PharmaSUG 2019 - Poster PO-253

F2Plots: Visualizing relative treatment effects in cancer clinical trials

Yellareddy Badduri, Quartesian.

ABSTRACT

In every year there were many clinical trials are conducting on different types of Cancers. With Cancer trials increasingly reporting nontime-to-event outcomes, data visualization has evolved to incorporate parameters such as responses to therapy, duration and degree of response, and novel representations of underlying tumor biology. Graphs and figures are excellent tools for data visualization and they have capability to display data figuratively and enables rapid interpretation. F2 plots (Forest and Funnel) were initially developed for presenting results of meta-analysis. Forest plot is an intuitive, convenient and used to show the relative treatment effect of an intervention between groups within the larger cohort. Forest plot is easily understood constitute several horizontal lines, which represent the 95% confidence interval, and a central symbol in the middle of the line segment, which represents a point estimate that is usually the median or mean. Funnel plots are scatter plots of the effect estimates from individual studies against some measure of each study's size or precision. Another advantage of funnel plots are that there is no spurious ranking of institutions, the eye is naturally drawn to important points that lie outside the funnel, there is allowance for increased variability of smaller units and it is easy to produce with standard spreadsheet. This paper will explain about different SAS programming approaches for producing both forest and the funnel plot, and representations that used to illustrate treatment effects.

INTRODUCTION

Cancer remains a leading cause of mortality worldwide and Cancer is a pathophysiologically heterogeneous disease that rapidly progresses to an uncontrollable stage after onset.The effectiveness of a new agent or combination therapy in a prospective cancer clinical trial is usually assessed by analyzing outcomes such as tumor response and overall survival. Although the measurements of clinical outcomes for cancer treatments have become diverse and complex, there remains a need for clear and easily interpreted representations of patients' experiences. Reporting nontime-to-event outcomes increased in Cancer clinical trials, due to this data visualization has been progressed to incorporate parameters such as responses to therapy, duration and degree of response, and novel representations of underlying tumor biology. Graphs and figures allow the illustration and visualization of data to demonstrate an intervention or treatment effect in oncology treatments.

VISUALIZATION OF TREATMENT EFFECTS

In cancer clinical trials understanding the treatment effects is important to assess the relative therapeutic efficacy between different groups. Forest plots and Funnel plots are examples of graphical representations of treatment effect and they are explained below.

FOREST PLOTS

Forest Plots were initially developed for presenting results of meta-analysis. However, since late 1990s they have gained popularity for presenting institutional risk-adjusted performance. The term forest plot was coined only in 1996 and derives from the appearance of this representation as a "forest" of lines. Forest plots constitute several horizontal lines, which represent the 95% confidence interval, and a central symbol in the middle of the line segment, which represents a point estimate that is usually the median or mean. The horizontal line provides a measure of the precision of the estimate, with line length being directly proportional to the variability in the data.

The visual representation allows for easy review and comparison across many factors. Forest plots are useful in considering the behaviors of subgroups within a larger dataset. For example, the benefit for a particular treatment may only be small in a large population, but separating out and analyzing the effect of the therapy in different subgroups may sometimes identify those who may benefit more. Such analyses can be subject to error, especially where small numbers of data points are present and confidence intervals are therefore wider than for the entire group.

METHODS TO CREATE FOREST PLOTS

Method 1: By Using PROC TEMPLATE

Stage 1: Creation of Input Data Set

The input dataset should contain Overall Survival data by Subgroups as given in the below table and the data format in input dataset should be as it is expected in Forest plot output.

Column 1: Subgroup Categories
Column 2: Plot
Column 3: Patients counts between two treatment groups
Column 4: Event counts between two treatment groups
Column 5: Median Overall Survival between two treatment groups
Column 6: HR (Hazard Ratio) and 95% CI (Confidence intervals)

Procedures PROC SQL is used to get the counts and PROC PHREG produced the HRs and 95% CI. In order to provide some desired formatting to the counts and percentages (for example placing the percentages within parentheses) some values are constructed by using concatenation.

