PhUSE US Connect 2018

Paper CT07

# Controlling the Drawing Space in ODS Graphics by Example

Max Cherny, GlaxoSmithKline, Collegeville, PA

## ABSTRACT

The SG annotation facility is a very powerful tool within ODS graphics. It is used to create or enhance ODS SAS graphics by placing graphical elements such as lines or symbols anywhere on a screen. However, with this power comes a certain degree of complexity, which makes learning how to annotate ODS graphics a difficult task.

One of the most challenging aspects of using the SG annotation dataset correctly is the control of drawing space. There are four different drawing areas and three different units which control placement of a graphical element. A SAS user needs to understand how these areas and units work together to identify the exact location of a coordinate in order to judiciously use the system.

This paper explains this aspect of the SG annotation dataset and provides easy and reusable examples for controlling the drawing space. These examples can function either as stand-alone programs or as the starting point for further experimentation with the SG annotation facility. These programs can also serve as the basic general introduction to the ODS graphics.

## INTRODUCTION

The ODS Statistical Graphics Procedures is a very powerful feature of SAS that allows SAS users to create different types of graphs. However, these procedures are not always sufficient enough to draw complex graphs. The SG Annotation facility can be used to enhance the output of ODS procedures or to create entire plots.

The SG annotation allows plotting of any graphical elements (i.e. lines, labels, symbols) by manually specifying their coordinates on the drawing area. Coordinates for such elements can be expressed as X and Y if they are numeric and XC and YC if they are characters. Depending on a type of the graphical element, its exact positions can also be expressed by using such variables as ANCHOR, ROTATE or WIDTH.

Coordinates are translated to a position on the drawing area according to the provided drawing spaces. There are four different drawing areas used by the SG Annotation facility: Graph, Layout, Wall and Data. They are used to place a graphical element relative to the position of another graphical element. Annotations are produced by applying one of the three units: Percentage, Pixels and Data.

Drawing space value is used to represent the combination of drawing space and units. For example, DATAPERCENT value is using PERCENT unit in DATA drawing area. There are currently 9 possible drawing space values because DATA VALUE units can be used with DATA area only.

| | UNIT | | |
|---|---|---|---|
| **DRAWING SPACE** | **DATA VALUE** | **PERCENTAGE** | **PIXELS** |
| **DATA** | DATAVALUE | DATAPERCENT | DATAPIXEL |
| **WALL** | | WALLPERCENT | WALLPIXEL |
| **LAYOUT** | | LAYOUTPERCENT | LAYOUTPIXEL |
| **GRAPH** | | GRAPHPERCENT | GRAPHPIXEL |

**Display 1. Drawing Space Values**

Drawing space values are set for the following annotation variables: DRAWSPACE, X1SPACE, X2SPACE, Y1SPACE and Y2SPACE. These annotation variables control the position of text labels, lines, rectangles, arrows and other graphical shapes used in SG annotation.

## BRIEF OVERVIEW OF THE SG ANNOTATION FACILITY

The annotation facility uses commands to place graphical elements on the drawing area. These commands are contained in an annotation data set, which is then invoked during the execution of ODS procedures such as SGPLOT

or SGPANEL.

Every observation in the annotation data set indicates which graphical element is drawn, as well as its position and characteristics. There are many functions and variables available in the facility. The most common annotation variables are:

| | |
|---|---|
| X and Y | Indicate x and y positions of a graphic element |
| DRAWSPACE | Indicates drawing space |
| TEXT | Text of a label |
| SIZE | Indicates size of a graphic element |
| FUNCTION | Indicates which type of a graphic element to draw |

The most frequently used annotation functions include:

| | |
|---|---|
| LABEL | Draws text on the drawing area |
| LINE | Connects two points with a line. Each point has x and y coordinates. |

These variables and functions can be used to create or enhance most graphs. They can be specified in a data set by using explicit statements such as:

```
function ="LABEL";
x1=10;
y1=20;
text="TEXT STRING";
```

The same function can be coded by using SG annotation macros as follows:

```
%sganno;
%sgtext(x1=10,y1=20, label="TEXT STRING");
```

These statements place text at position where X is 10 and Y as 20. Note that using annotation macros requires invoking %SGANNO macro.

The most common SG annotation unit is Data value. This unit directly corresponds to the plotted data and thus can be easily used to enhance existing plots. The Percentage unit is independent of the plotted data and corresponds to the percentage of the drawing area. This unit allows placing of a graphical element anywhere in the drawing area or even outside of it. The Pixel value unit is used independently of the plotted data or drawing area and uses pixel values to render annotations anywhere.

## DATA AREA

The Data Area is the drawing area of the plot itself. Figure 1 displays a simple scatter plot with grouped data. The shaded area represents data drawing space.

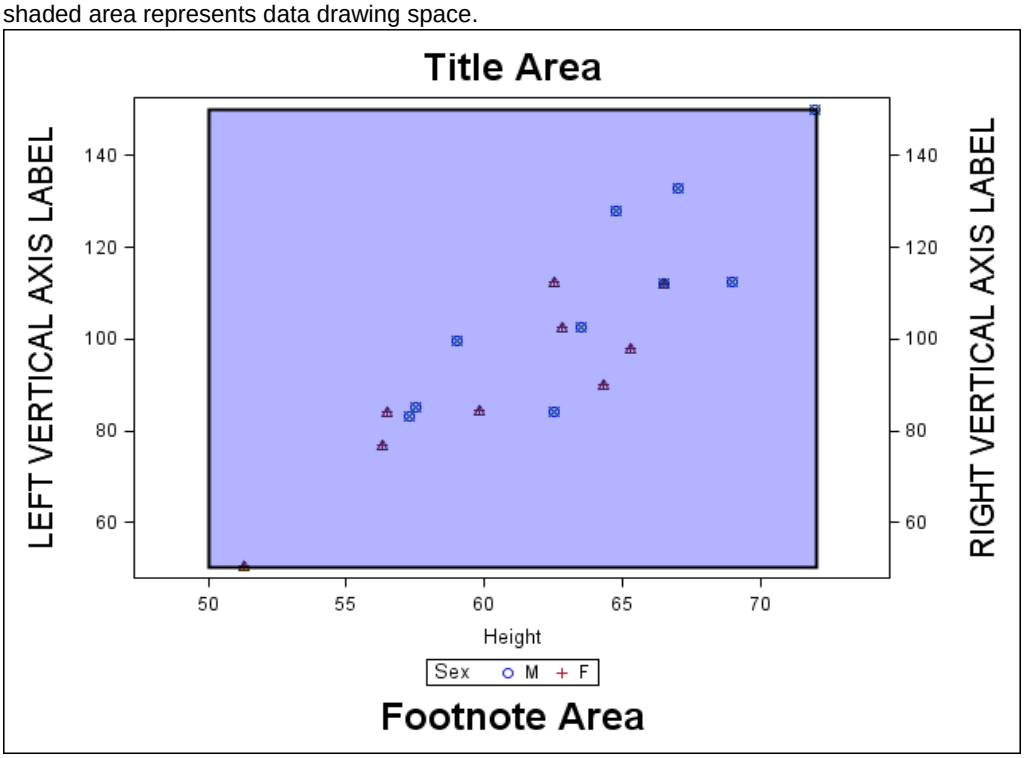


**Figure 1: Data Area**

Data area is the most useful drawing area since most annotations are plotted within the axis. The combination of data area and percentage unit is one of the easiest systems to use as it refers to data area and uses percentages of the axis within the area to establish the position of a graphical element. The value of DRAWSPACE variable for such system is DATAPERCENT.

