. With the plasma creatinine values set to missing, SAS GPLOT simply will not display those observations.

```
data urine;
  set urine__;
  by site randno__ asdt;
  urine1=urine__;
  output;
  if urine__>.then do;
    urine1=0;
    labval=.;
    output;
  end;
run;
```

In PROC GPLOT, we can produce box and whisker plots by specifying INTERPOL=BOX in the SYMBOL statement (SAS Institute, 2002). The INTERPOL=BOX00F will display a box ranges from the high and low extremes without displaying the whiskers, and the box will be filled by the color specified in the CV option.

```
/*Symbol statement for the bar chart.*/
Symbol1 v=none interpol=box00f cv=blue bwidth=2;
```

This above SYMBOL statement will connect the high and low extreme values on the x-axis (INTERPOL=BOX00F) with no plot symbol (V=NONE). The box will be filled with the color black (CV=BLACK). The BIWIDTH option specifies the width of the box. The width is about 0.4 cm in the current graphical display.

CREATING A SCATTER-PLOT SUPERIMPOSED ON THE BAR-CHART: A SECOND PLOT STATEMENT

By employing two PLOT statements in GPLOT, we can superimpose a scatter plot of plasma creatinine on the bar chart of urine output with different scales on the same graph (SAS Institute, 2002). The first PLOT statement (PLOT URINE1*ASDT) displays the bar chart of urine output by assessment date. The second PLOT statement (PLOT2 LABVAL*ASDT) displays a scatter plot of plasma creatinine by assessment date. Figure 1 shows the graphical display using DATA MANIPULATION

```
/*Symbol statement for the scatter plot.*/
symbol2 h=0.17 in v='P' color=green cv=green interpol=join l=1;

/*Axis statement for the X-axis.*/
axis1 label=(h=0.17 in f=triplex "Assessment date and time") value=(h=0.135 in
f=triplex) minor=none order=("14MAR2004:12:00:00"dt to "19MAR2004:12:00:00"dt
by dtday) offset=(1 pct) color=black;

/*Axis statement for the left axis.*/
axis2 label=(h=0.17 in a=90 f=triplex "Urine Output (CC)") minor=none order=(0
to 2100 by 100) offset=(5 pct) value=(h=0.135 in f=triplex) color=black;

/*Axis statement for the right axis.*/
axis3 label=(h=0.17 in a=270 f=triplex "Plasma Creatinine (mg/dl)") minor=none
order=(0 to 2 by 1) offset=(5 pct) value=(h=0.135 in f=triplex) color=black;
```