The categorical groups being analyzed shown in the following table:

COL1	**COL3**	**COL4**	**COL5**	**COL6**
Geographical region				
North America	412 vs 416	138 vs 139	14.36 vs 13.34	0.951 (0.751,1.204)
Japan/Korea	119 vs 114	59 vs 52	15.61 vs 16.99	1.097 (0.754,1.596)
Rest of the World	77 vs 78	22 vs 35	NA vs 10.81	0.713 (0.416,1.222)
Location of the primary tumor				
Left colon	435 vs 409	150 vs 141	15.11 vs 16.56	0.992 (0.788,1.249)
Right colon	173 vs 199	69 vs 85	14.06 vs 11.93	0.881 (0.640,1.212)
Age				
<65	372 vs 395	125 vs 134	14.55 vs 15.34	1.025 (0.803,1.308)
>=65	236 vs 213	94 vs 92	14.26 vs 12.09	0.837 (0.627,1.116)
Sex				
Male	365 vs 368	135 vs 139	14.03 vs 14.29	1.012 (0.798,1.283)
Female	243 vs 240	84 vs 87	15.64 vs 13.34	0.872 (0.645,1.178)
Primary tumor site				
Colon	397 vs 424	141 vs 160	14.03 vs 13.17	0.939 (0.748,1.177)
Rectum	211 vs 184	78 vs 66	17.58 vs 17.15	0.989 (0.711,1.374)

COL1	COL3	COL4	COL5	COL6
ECOG Performance Status				
0	311 vs 329	109 vs 100	16.13 vs 16.99	1.065 (0.811,1.398)
1	294 vs 279	109 vs 126	12.16 vs 11.40	0.844 (0.653,1.091)
Race				
White	349 vs 371	118 vs 125	14.06 vs 13.34	0.934 (0.726,1.202)
Black	40 vs 23	12 vs 9	14.36 vs 13.17	0.782 (0.327,1.871)
Asian	193 vs 193	81 vs 86	15.67 vs 14.29	0.947 (0.698,1.284)
Other	26 vs 21	8 vs 6	11.86 vs NA	1.514 (0.507,4.521)
Country				
Canada	23 vs 13	8 vs 3	15.11 vs NA	1.557 (0.401,6.042)
USA	191 vs 209	56 vs 70	15.64 vs 13.44	0.808 (0.568,1.150)
Western Europe	168 vs 166	62 vs 55	12.16 vs 13.34	1.070 (0.744,1.539)
Australia	30 vs 28	12 vs 11	11.86 vs 12.02	1.117 (0.491,2.540)
Japan	65 vs 61	41 vs 36	16.23 vs 17.15	1.027 (0.653,1.614)
Korea	39 vs 40	13 vs 10	15.67 vs 18.79	1.284 (0.560,2.947)
Other	92 vs 91	27 vs 41	NA vs 10.81	0.790 (0.483,1.293)

Stage 2: Graph Template Language (GTL)

The Graph Template Language (GTL) offers a wide range of plot statements, layouts, and other options and it is a powerful tool and it is the foundation for the ODS Graphics System.

Creating a graph using GTL is a two-step process:

The first step is to define a statgraph template that describes the structure and appearance of a graph to be produced (see Appendix).

The second step is to use the SGRENDER procedure to associate the template and the data object to create the graph (see Appendix).