The following example illustrates how to use such system. The test data set below represents change from the baseline in BMI over time.

```
  /* create test data set */
  data bmi_data;
    length Drug $1;
    input Drug $1 Month 3 Mean;
    datalines;
A 1 0
B 1 0
A 2 -12.5
B 2 -4.08
A 3 -19.4
B 3 -4.15
A 4 -22.5 7
B 4 -0.15
  ;
  run;
```

The following code creates an annotation data set containing commands to draw “Change in BMI” label:

```
   /* create annotation data set */
  data anno;
    function="TEXT";
    label="Change in BMI";
    anchor="BOTTOMLEFT"; /* controls the starting position of text */
    drawspace="DATAPERCENT";
    textweight="BOLD";
    x1=30;
    y1=70;
    width=50;
  run;
```

The code below creates a plot of BMI over time shown on Figure 2:

```
  proc sgplot data=bmi_data sganno=anno;
    title height=2 "Sample Plot";
    series x=month y=mean/group=drug;
    xaxis offsetmin=0.1 offsetmax=0.1;
    xaxis type=discrete;
    label mean="Mean BMI";
  run;
```

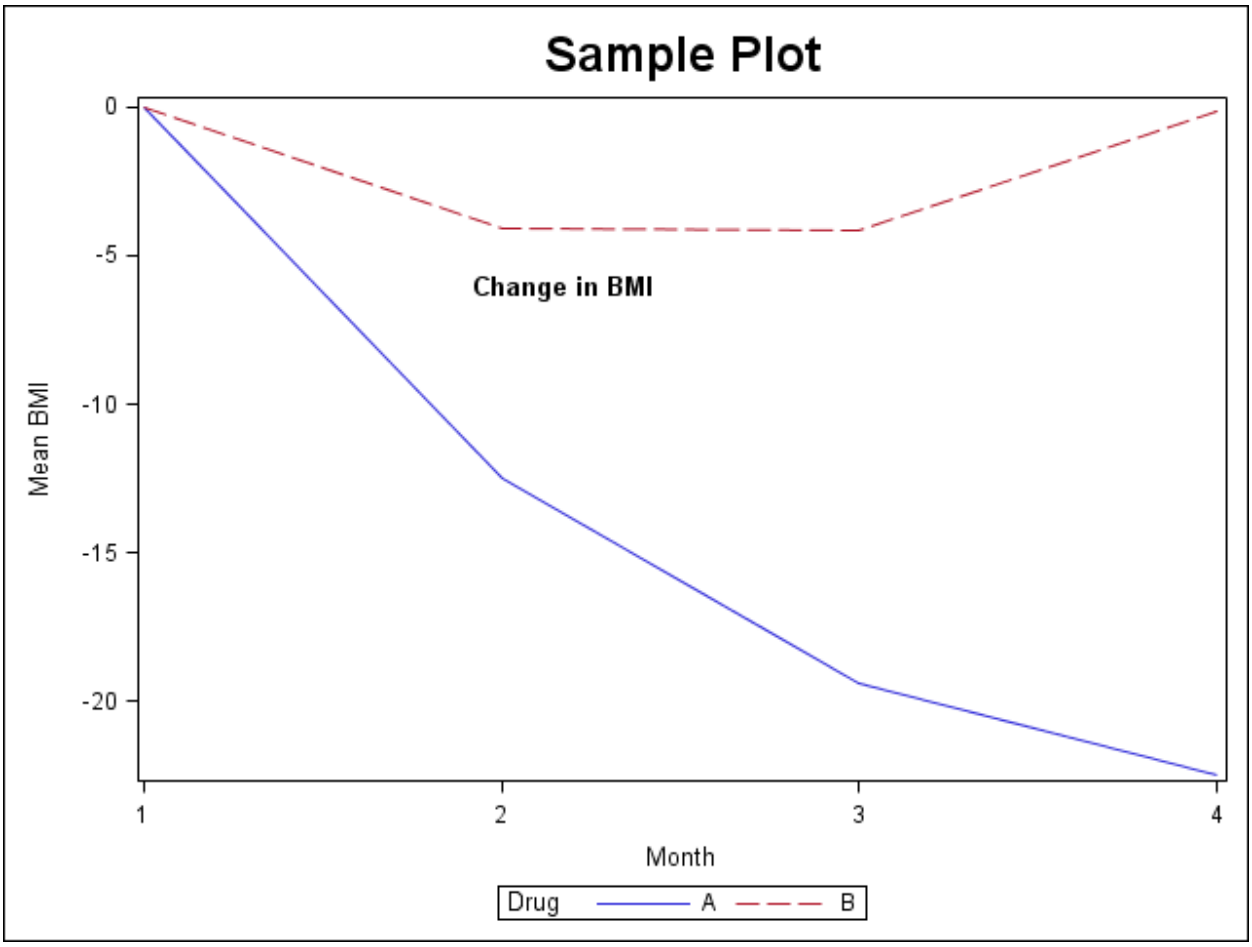


**Figure 2: Example of DATAPERCENT use**

This example demonstrates how using 30% of the horizontal axis and 70% of the vertical axis creates a text label in that position.