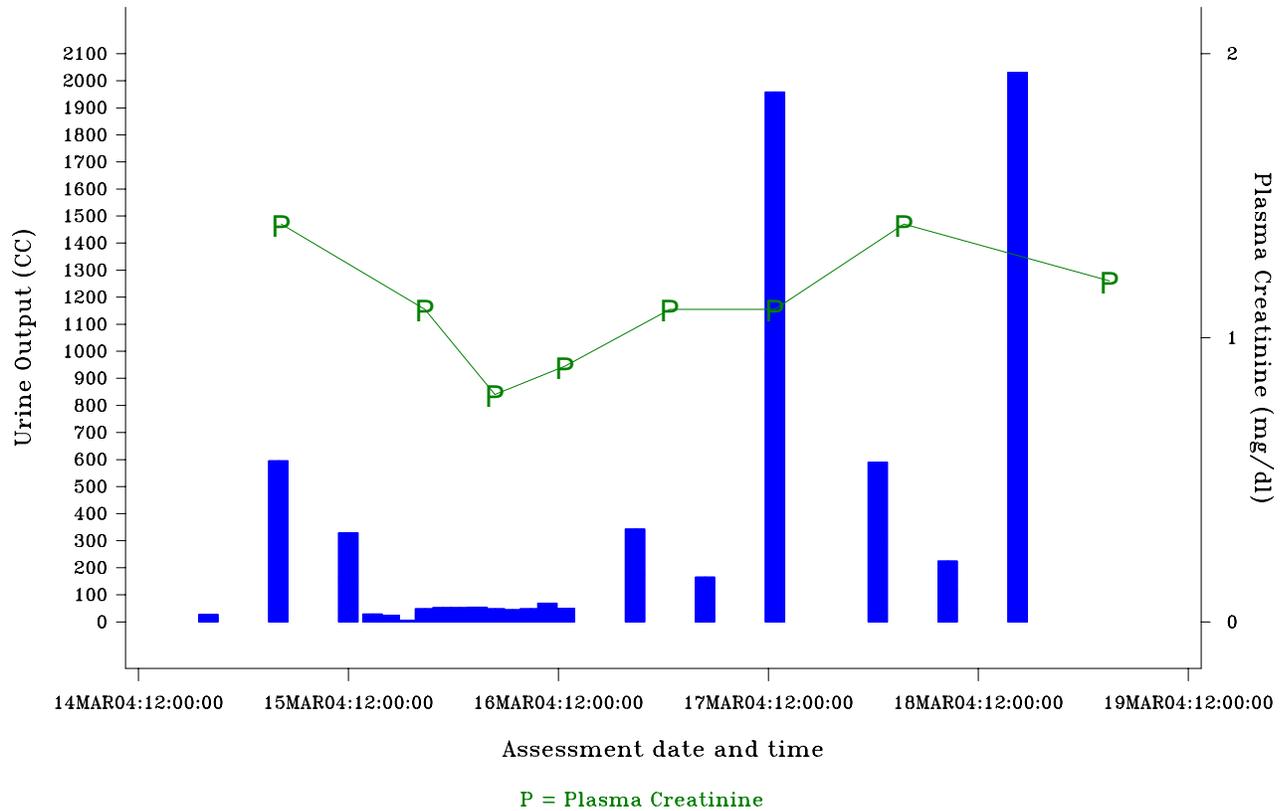
```

proc gplot data=urine;
  plot urine1*asdt/ noframe nolegend haxis=axis1 vaxis=axis2;
  plot2 labval*asdt / noframe nolegend haxis=axis1 vaxis=axis3;
run;
quit;

```

Figure 1 Creating Bar Charts and Scatter Plots using Data Manipulation

Urine Output, and Creatinine Levels for Patient 1101



2. SAS ANNOTATE

The alternative approach in creating bar charts is by employing the ANNOTATE data set to customize the graphical display in SAS GPLOT. In an ANNOTATE data set, each observation represents a command to draw a graphic element or to perform an action. The SAS ANNOTATE macros can be used within a DATA step to simplify the process of creating ANNOTATE observations (SAS Institute, 2002). The BAR annotate macro can easily create ANNOTATE observations that draws a bar chart in GPLOT.

The following table shows the structure of an annotate data set that is required for the BAR annotate macro.

<i>Summary of Annotate Variables</i>		
Task Group	Variable	Description
Variable that defines an action	FUNCTION	specifies a drawing or programming action.
Positioning variables that determine coordinate values	X	specifies a numeric horizontal coordinate
	Y	specifies a numeric vertical coordinate
Positioning variables that specify coordinate systems	XSYS	specifies coordinate system for X or XC coordinates
	YSYS	specifies coordinate system for Y or YC coordinates
Attribute variables	COLOR	color of a graphics primitive
	LINE	line type to use in drawing or special control over pies and bars
	STYLE	font or pattern for a graphics element, depends on the FUNCTION variable

The ANNOMAC macro needs to be first compiled (%ANNOMAC) before any ANNOTATE macros can be accessed in SAS. The following DATA step shows how an ANNOTATE data set can be easily created by using the BAR macro, %BAR (x1, y1, x2, y2, color, line, style), in the ANNOTATE facility.

```
%annomac ;

data annot;
  set urine__;
  by site randno__ asdt;
  if urine__>. ;
  retain xsys ysys '2';
  %bar(asdt-3600,0,asdt+3600,urine__,green,0,solid);
  format x datetimel6.;
run;
```

