PhUSE 2016 - Paper DV04

Clinical Graphs Using SAS

Sanjay Matange, SAS Institute, Inc.

A Brief History of ODS Graphics

ODS Graphics was first released with SAS 9.2

• The feature set focused on the needs for creating automatic graphs from SAS analytical procedures.

• The feature set was enhanced in SAS 9.3 and SAS 9.4 to make graphing easy for the general audience.

• Most clinical graphs are easy to create using SAS 9.4.

ODS Graphics Software

ODS Graphics includes the following components:

- Automatic graphs from SAS analytical procedures.

    No knowledge of graph syntax is required.

- SGRENDER procedure to create graphs using GTL templates.

    Audience – Advanced Graph Programmer

- SG procedures to create graphs with syntax.

    Audience – Graph Programmer

- ODS Graphics Designer to create graphs using interactive tool.

    Audience – Analyst.

All of the above use the Graph Template Language (GTL) to create the graphs, so all graphs produced are consistent in appearance.

SAS
Federal

The Statistical Graphics (SG) Procedures

The SG Procedures provide an easy to use procedure type syntax for using the GTL features.

- SGPLOT procedure for single-cell graphs.
- SGPANEL procedure for Classification Panels.
- SGSCATTER procedure for comparative scatter plots.

In this presentation, We will primarily focus on the use of the SGPLOT procedure to create clinical graphs.

Structure of a Graph

A SGPLOT graph has the following features:

- Zero or more titles and or footnotes.

- One region in the middle that is used to display the data.

- One or more plots that are used to display the data.

- A set of axes that are shared by the plots in the cell.

- Zero or more legends and or Insets.

In this graph the axis labels are colored to match the data they represent.
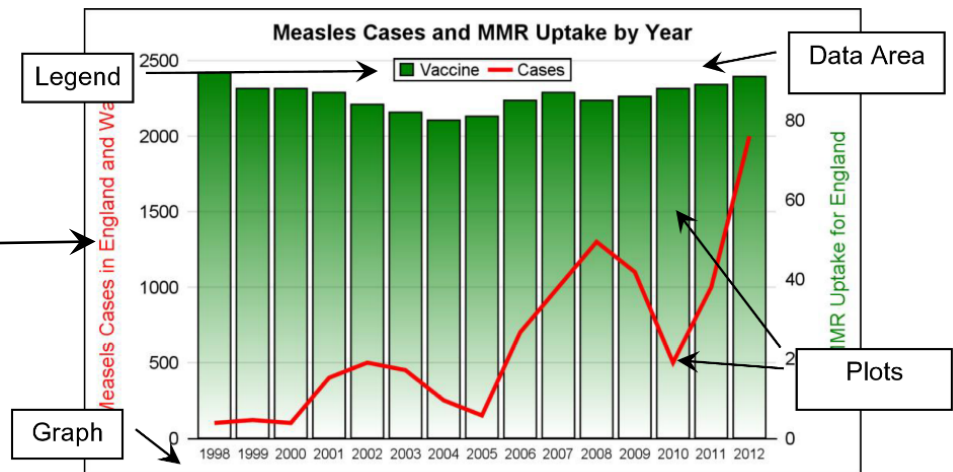
SGPLOT Procedure Syntax

PROC SGPLOT <DATA=data-set> <SGANNO=data-set>

        <DATTRMAP=data-set> <options>;

 plot-statement(s);   */* One or more */*

 <styleattrs-statement(s);>

  <refline-statement(s);>

  <dropline-statement(s);>

  <inset-statement(s);>

  <axis-statement(s);>

 <keylegend-statement(s);>

 RUN;

SGPLOT Statement Groups

Plot statements can be grouped as follows:

- **Basic Plots:** Such as scatter, series, and so on.

- **Fit and Confidence Plots:** Such as regression and loess plots.

- **Distribution plots:** Such as histograms and box plots.

- **Categorical Plots:** Such as bar charts and dot plots.

- **Other:** Reference line, Inset, Legend, etc.

Simple graphs can be created using just one plot statement.

Complex graphs can be created by combining compatible plot statements in layers to create the final graph.

§sas
Federal

SGPLOT Statement Groups

Plot Combinations:

- **Basic and Fit plots** can be combined with each other and "Other" plots.

- **Distribution plots** can be combined with plots in same group and "Other" plots.

- **Categorical plots** can be combined with plots in same group and "Other" plots.

Now let us review how to create some popular clinical graphs using SAS 9.4 features.

Commonly Requested Clinical Graphs

Let us review how to build the following clinical graphs:

- Median of Lipid Profile by Visit and Treatment.

- Swimmer Plot for Tumor Response.

- Survival Plot.

- Forest Plot by Subgroups.

- Waterfall Chart of Change in Tumor Size.

Note: All these graphs will be built using plot statements only.
Annotation is not used. This makes the graphs scalable to other data.

Median of Lipid Profile by Visit and Treatment

- This graph displays the median and CL of lipid data by visit and treatment.

- Note the visits are treated as categorical data, and equally spaced on the x-axis.

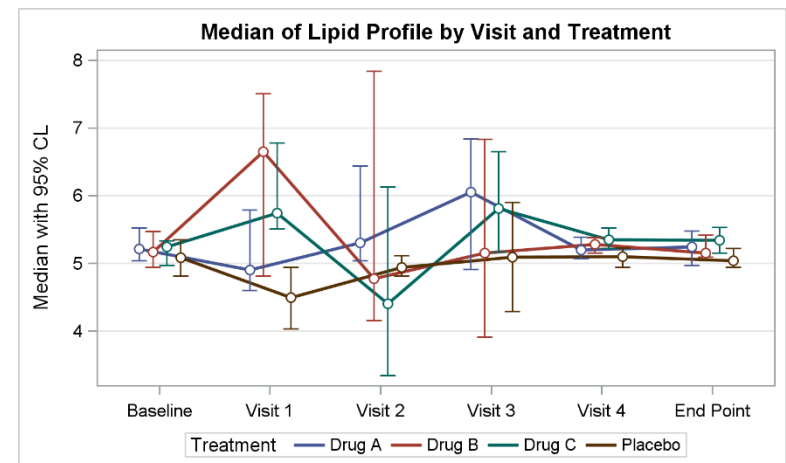Median of Lipid Profile by Visit and Treatment
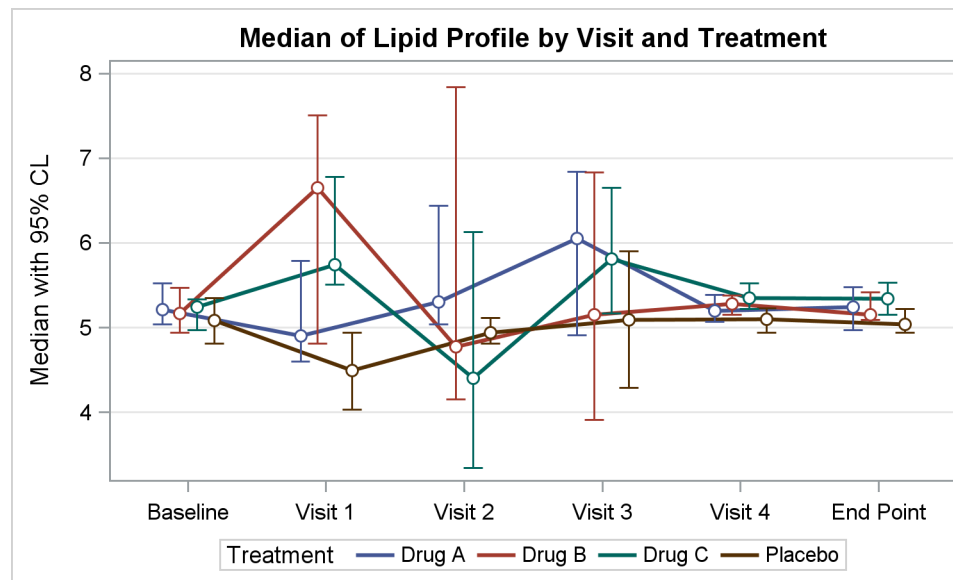
Median of Lipid Profile by Visit and Treatment