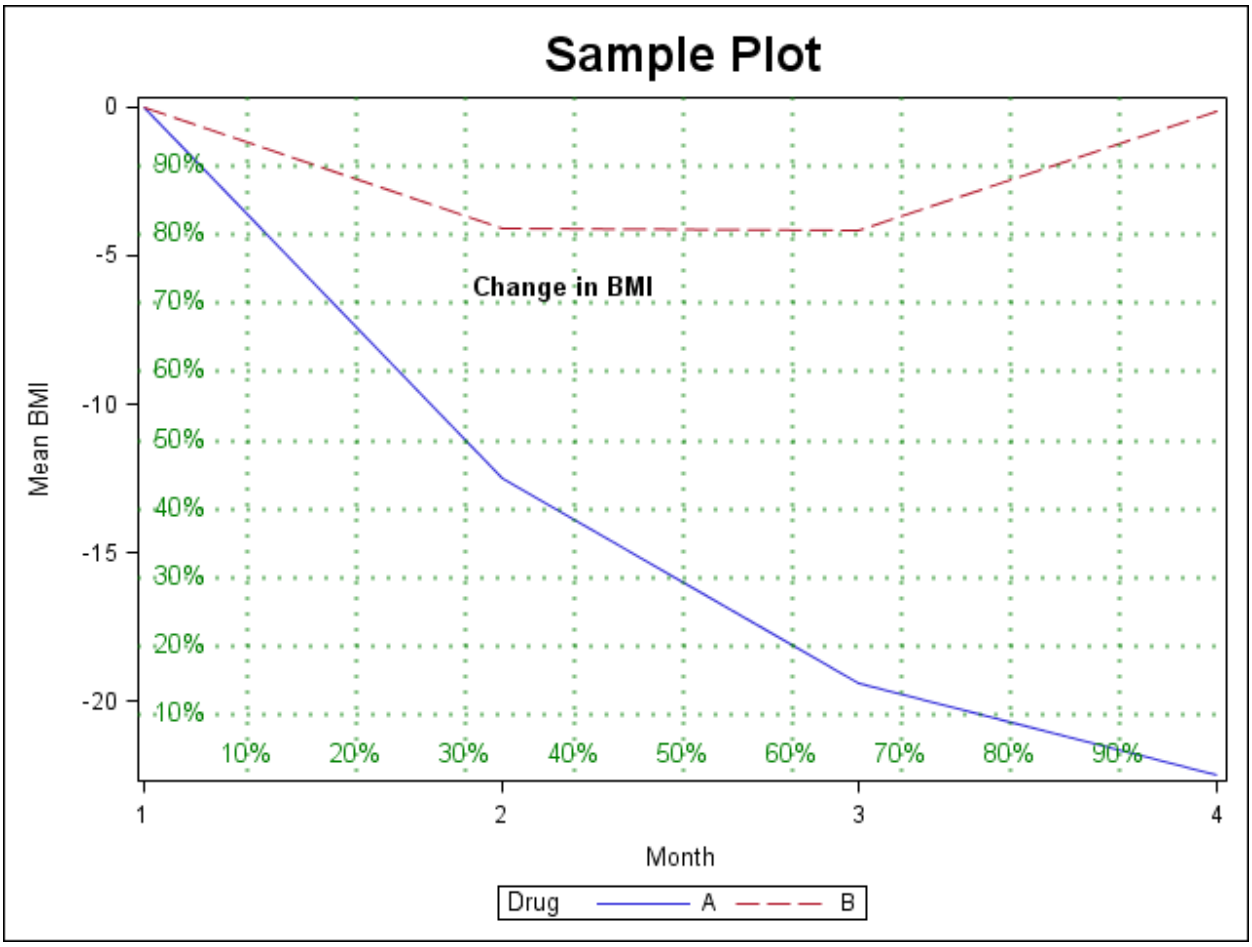


**Figure 3: Example of DATAPERCENT with grid lines**

Figure 3 displays the same plot as Figure 2 with an added grid representing the percentages of data area. The text is left aligned by setting the ANCHOR variable to “BOTTOMLEFT”. The use of ANCHOR variable is critical in setting the exact position of a graphical element.

### WALL AREA

The area that includes the entire space within the axis is called Wall area and is represented on Figure 4 as the shaded area.

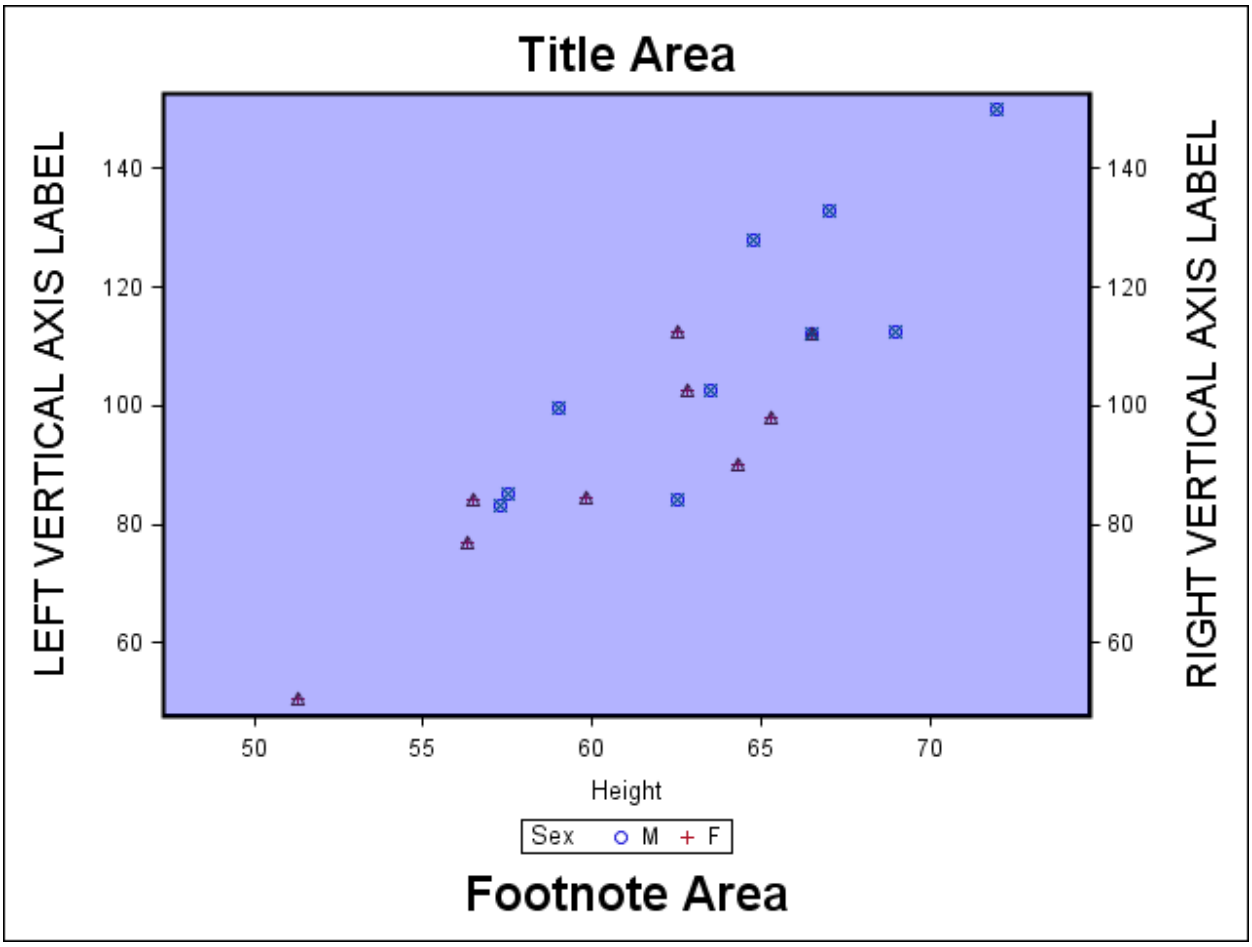


**Figure 4: Wall Area**

It is important to keep in mind that x and y coordinates can each use different data areas and units. The following example illustrates the use of data area and value units for X coordinate and wall area and percent units for Y coordinates.
The program below creates a sample profile plot of oral corticosteroid (OCS) daily doses over time. Each shaded area represents a period when a subject was in remission.