By using the BAR annotate macro, the values for the ANNOTATE variables are automatically created. Here is a sample display of the ANNOTATE data set:

Function	X	Y	Xsys	Ysys	Color	Line	Style
MOVE	14MAR04:19:00:00	0	2	2	Green		
BAR	14MAR04:21:00:00	27	2	2	Green	0	Solid
MOVE	15MAR04:03:00:00	0	2	2	Green		
BAR	15MAR04:05:00:00	595	2	2	Green	0	Solid
MOVE	15MAR04:11:00:00	0	2	2	Green		

Function:

- The MOVE function in the ANNOTATE facility will move the starting point to the specified (X, Y) values.
- The BAR function will draw a green bar from the current (X, Y) position.

X:

- Created by parameters X1 and X2.
- X1 = Assessment Date/Time – 3600 seconds (1 hour)
- X2 = Assessment Date/Time + 3600 seconds.

Y:

- Created by parameters Y1 and Y2.
- Y1 = 0
- Y2 = Urine output value

Xsys/Ysys:

- The value 2 references the minimum to maximum of the data values on the X-axis/Y-axis as the area for drawing the graphics elements.

Color:

- The color of the bars will be green.

Line:

- The value 0 draws an outline all around the bar.

Style:

- Solid will fill the bar with the color specified.

CREATING A SCATTER-PLOT SUPERIMPOSED ON THE BAR-CHART: A SECOND PLOT STATEMENT AND THE ANNOTATE= OPTION

By employing the ANNOTATE= option in the first PLOT statement, the ANNOTATE data set ANNOT will be used to create bar chart for the urine output data. The second PLOT statement (PLOT2) will display the scatter plot of plasma creatinine. Figure 2 shows the graphical display using SAS ANNOTATE in PROC GPLOT.

```

/*Symbol statement for the bar chart.*/
Symbol1 v=none interpol=none cv=green;

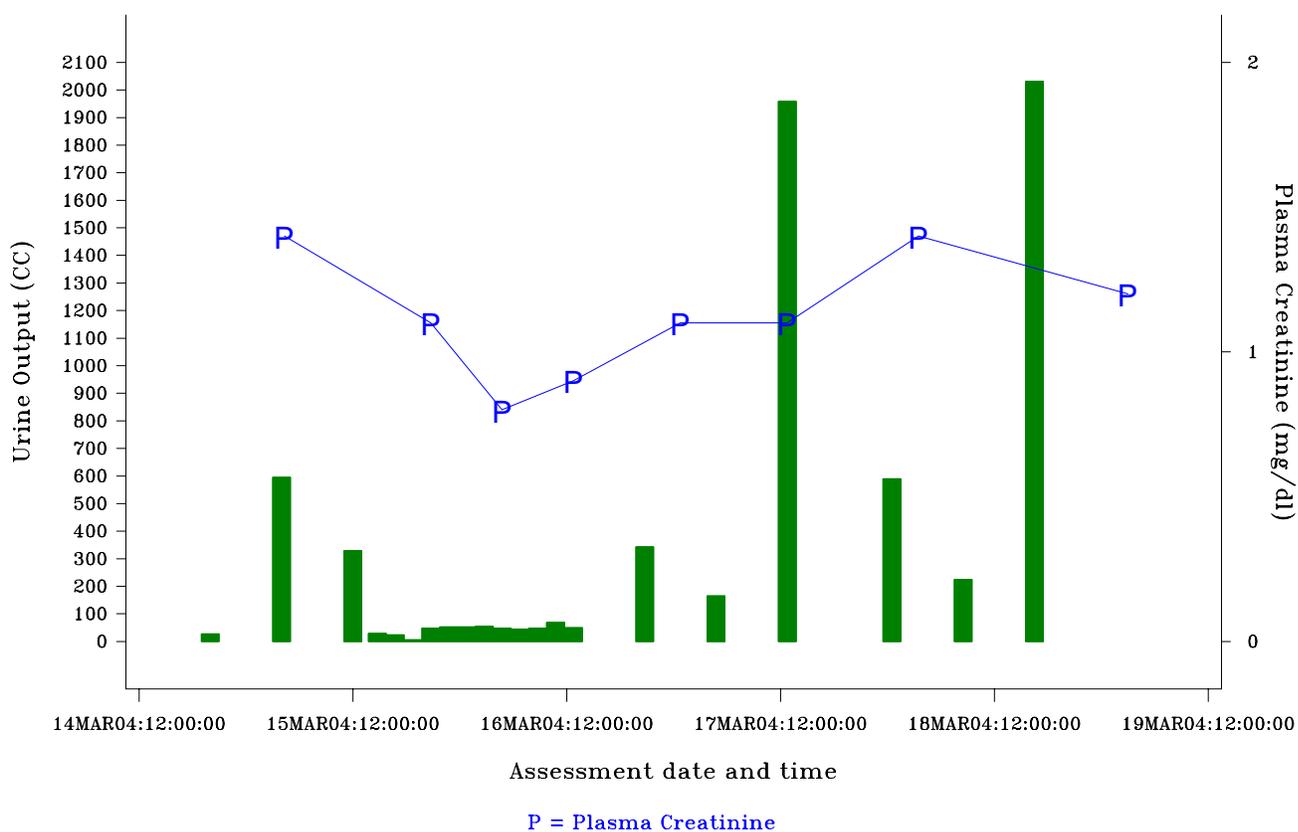
/*Symbol statement for the scatter plot.*/
symbol2 h=0.17 in v='P' color=blue cv=blue interpol=join l=1;

proc gplot data=urine__;
plot urine__*asdt/ annotate=annot overlay noframe nolegend haxis=axis1
vaxis=axis2;
plot2 labval*asdt / noframe nolegend haxis=axis1 vaxis=axis3;
run;
quit;