- The median values and CL by treatment for each visit are clustered around each x-axis tick value, so they can be seen clearly.

- The trend line joining the median values by treatment is also clustered so each trend line meets the appropriate value of the median.

Median of Lipid Profile by Visit and Treatment
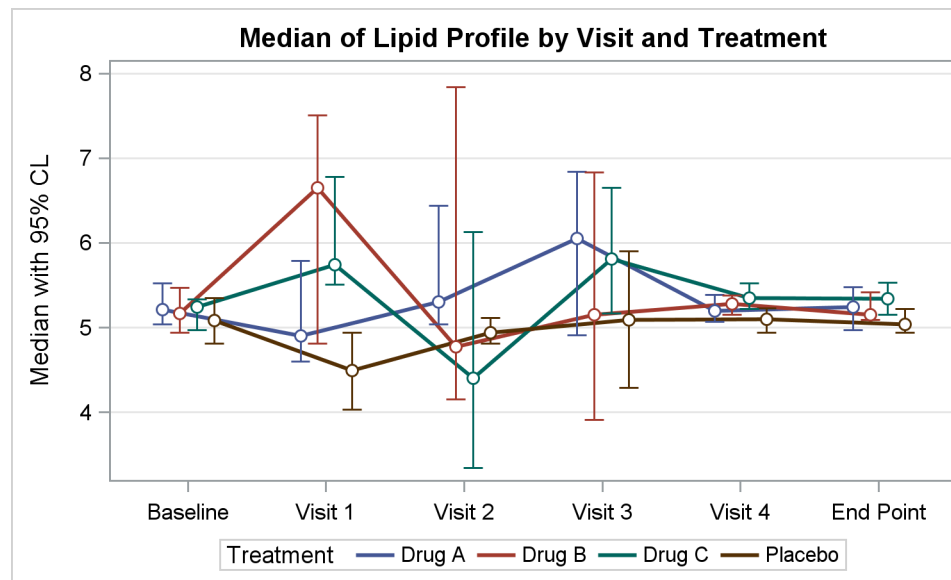
Median of Lipid Profile by Visit and Treatment

```
title 'Median of Lipid Profile by Visit and Treatment';
proc sgplot data=lipid_grp;
   series  x=day y=median / lineattrs=(pattern=solid) group=trt name='s'
           groupdisplay=cluster clusterwidth=0.5 lineattrs=(thickness=2);
   scatter x=day y=median / yerrorlower=lcl yerrorupper=ucl group=trt
           groupdisplay=cluster clusterwidth=0.5
           errorbarattrs=(thickness=1) <options>;
   keylegend 's' / title='Treatment' linelength=20;
   yaxis label='Median with 95% CL' grid;
   xaxis display=(nolabel);
run;
```


Median of Lipid Profile by Visit and Treatment

Median of Lipid Profile by Visit and Treatment

- When the visits are not at equal time intervals, it may be desirable to see the visits scaled by time.

- This graph uses a numeric "week" value for visit, and a numeric x-axis to view the data scaled by the visit interval.


Median of Lipid Profile by Week and Treatment
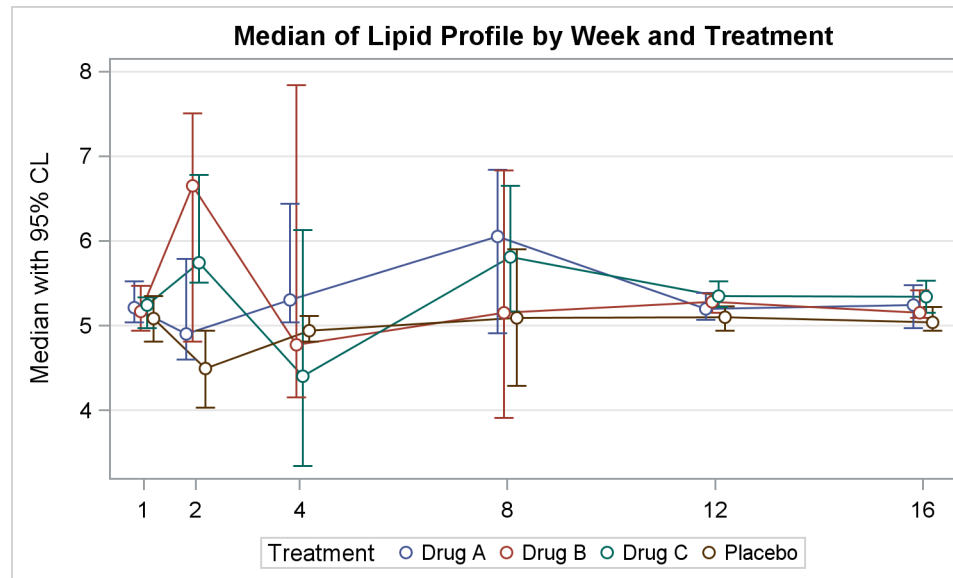
Swimmer Plot of Tumor Response

Stacey Phillips: The Swimmer plot is a graphical way of showing multiple parts of a subject's time to response "story" in one graph.