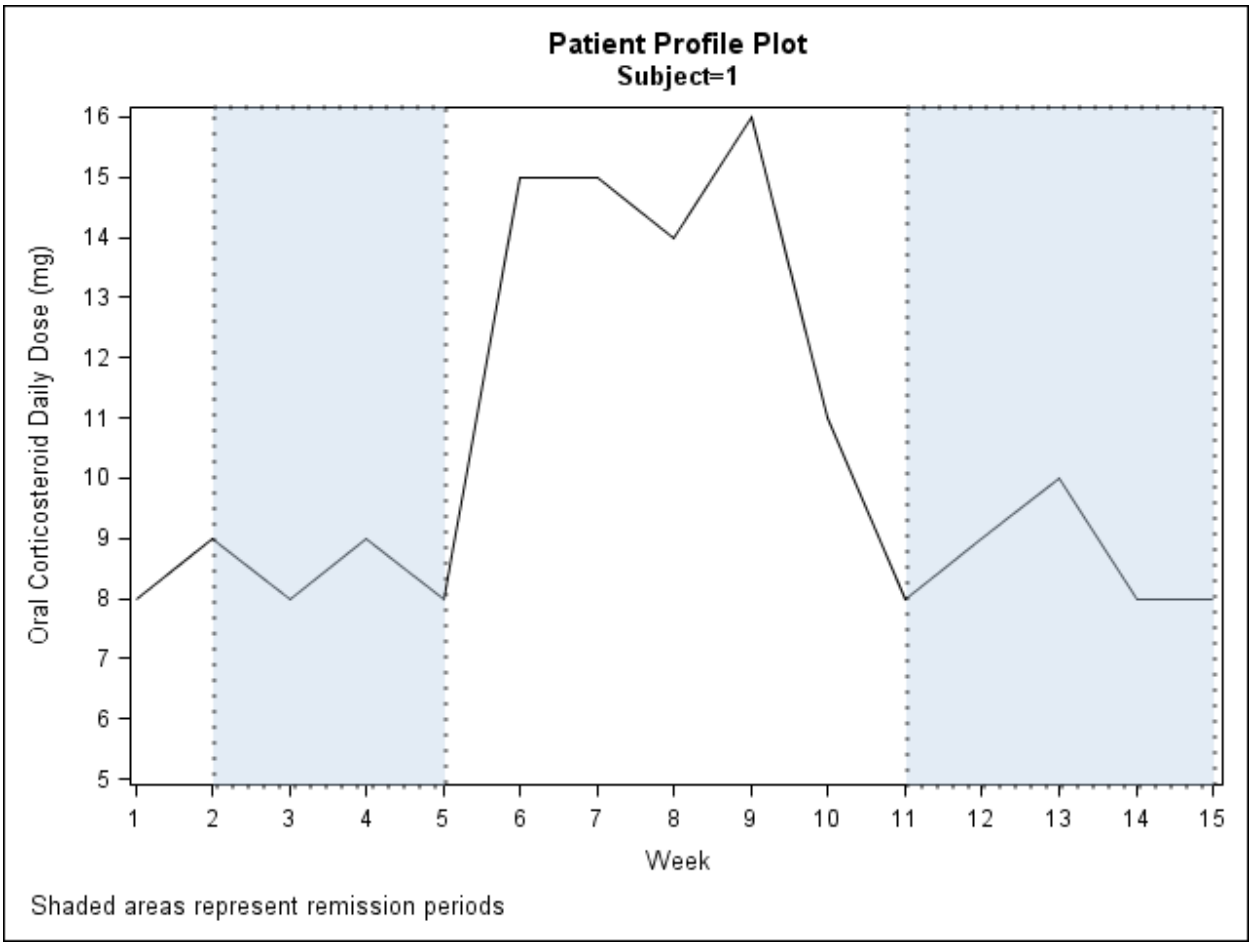


**Figure 5: Example of using Wall data area**

The sample data set is generated by the following program:

```
/* create test data for OCS daily and events. */
data patient_data;
  length relapse asthma hosp $1;
  input usubjid week dose relapse asthma hosp;
datalines;
1 1 8 N N N
1 2 9 N N N
1 3 8 N Y N
1 4 9 N N N
1 5 8 N N N
1 6 15 N N N
1 7 15 Y Y Y
1 8 14 N Y N
1 9 16 Y Y N
1 10 11 N Y N
1 11 8 N N N
1 12 9 N N N
1 13 10 N N N
1 14 8 N N N
1 15 8 N N N
;
  run;

/* dataset for remission time points */
data remission;
  /* start and end represent weeks when remission started and ended */
```

```
    start=2;
    end=5;
    output;
    start=11;
    end=15;
    output;
  run;
```

The data step below generates the annotation data using RECTANGLE function:

```
  /* use annotation to create a rectangle. Similar rectangle can be created by using
POLYGON statement in PROC SGPLOT or PROC SGPANEL, however annotation allows for more
customization options */
  data anno;
    set remission;
    function="RECTANGLE";
    x1space="DATAVALUE";
    x1=start;
    widthunit="DATA";

    y1space="WALLPERCENT";
    y1=100;
    height=100;

    display="ALL";
    filltransparency=0.6;
    linepattern="THINDOT";

    linecolor="BLACK";
    anchor="TOPLEFT"; /* to control the exact placement of the rectangle */
    width=end-start; /* calculate the width of remission in weeks*/
  run;
```

The width of the shaded area corresponds to remission periods. This is done by setting value of X1SPACE variable to “DATAVALUE”, which means it corresponds to the values on the X axis. It is important to note that WIDTHUNIT variable must be set to DATA for the width of the bar to match the values of X axis. The Y1SPACE is set to WALLPERCENT. Y1 variable is set to 100 and HEIGHT variable to 100 for the rectangle to cover the entire wall area.

Figure 5 is generated by the following SGPLOT procedure:

```
  proc sgplot data=patient_data sganno=anno;
    series x=week y=dose;
    xaxis type=discrete;
    yaxis values = (5 to 16);
    label usubjid="Subject"
    week="Week"
    dose="Oral Corticosteroid Daily Dose (mg)";
    title "Patient Profile Plot";
    by usubjid;
    footnote justify=left "Shaded areas represent remission periods";
  run;
```

## LAYOUT AREA

The Layout drawing area includes most of the graph except for titles and footnotes and is represented on Figure 6