```

Figure 2 Creating Bar Charts and Scatter Plots using SAS ANNOTATE

Urine Output, and Creatinine Levels for Patient 1101

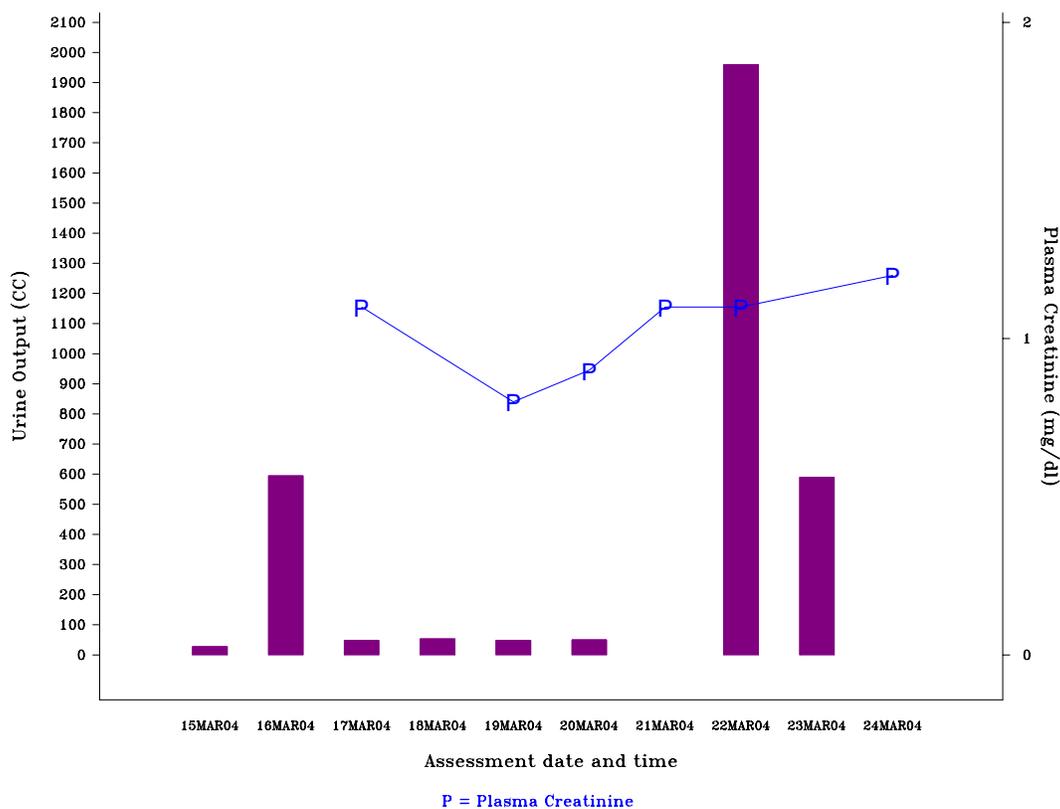


A COMPARISON: USING GPLOT OR GBARLINE?

The answer to the question is: **it all depends on your data**. The GBARLINE procedure does not perform well in analyzing the trend of our safety data. However it provides an effective one-step solution to **data with limited groupings and same intervals on the X-axis**. To illustrate this point, the GBARLINE procedure will be rerun using the following revised data and displayed in figure 4.

Site	Randno__	Asdt	Labval	Urine__
3	1101	15MAR2004		27
3	1101	16MAR2004		595
3	1101	17MAR2004	1.1	48
3	1101	18MAR2004		53
3	1101	19MAR2004	0.8	48
3	1101	20MAR2004	0.9	50
3	1101	21MAR2004	1.1	
3	1101	22MAR2004	1.1	1959
3	1101	23MAR2004		590
3	1101	24MAR2004	1.2	

Figure 4 Creating Bar Charts and Scatter Plots using PROC GBARLINE
Data with Limited Distinct Groupings and Same Intervals on the X-axis
Urine Output, and Creatinine Levels for Patient 1101



CONCLUSION

This paper has shown how a bar chart and a scatter plot can be created on the same graphical output using DATA MANIPULATION and SAS ANNOTATE in PROC GPLOT. The SAS version 9 PROC GBARLINE provides a viable alternative for accomplishing the same task for small distinct groupings of data that are evenly-spaced on the X-axis. For data that are clustered or irregularly spaced, the GPLOT procedure produces a more meaningful and interpretable graphical display.

REFERENCE

SAS Institute (2005), What's New in SAS GRAPH (V9 and Beyond). A presentation offered by the SAS Institute at the 30th Annual Conference of the SAS Users Group International, Philadelphia, PA.

SAS Institute (2002), SAS/GRAPH Software: Reference Volume 1, Version 8, Cary, NC: SAS Institute Inc.

SAS Institute (2002), SAS/GRAPH Software: Reference Volume 2 Version 8, Cary, NC: SAS Institute Inc.

Tsang, Rita (2005), Creating Bar Charts and Scatter Plots Using the SAS GPLOT Procedure – Both on the Same Page? It is Easier Than You Think! A paper presented at the 30th Annual Conference of the SAS Users Group International, Philadelphia, PA.

Tsang, Rita (2005), Creating Bar Charts and Scatter Plots on the Same Page, Deciphering SAS ANNOTATE in PROC GPLOT. A paper presented at the Annual Conference of the North East SAS Users Group 2005, Portland, ME.

ACKNOWLEDGEMENT

The author would like to express her appreciation to the following individual for her invaluable comments and suggestions in this paper:

Pat Wesolowski, Averion Inc.

CONTACT INFORMATION

The author may be contacted at:

Rita Tsang
Averion Inc.
4 California Avenue
Framingham, MA 01701
(508)416-2694
Fax: (508)416-2797
E-mail: rtsang@averioninc.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.