The graph includes:
- A bar for each subject showing the length of treatment classified by the disease stage.
- Indicators for the start and end of each response episode, classified by complete or partial response.
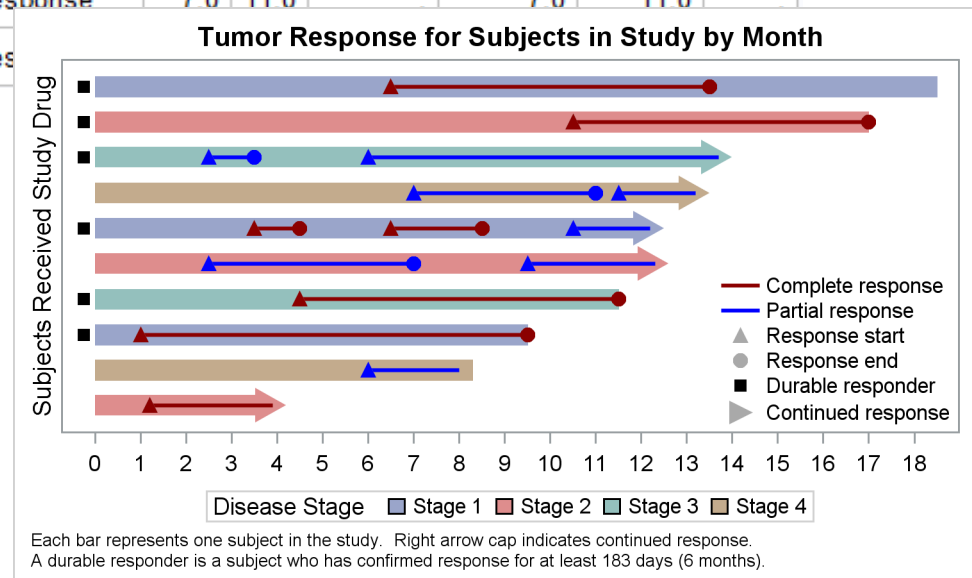- Indicators showing "Durable responder".
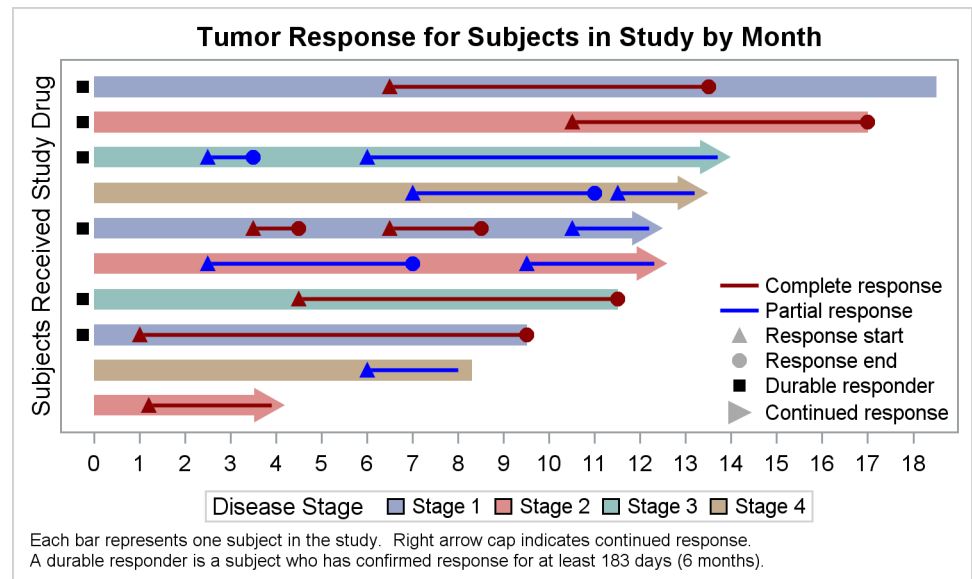
Phillips, Stacey. 2014. "Swimmer Plot: Tell a Graphical Story of Your Time to Response Data Using PROC SGPLOT." PharmaSUG, San Diego. Available at http://www.pharmasug.org/proceedings/2014/DG/PharmaSUG-2014-DG07.pdf

Swimmer Plot of Tumor Response

Data for plot

Obs	Item	Stage	Low	High	Highcap	Status	Start	End	Durable	Startline	Endline	Xmin
1	1	Stage 1	0	18.5		Complete response	6.5	13.5	-0.25	6.5	13.5	.
2	2	Stage 2	0	17.0		Complete response	10.5	17.0	-0.25	10.5	17.0	.
3	3	Stage 3	0	14.0	FilledArrow	Partial response	2.5	3.5	-0.25	2.5	3.5	.
4	3		0	14.0	FilledArrow	Partial response	6.0	.	.	6.0	13.7	.
5	4	Stage 4	0	13.5	FilledArrow	Partial response	7.0	11.0		7.0	11.0	.
6	4		0	13.5	FilledArrow	Partial res...						

Let us see how we build
this graph using SGPLOT.

Tumor Response for Subjects in Study by Month

Each bar represents one subject in the study. Right arrow cap indicates continued response.
A durable responder is a subject who has confirmed response for at least 183 days (6 months).
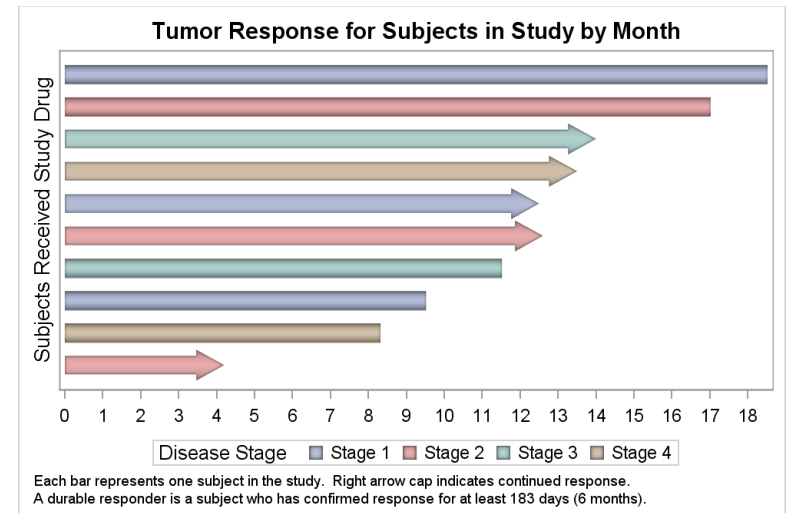
Swimmer Plot of Tumor Response

First we represent the duration for each subject by stage in the study using the HighLow plot.