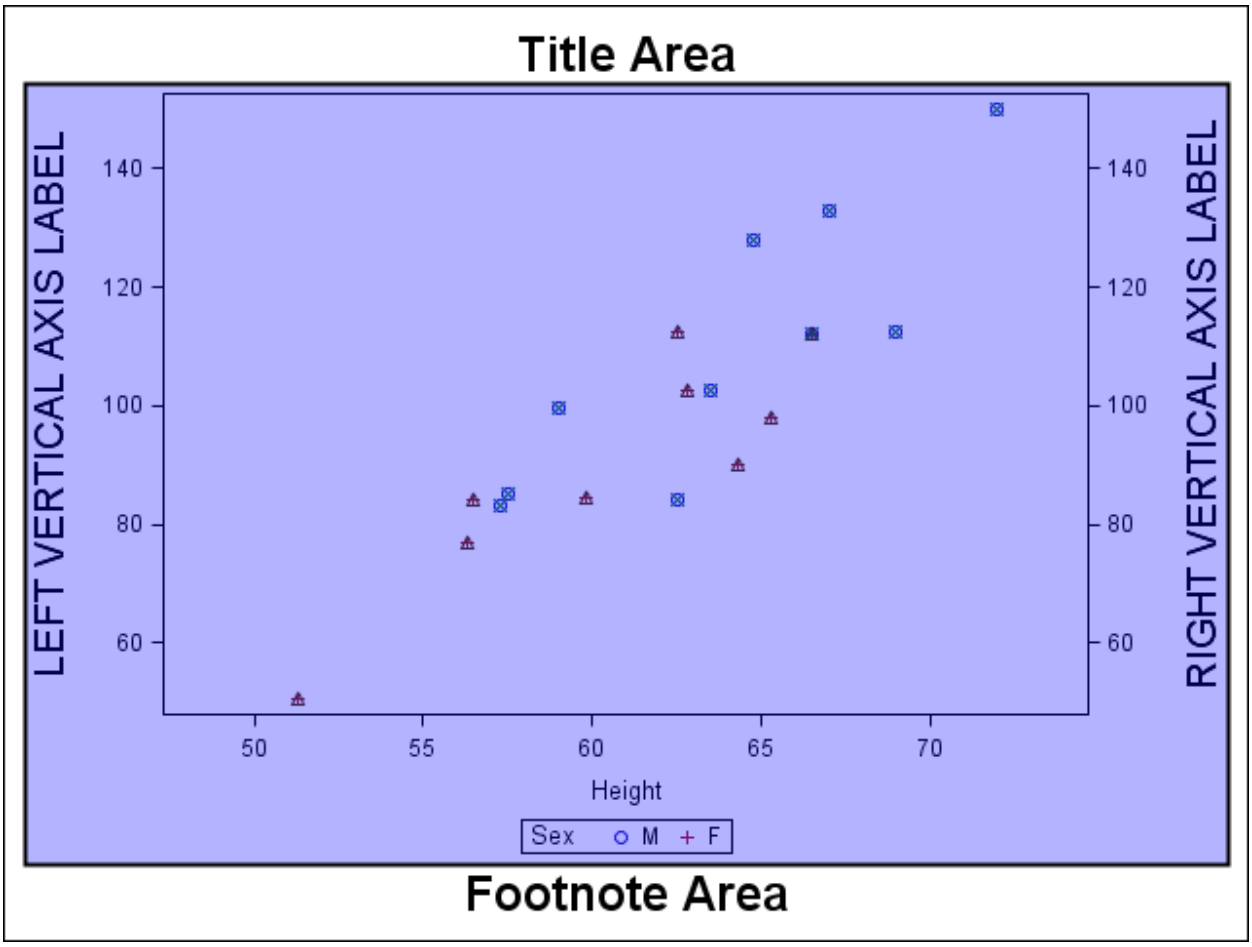


**Figure 6: Layout Area**

This drawing area can be useful for annotations which need to be placed outside of wall area. The following example demonstrates the use of layout area. The program below enhances Figure 5 by adding lines to indicate times of relapse, asthma occurrences and hospitalization. The layout area is used to add text labels to the left of the Y axis.