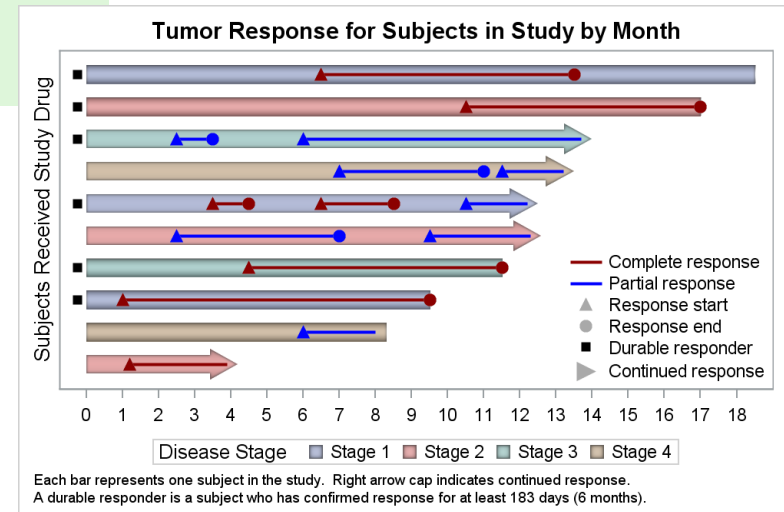- A Highcap is displayed to indicate continuing response.

```
footnote  J=l h=0.8 'Each bar..";
footnote2 J=l h=0.8 'A durable..';
title 'Tumor Response for ..';
proc sgplot data= swimmer
     dattrmap=attrmap;
highlow y=item low=low high=high /
          group=stage highcap=highcap;
xaxis display=(nolabel);
yaxis reverse display=(<options>);
keylegend 'stage';
run;
```


Tumor Response for Subjects in Study by Month

Each bar represents one subject in the study. Right arrow cap indicates continued response.
A durable responder is a subject who has confirmed response for at least 183 days (6 months).

Swimmer Plot of Tumor Response

Then we layer the complete or partial response duration and other details.

```
proc sgplot data= swimmer dattrmap=attrmap;
  highlow y=item low=low high=high / group=stage highcap=highcap;
  highlow y=item low=startline high=endline / group=status;
  scatter y=item x=start / name='s' <options>;
  scatter y=item x=end / name='e' <options>;
  scatter y=item x=start / <options>;
  scatter y=item x=end / <options>;
  scatter y=item x=xmin / <options>;
  scatter y=item x=durable / <options>;
  xaxis display=(nolabel) label='Months';
  yaxis reverse display=(<options>);
  keylegend 'stage';
  keylegend 'status' 's' 'e' 'd' 'x';
  run;
```


Tumor Response for Subjects in Study by Month

Disease Stage: Stage 1, Stage 2, Stage 3, Stage 4

Each bar represents one subject in the study. Right arrow cap indicates continued response.
A durable responder is a subject who has confirmed response for at least 183 days (6 months).
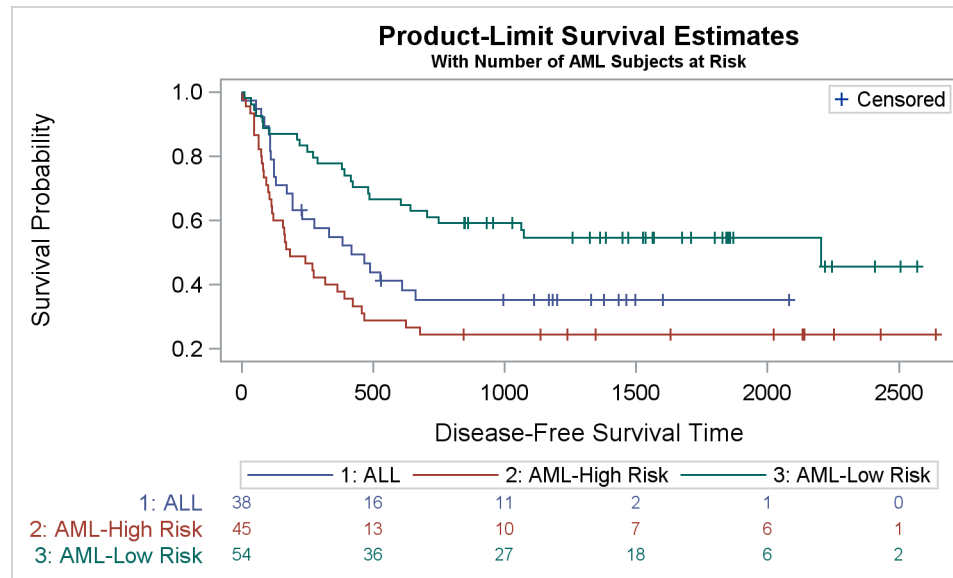
Swimmer Plot of Tumor Response

The graph features can be customize for a gray scale output.

- A YAxisTable is used to display "Stage" on the left of each bar.
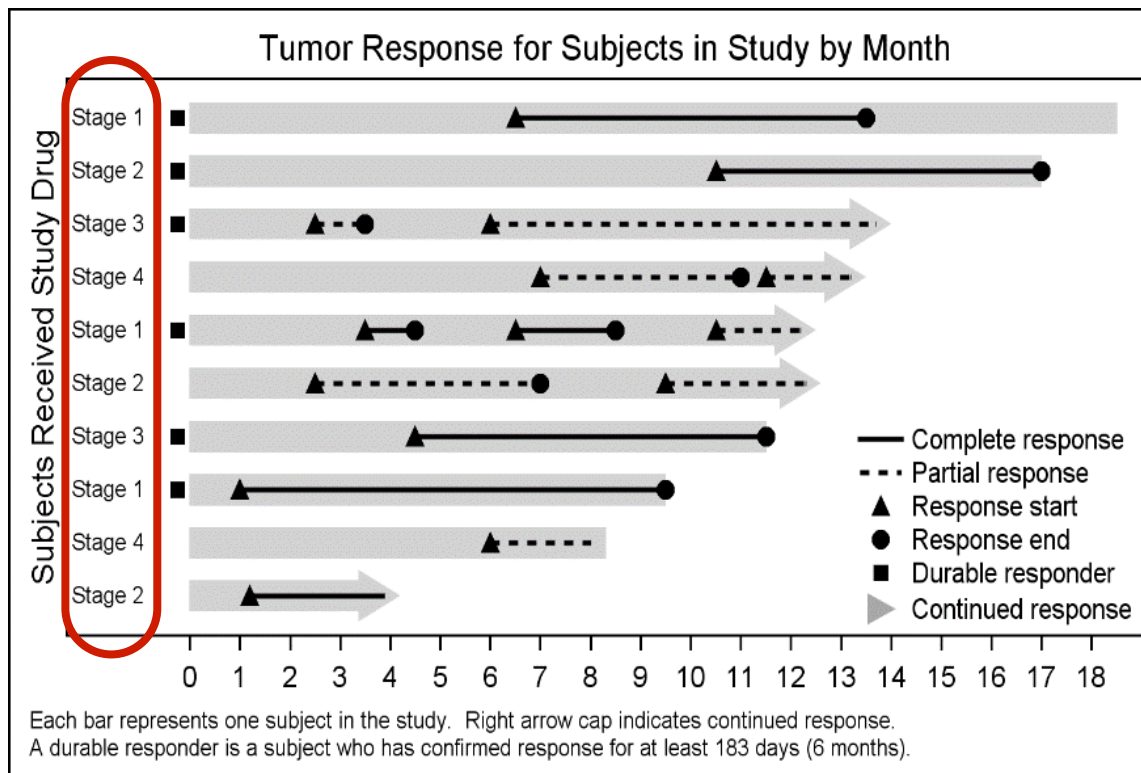
Survival Plot

- The Product-Limit Survival Estimates plot displays the survival probabilities by stratum over time.
- This plot displays the subjects in the study by stratum over time along the bottom of the graph.

Product-Limit Survival Estimates
With Number of AML Subjects at Risk

1: ALL	38	16	11	2	1	0
2: AML-High Risk	45	13	10	7	6	1
3: AML-Low Risk	54	36	27	18	6	2

Survival Plot – Get the Data

The data for the graph is obtained from the LIFETEST procedure.

```
ods graphics on;
ods output Survivalplot=SurvivalPlotData;
proc lifetest data=sashelp.BMT
  plots=survival(atrisk=0 to 2500 by 500);
  time T * Status(0);
  strata group / test=logrankadjust=sidak;
run;
```

Obs	Time	Survival	AtRisk	Event	Censored	tAtRisk	Stratum	StratumNum
1	0	1.00000	38	0	.	.	1: ALL	1
2	0	.	38	.	.	0	1: ALL	1
3	1	0.97368	38	1	.	.	1: ALL	1
4	55	0.94737	37	1	.	.	1: ALL	1
5	74	0.92105	36	1	.	.	1: ALL	1

Note: The LIFETEST procedure creates a Survival Plot automatically. The goal here is to show how to create a customized version using data from the procedure.

Survival Plot

Here is a traditional layout with "at-risk" table at the bottom.

- Note the use of solid line pattern for all curves.

```
title 'Product-Limit Survival Estimates';
title2  h=0.8 'With Number of AML Subjects at Risk';
proc sgplot data=SurvivalPlotData;
   step x=time y=survival / group=stratum
        lineattrs=(pattern=solid) name='s';
   scatter x=time y=censored / name='c';
   scatter x=time y=censored / group=stratu
   xaxistable atrisk / x=tatrisk
        class=stratum colorgroup=stratum;
   keylegend 'c' / location=inside;
   keylegend 's';
run;
```
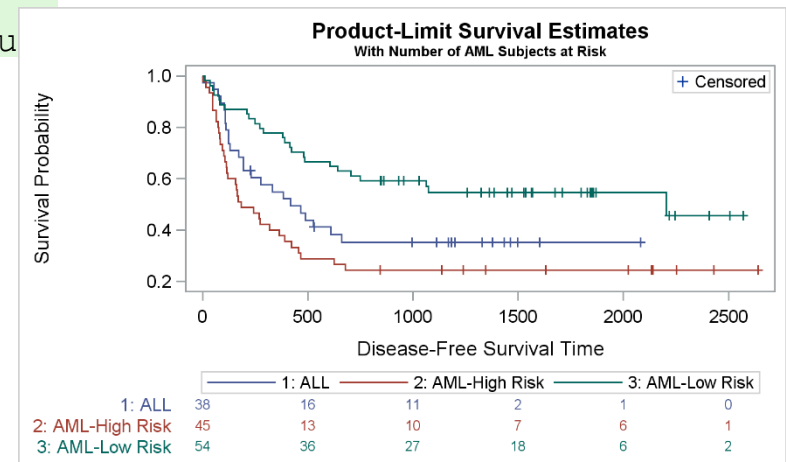
Survival Plot with Alternative Layout

The "At-Risk" table is traditionally displayed at the bottom of the graph.

- This normally places the text values far from the curves with a lot in intervening ink.

- An alternative is to place the table closer to the curves, reducing the distractions.

- This is easily done using the LOCATION=INSIDE option.
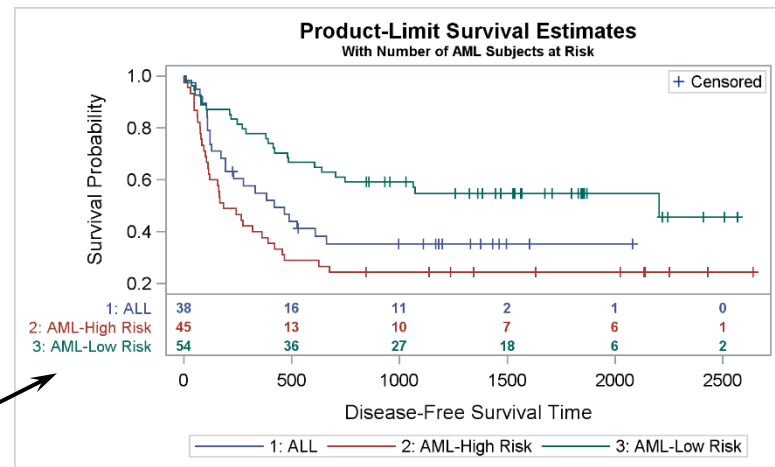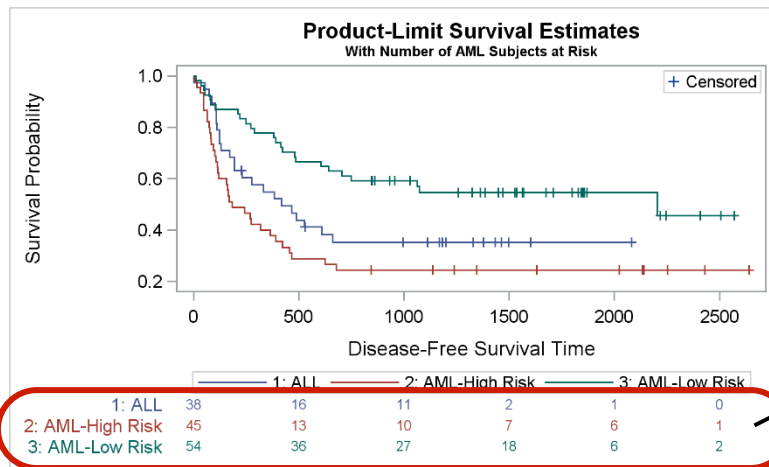

```
xaxistable atrisk / x=tatrisk class=stratum
          colorgroup=stratum location=inside;
```

Survival Plot in Gray Scale

Survival plot for a gray scale medium has special needs.

- Line patterns can be used to distinguish the stratum levels.

- However, line patterns may not be optimal with a step plot.