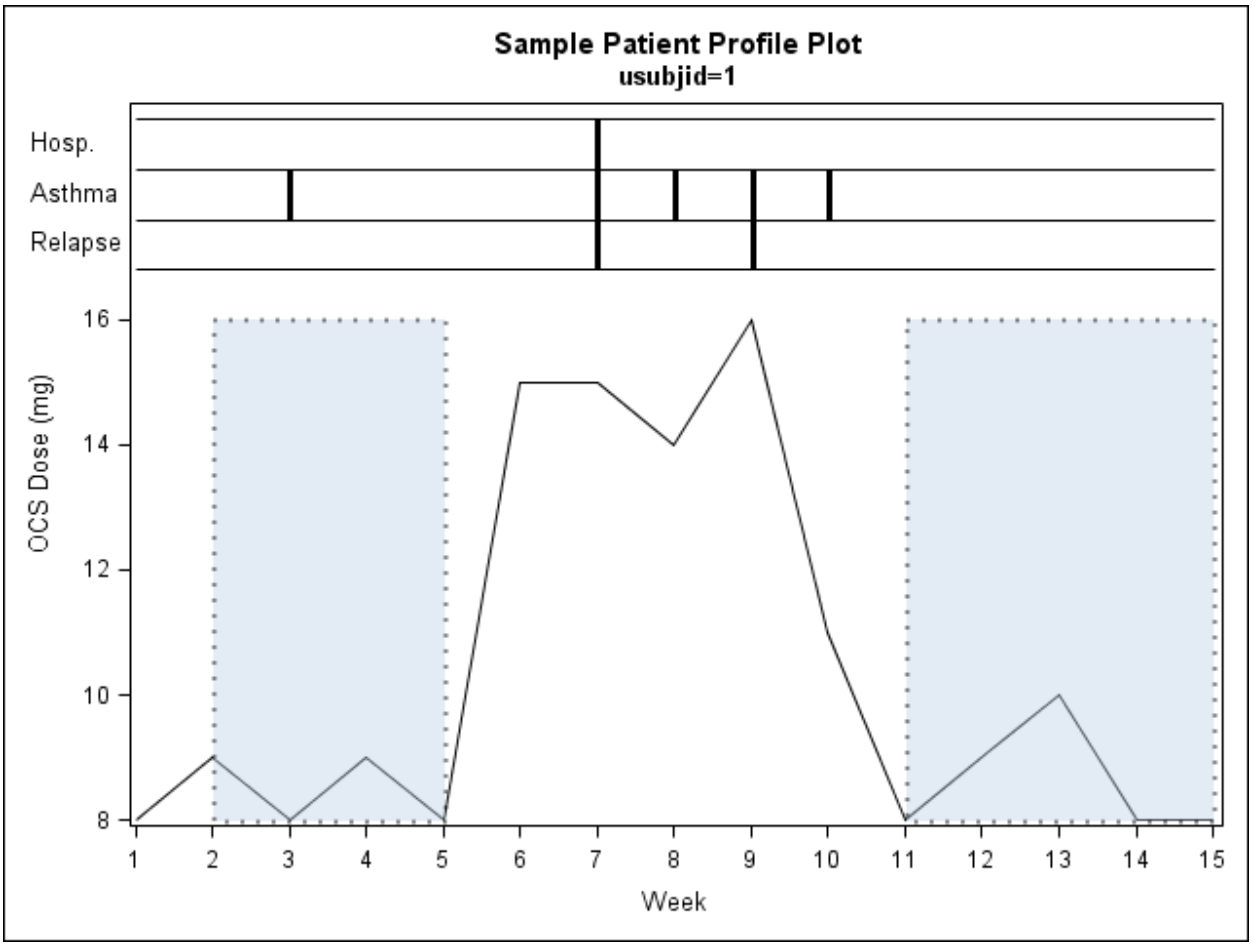


**Figure 7 LAYOUT data area example**

The data steps below generate several annotation data sets:

```
/* create text labels */
data anno_label;
  length anchor label function x1space y1space $15;
  /* add relapse labels */
  function="TEXT";
  label="Relapse";
  anchor="LEFT";
  x1=0;
  y1=115;
  x1space="LAYOUTPERCENT";
  y1space="DATAPERCENT";
  output;
  ;

  /* add asthma labels */
  label="Asthma";
  y1=125;
  output;
  ;

  /* add hosp. labels */
  label="Hosp.";
  y1=135;
  output;
  ;
run;
```

```
/* create remission rectangles */
data anno_remission;
  set remission;
  function="RECTANGLE";
  x1space="DATAVALUE";
  y1space="DATAPERCENT";
  x1=start;
  y1=100;
  widthunit="DATA";
  height=100;
  display="ALL";
  filltransparency=0.6;
  linepattern="THINDOT";
  linecolor="BLACK";
  anchor="TOPLEFT";
  width=end-start;
run;

/* create lines for asthma, relapse and hosp. */
data anno_lines;
  set patient_data;
  length function x1space x2space y1space y2space $15;
  function="LINE";
  x1space="DATAVALUE"; /* can be replaced by DRAWSPACE variable*/
  x2space="DATAVALUE";
  y1space="DATAPERCENT";
  y2space="DATAPERCENT";
  linethickness=3;
  x1=week;
  x2=week;

  /* draw relapse lines */
  if relapse="Y" then do;
    y1=110;
    y2=120;
    output;
  end;

  /* draw Asthma lines */
  if Asthma="Y" then do;
    y1=120;
    y2=130;
    output;
  end;

  /* draw hospitalization lines */
  if hosp="Y" then do;
    y1=140;
    y2=130;
    output;
  end;
run;

/* draw rectangles around relapse, asthma and hosp. lines */
data anno_rectangle_lines;
  length function x1space y1space $15;
  function="LINE";
  x1space="DATAPERCENT";
  x2space="DATAPERCENT";
```

```
    y1space="DATAPERCENT";
    y2space="DATAPERCENT";
    x1=0;
    x2=100;
    linethickness=1;

    y1=110;
    y2=110;
    output;

    y1=120;
    y2=120;
    output;

    y1=130;
    y2=130;
    output;

    y1=140;
    y2=140;
    output;
  run;

  /* combine all annotation datasets together */
  data anno;
    set anno_label
        anno_lines
        anno_rectangle_lines
        anno_remission;
  run;
```

Figure 7 is generated by the following SGPLOT procedure:

```
  proc sgplot data=patient_data sganno=anno;
    series x=week y=dose;
    xaxis type=discrete;
    yaxis values =(5 to 20);
    label week="Week"
    dose="OCS Dose (mg)";
    title "Sample Patient Profile Plot";
    yaxis offsetmax=0.3;
    by usubjid;
  run;
```

'Asthma', 'Relapse' and 'Hosp.' labels are placed outside of the plot. Their X1 coordinates are to set to 0. Since the drawing area is set as LAYOUT, the labels are placed in the same starting position as the Y axis label (i.e. 0% of the data area). The Y1 coordinates of labels are set to 115%, 125% and 125%, respectively for the labels to appear above the plot. The lines representing weeks of asthma, hospitalization and relapse are drawn by setting the X1 coordinates to values of X axis using data area and value units corresponding to WEEK variable in the input data set.

As can be seen in the previous example where we set y1 coordinate of labels to more than 100% of data area, graphical elements can be placed anywhere on the screen using any values. For example, labels are placed even further to the left of the Y axis in Figure 8 by setting the X1 coordinate to -2%.

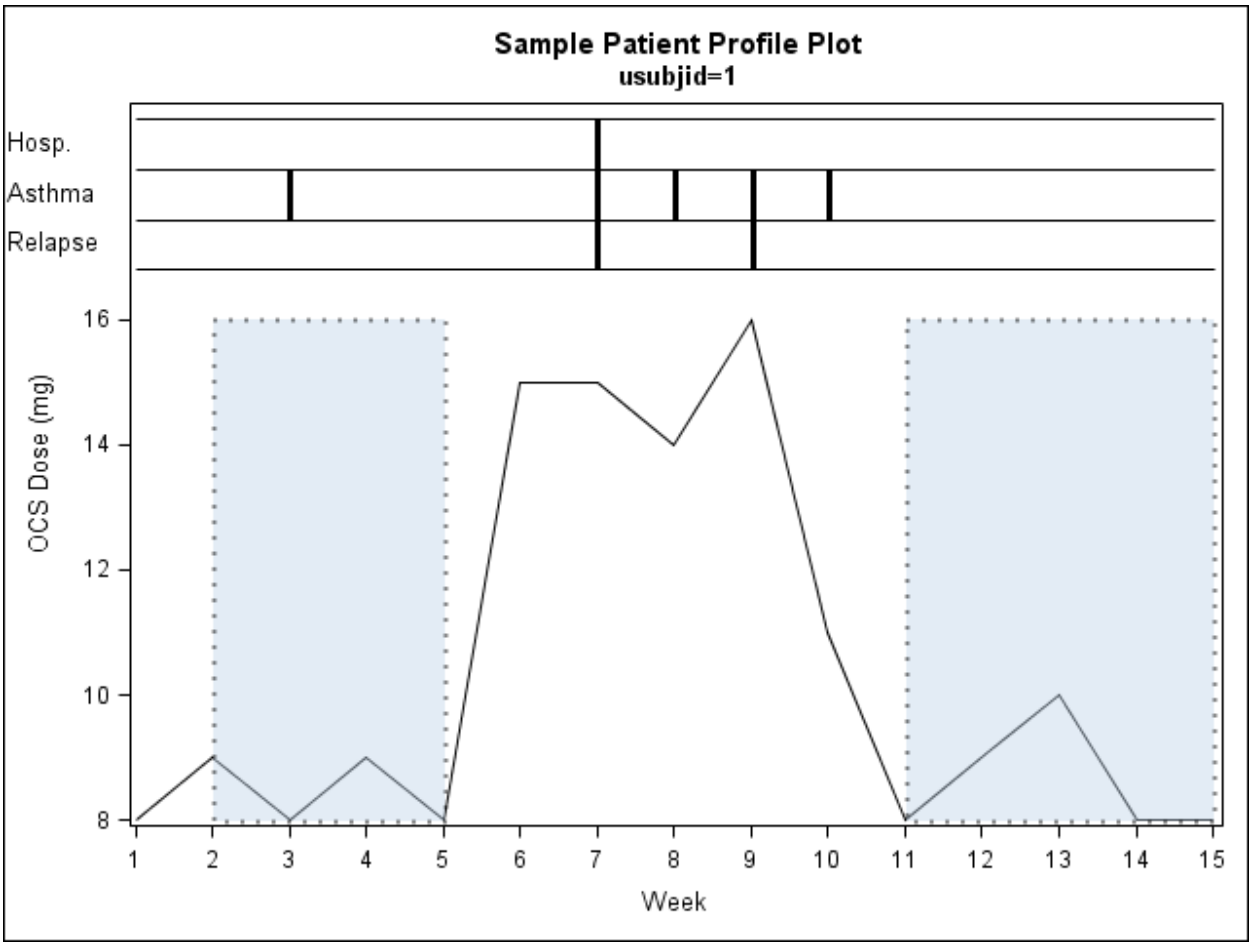


**Figure 8 Moving text labels further to the left by setting x1 to -2%**

```
data anno_label;
    length anchor label function x1space y1space $15;
    /* add relapse labels */
    function="TEXT";
    label="Relapse";
    anchor="LEFT";
    x1=-2; /* -2% is an acceptable coordinate*/
    y1=115;
    x1space="LAYOUTPERCENT";
    y1space="DATAPERCENT";
    output;
    ;
```

It is up to the programmer to make sure graphical elements are correctly rendered on the screen. The x1 coordinate of text labels is set to -5% in Figure 9, which causes the text to be placed off screen; however, no SAS errors or warnings are generated.