- So, here I have used solid patterns with curve labels to identify the stratum levels.
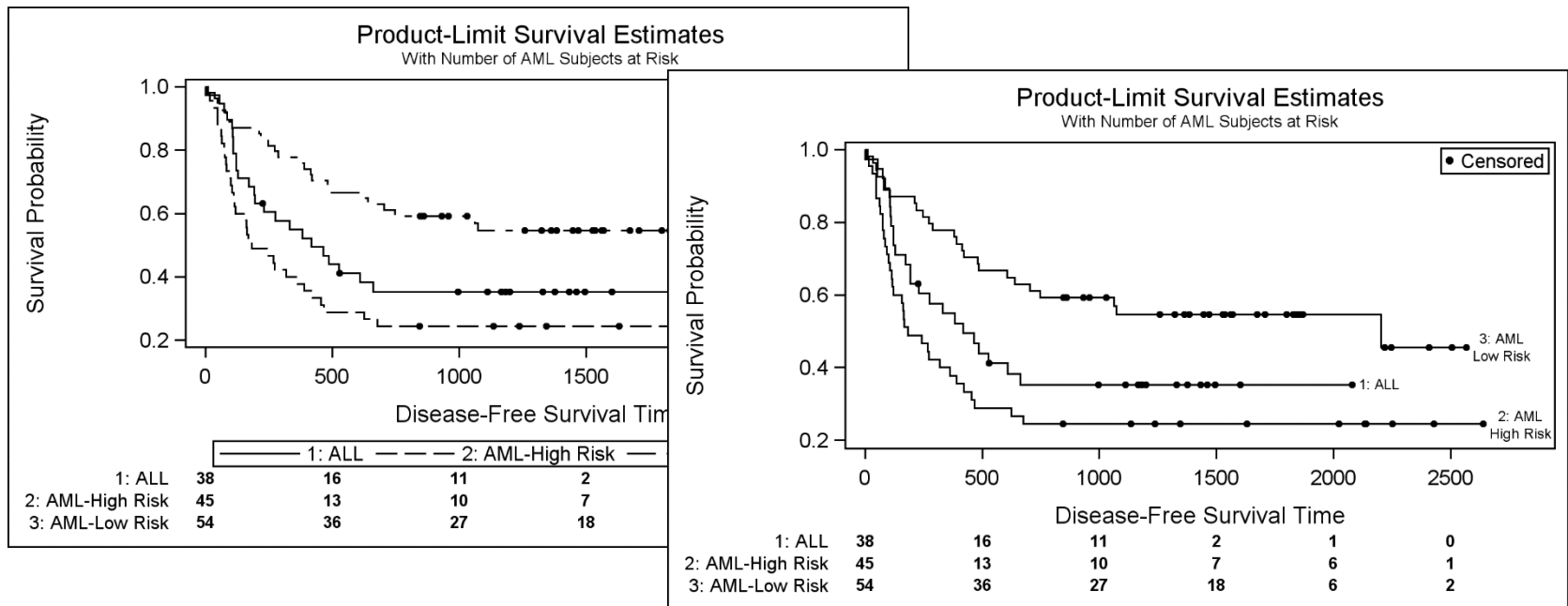
Forest Plot by Subgroups

- A forest plot is used for meta-analysis of the results of randomized controlled trials.
- This plot displays the Hazard Ratio Plot along with other statistics for the study in one plot.
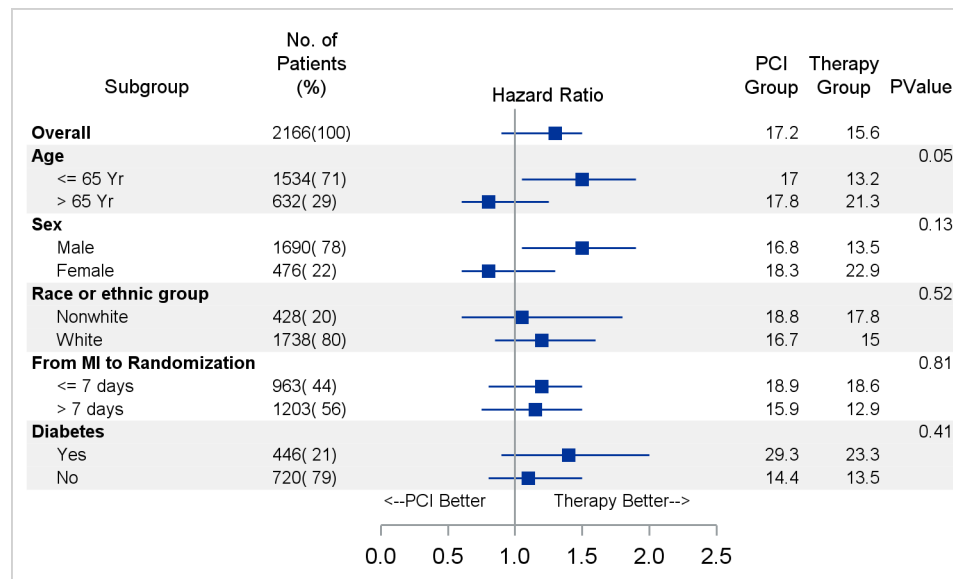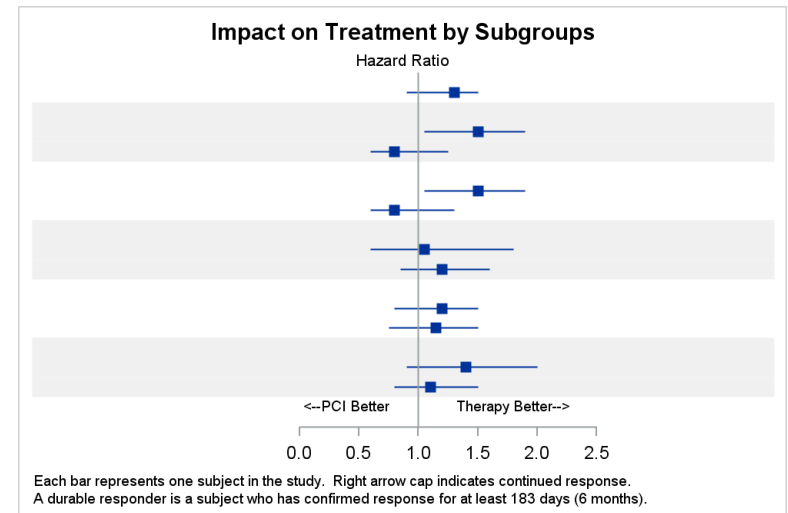
Forest Plot by Subgroups - Data

- Data for the Forest Plot by ObsId.

- Id determines if the observation is a Group header or values.

- IndentWt determines the amount of indentation.

- Ref column values are missing where horizontal band is not shown.

Obs	ObsId	Id	Subgroup	indentWt	Count	Percent	Mean	Low	High	PCIGroup	Group	PValue	ref
1	1	1	Overall	0	2166	100	1.3	0.90	1.50	17.2	15.6		
2	2	1	Age	0								0.05	2
3	3	2	<= 65 Yr	1	1534	71	1.5	1.05	1.90	17.0	13.2		3
4	4	2	> 65 Yr	1	632	29	0.8	0.60	1.25	17.8	21.3		4
5	5	1	Sex	0								0.13	
6	6	2	Male	1	1690	78	1.5	1.05	1.90	16.8	13.5		
7	7	2	Female	1	476	22	0.8	0.60	1.30	18.3	22.9		

Forest Plot by Subgroups – Hazard Ratio

First we display the graphical Hazard Ratio information by Study

```
title 'Impact on Treatment by Subgroups';
proc sgplot data=forest_subgroup_2;
   styleattrs axisextent=data;
   refline ref / <options>;
   highlow y=obsid low=low high=high;
   scatter y=obsid x=mean / <options>;
   refline 1 / axis=x;
   text x=xl y=obsid text=text / <options>;
   yaxis reverse display=none;
   xaxis display=(nolabel) <options>;
   x2axis label='Hazard Ratio' / <options>;
run;
```
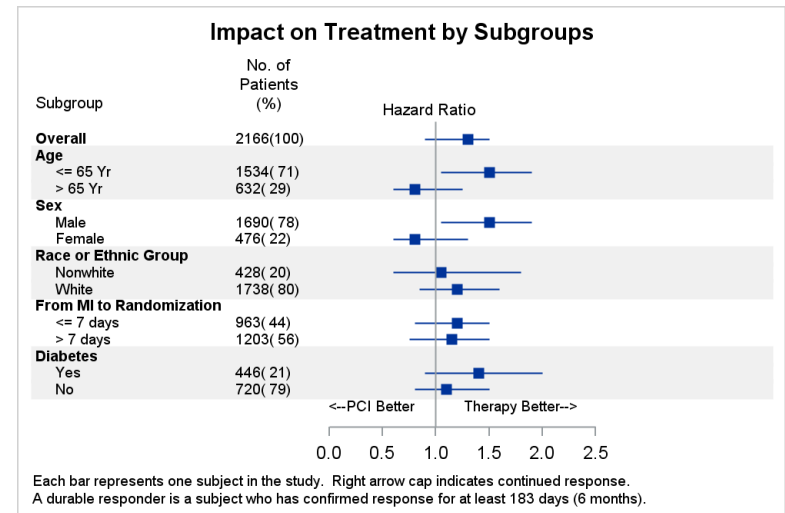

Some options are trimmed to fit the space.

Forest Plot by Subgroups – Left Columns

Add display of the columns on the left side

```
title 'Impact on Treatment by Subgroups';
proc sgplot data=forest_subgroup_2;
  styleattrs axisextent=data;
  refline ref / <options>;
  highlow y=obsid low=low high=high;
  scatter y=obsid x=mean / <options>;
  refline 1 / axis=x;
  text x=xl y=obsid text=text / <options>;
  yaxistable subgroup  / <options>;
  yaxistable countpct / <options>;
  yaxis reverse display=none;
  xaxis display=(nolabel) <options>;
  x2axis label='Hazard Ratio' / <options>;
run;
```
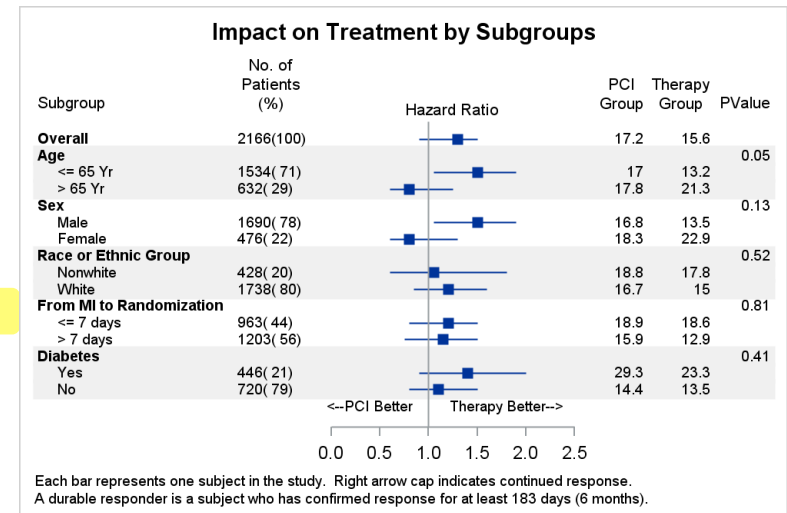

Some options are trimmed to fit the space.

Forest Plot by Subgroups – Left Columns

Add display of the columns on the rightside

```
title 'Impact on Treatment by Subgroups';
proc sgplot data=forest_subgroup_2;
  styleattrs axisextent=data;
  refline ref / <options>;
  highlow y=obsid low=low high=high;
  scatter y=obsid x=mean / <options>;
  refline 1 / axis=x;
  text x=xl y=obsid text=text / <options>;
  yaxistable subgroup  / <options>;
  yaxistable countpct / <options>;
  yaxistable PCIGrp group pval / <options>;
  yaxis reverse display=none;
  xaxis display=(nolabel) <options>;
  x2axis label='Hazard Ratio' / <options>;
run;
```


Impact on Treatment by Subgroups

Some options are trimmed to fit the space.

Waterfall Chart of Change in Tumor Size

- A waterfall chart is commonly used in the Oncology domain to track the change in tumor size for subjects in a study by treatment.