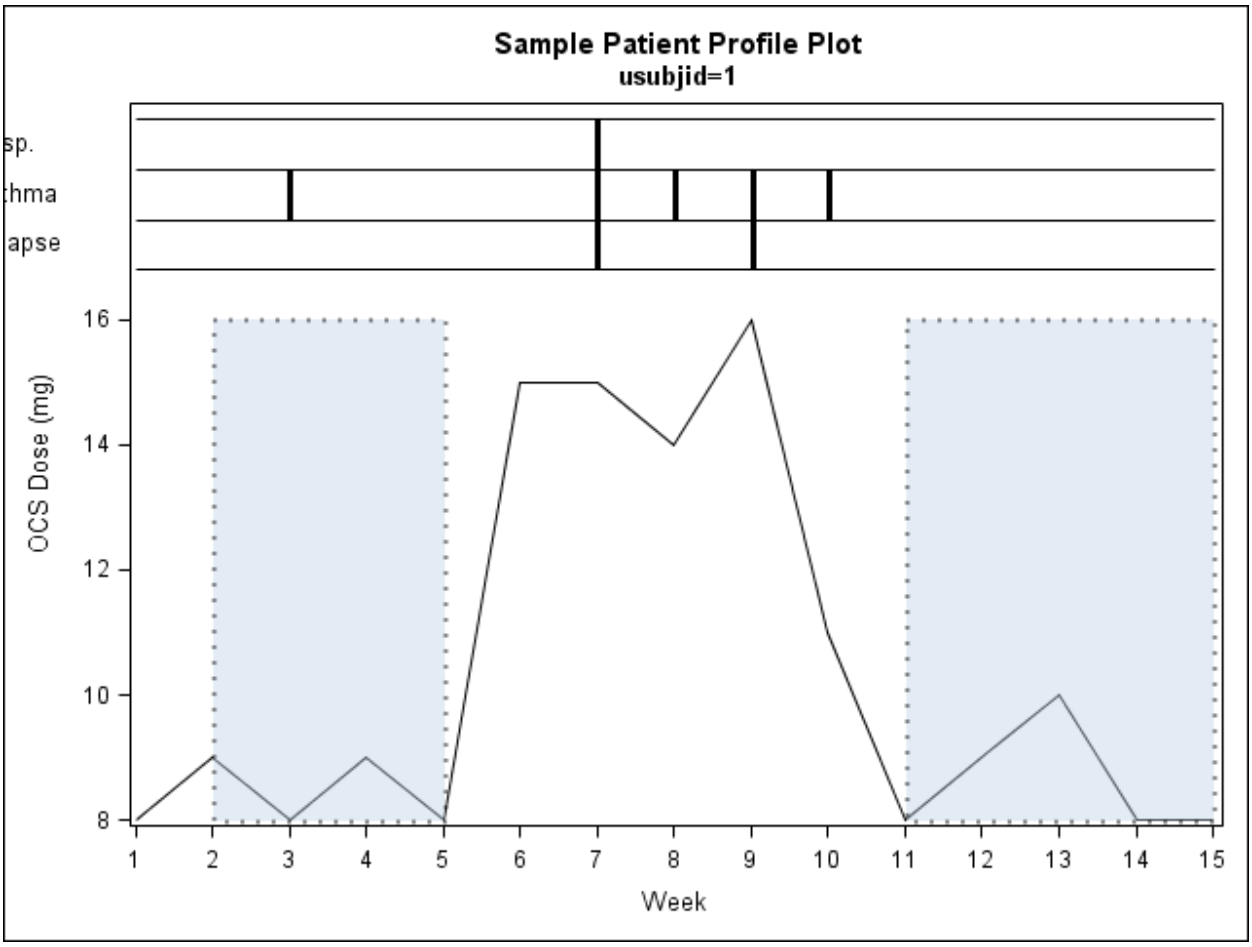


**Figure 9 Incorrect placement of labels**

## GRAPH AREA

The entire drawing area is called the Graph area. It includes the data and the wall areas, as well as areas reserved for footnotes and titles. That is why Figure 10 is completely shaded.