- This plot displays the change in tumor size for each patient in the study as a bar classified by treatment, sorted by percent change.
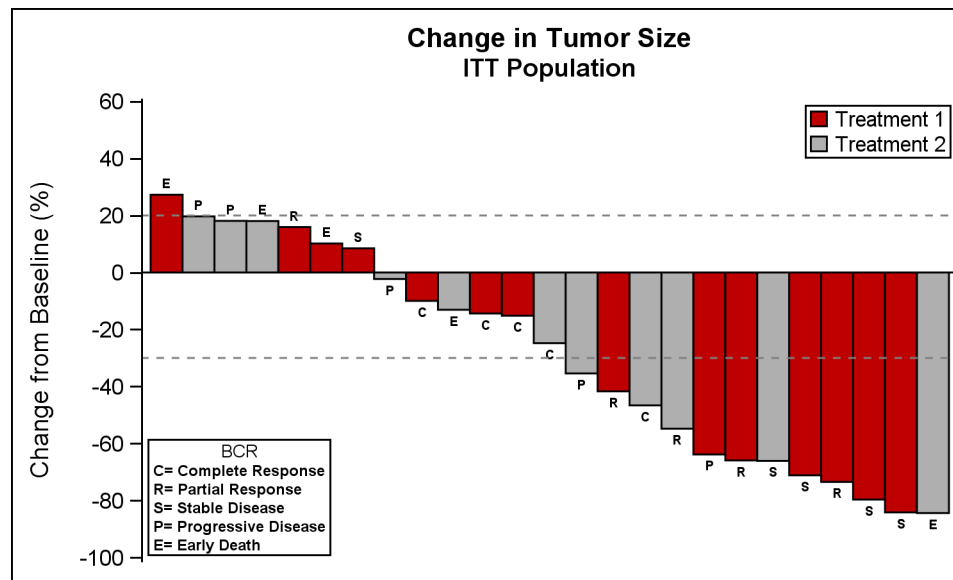
Waterfall Chart of Change in Tumor Size

```
proc sgplot data=TumorSize noborder;
   styleattrs datacolors=(cxbf0000 cx4f4f4f);
   vbarparm category=cid response=change /
           group=group;
   refline 20 -30;
   xaxis display=none;
   yaxis values=(60 to -100 by -20);
   inset "C= Complete Response"
       "R= Partial Response " ...
       position=bottomleft border;
   keylegend / location=inside
       across=1 border;
run;
```
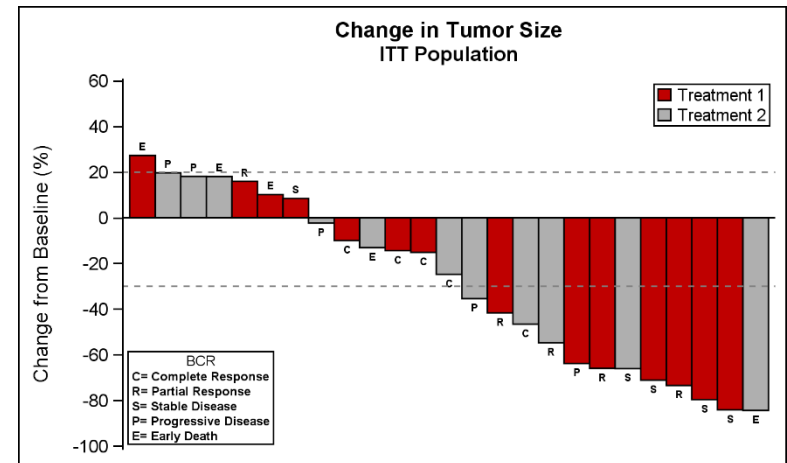
Waterfall Chart of Change in Tumor Size

Here is an alternate view of the same graph with a band plot.
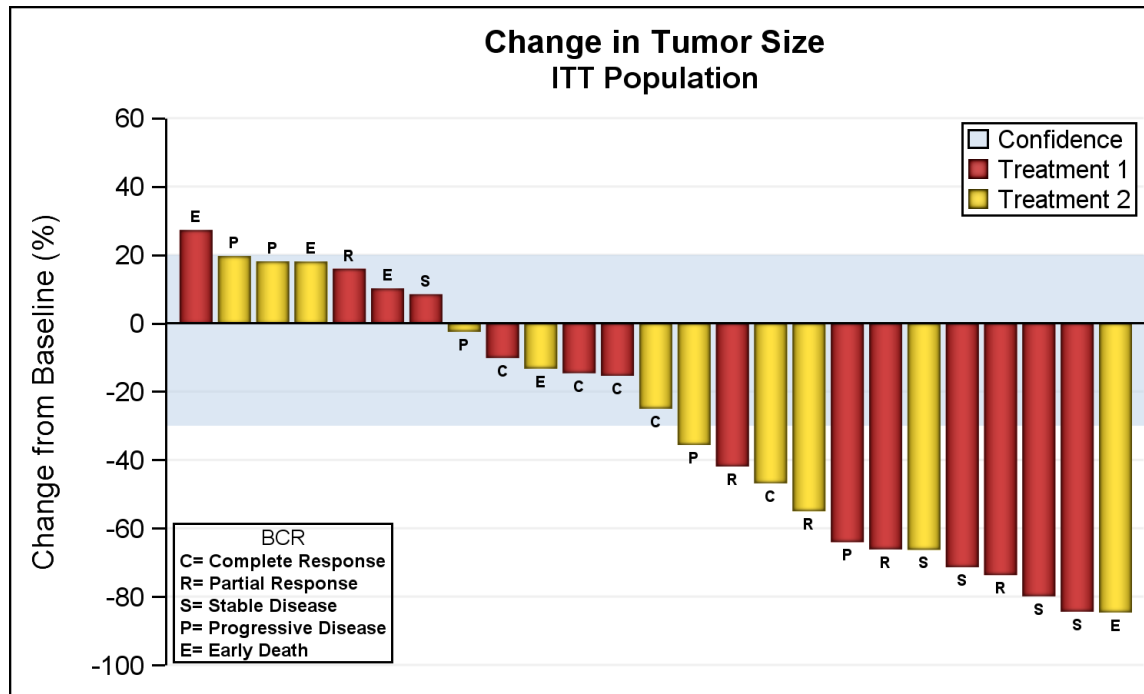
Conclusion

- The SAS 9.4 SGPLOT procedure provides new features that simplify creating clinical graphs.

- The XAxisTable and YAxisTable statement included in SAS 9.4 SGPLOT make it easy to add tabular data to graphs.

- The SAS 9.3 SGPLOT procedure can be used to create such graphs using the SGAnnotate feature. For more information on this, see:

   http://support.sas.com/rnd/datavisualization/papers/sgf2016/SAS4321-2016.pdf

References

- Phillips, Stacey. 2014. "Swimmer Plot: Tell a Graphical Story of Your Time to Response Data Using PROC SGPLOT." PharmaSUG, San Diego. Available at http://www.pharmasug.org/proceedings/2014/DG/PharmaSUG-2014-DG07.pdf

- Pandya, Niraj. 2012. "Waterfall Charts in Oncology Trials – Ride the Wave." PharmaSUG. Available at http://www.pharmasug.org/proceedings/2012/DG/PharmaSUG-2012-DG13.pdf

- Matange, Sanjay. 2016. Clinical Graphs Using SAS. SAS Global Forum, 2016. Available at: http://support.sas.com/rnd/datavisualization/papers/sgf2016/SAS4321-2016.pdf

- Matange, Sanjay. 2016. Clinical Graphs Using SAS. SAS Institute. Available at: https://www.sas.com/store/prodBK_68179_en.html

- Graphically Speaking Blog by Sanjay Matange. Available at: http://blogs.sas.com/content/graphicallyspeaking/