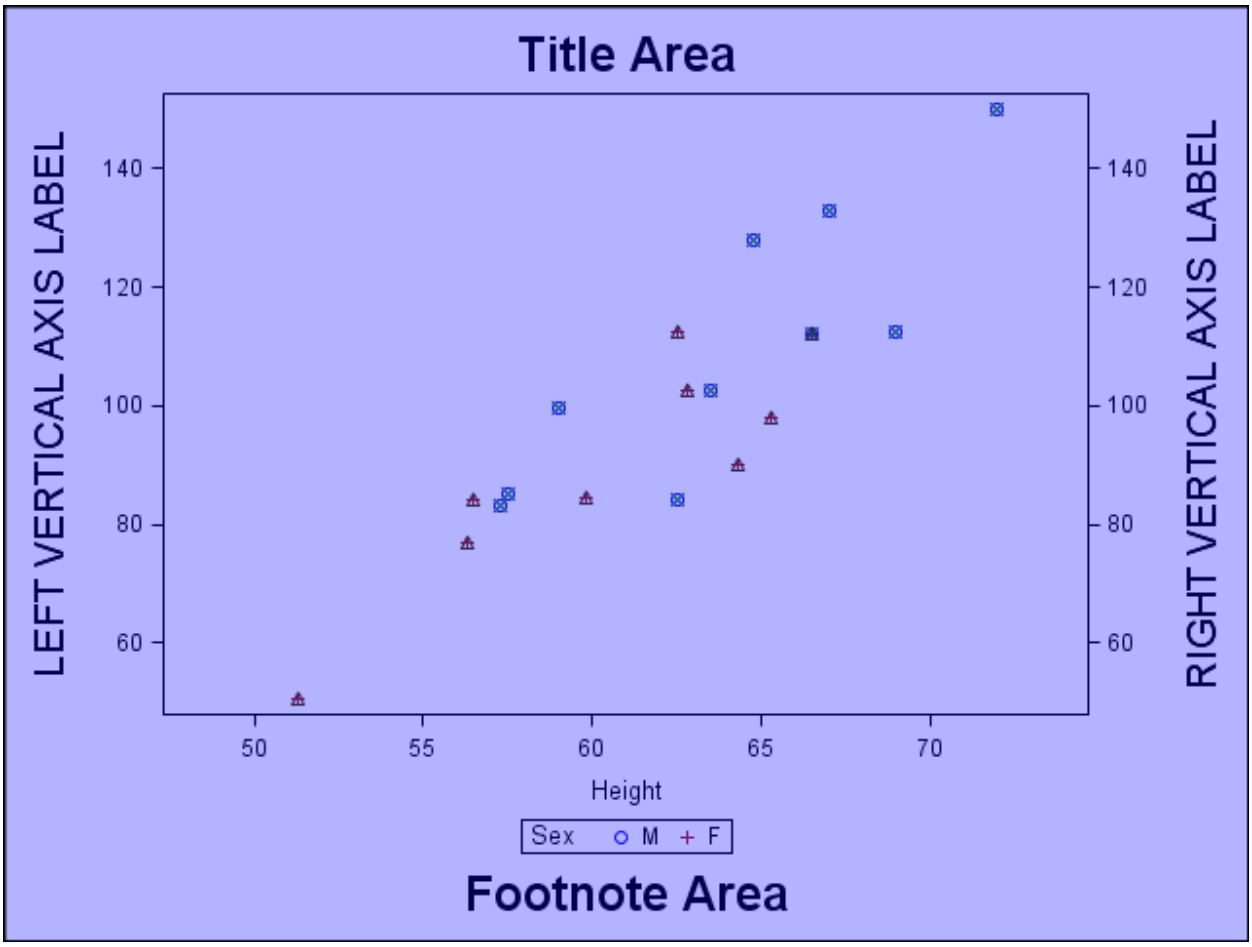


**Figure 10: Graph Area**

The graph area is similar to layout area since it includes most of screen. However, while layout area does not cover the areas beyond footnotes and titles, the graph area covers the entire screen. A graphical element can be placed anywhere in the graph image. For example, by changing the x1 coordinate to 0 and x1space to GRAPHPERCENT, the labels will be placed at the leftmost position of Figure 7.

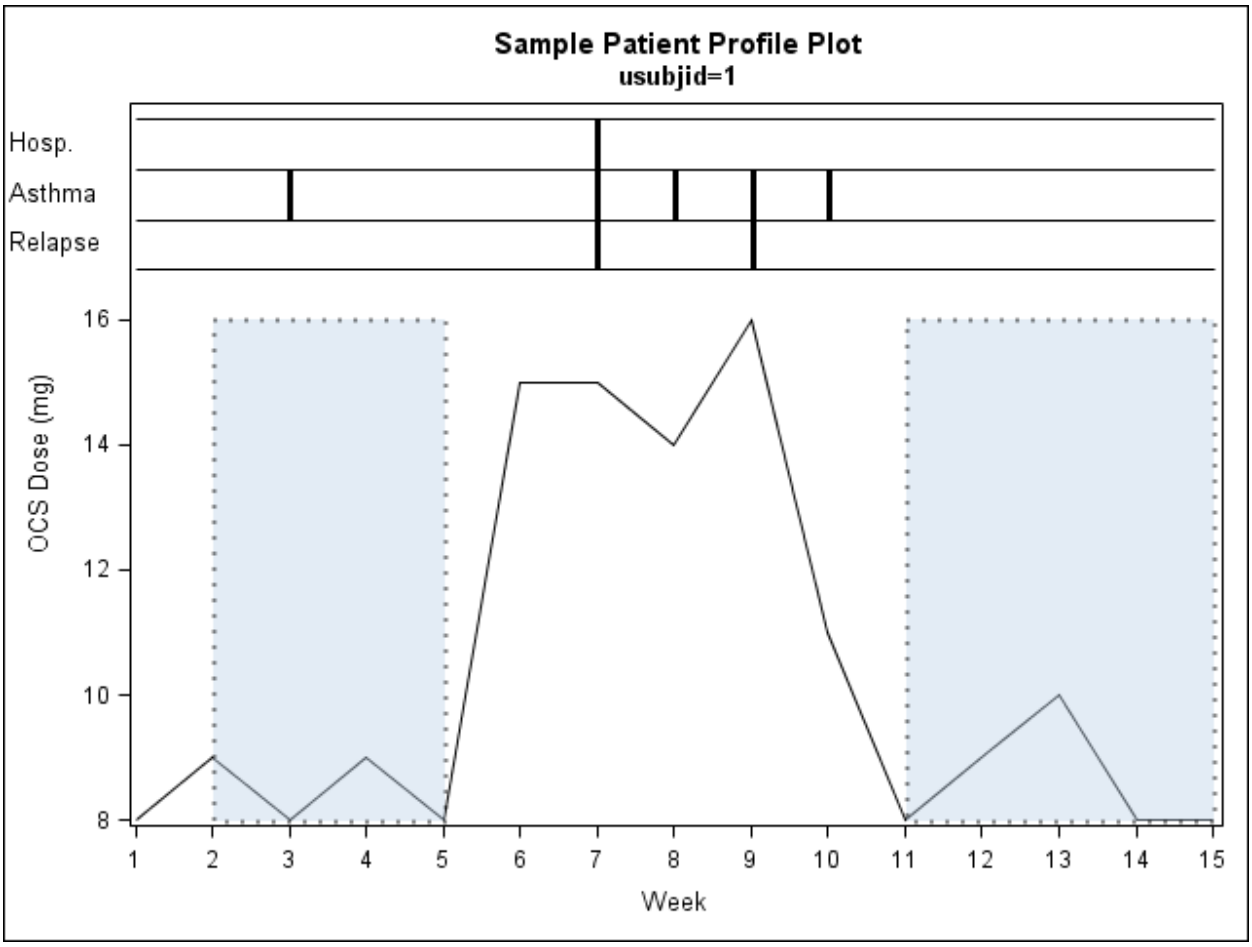


**Figure 11 Example of Graph area**

## LIMITATIONS OF WALL AND GRAPH AREAS

Unfortunately, data and wall draw areas are not available when using SGPANEL and SGCATTER procedures in SAS 9.4. Accordingly, paneled graphs can only be annotated with graph and layout draw areas.

## CONCLUSION

The SG Annotation facility offers a variety of ways to reference coordinates of a graphical element. Using an appropriate drawing space for a given task can make SAS programs more efficient and portable. The most common drawing space is Data area because it uses values of axis and thus allows referencing the plotted data. Wall, layout and graph drawing areas are normally used to plot graphical elements that are independent of the plotted data. Annotations are produced by using pixel, percentage and data units.

## ACKNOWLEDGMENTS

Author would like to thank Jessica Mandel for her help with this paper.

## REFERENCES

Heath, Dan. 2011. “Now You Can Annotate Your Statistical Graphics Procedure Graphs”. SAS Global Forum 2011, Las Vegas, NV https://www.lexjansen.com/wuss/2012/170.pdf

## CONTACT INFORMATION

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

Max Cherny
GlaxoSmithKline
Email: chernym@yahoo.com