Method 2: By Using PROC SGPLOT

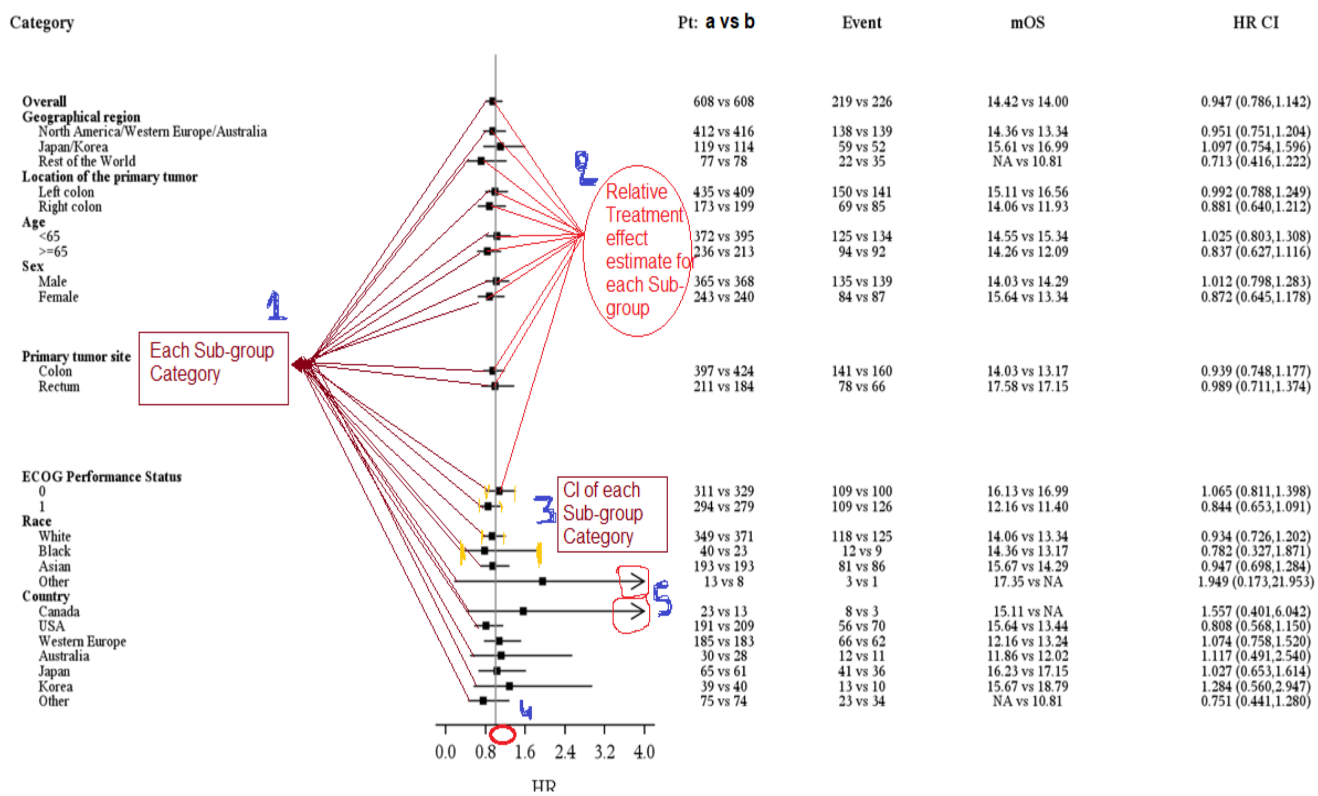
The SG Procedures provide an easy to use procedure type syntax for using the GTL features. The SGPLOT procedure produces a variety of graphs including bar charts, scatter plots, and line graphs.

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.

To create Forest plot by using SGPLOT code (see Appendix).

FOREST PLOT Output:



Understanding of FOREST PLOT:

Point 1: Each Sub-group Category represented by a line.

Point 2: The mid-point of the box represents the point effect estimate. The area of the box represents the weight given to the Sub-group Category.

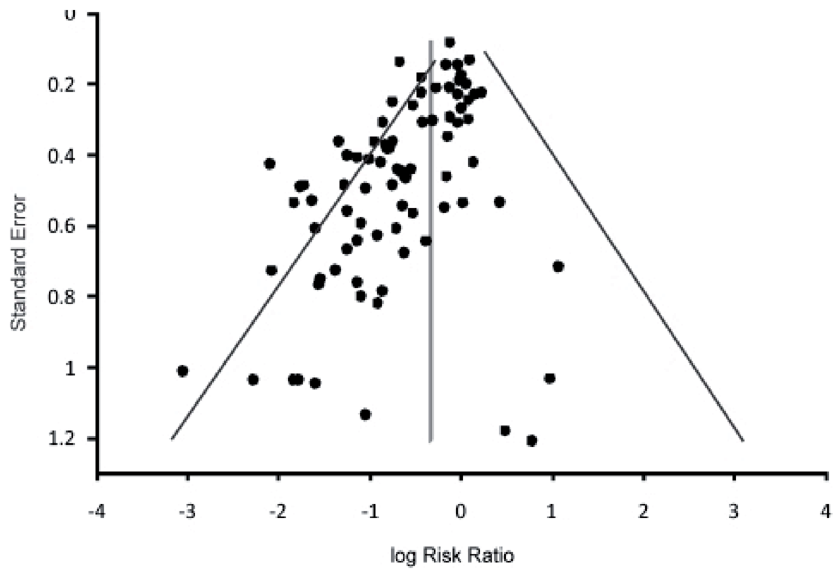
Point 3: The width of the line shows the confidence intervals of the effect estimate of each Sub-group Category. Point estimate is best guess of the true effect in the population. 95% confidence intervals mean that there is a 95% chance that the true effect in the population will lie within the range. They also mean that if the trial is repeated, there is a 95% chance that the point estimate from the trial lies within the 95% confidence intervals obtained in the systematic review.

Point 4: All ratios commonly used as effect measures in meta-analyses are relative measures. 1 indicates no effect. If 1 is included in the 95% confidence intervals, it indicates that there is no statistical significance at 5% significance levels. If 1 is not included in the 95% confidence intervals, the results are statistically significant at 5% significance levels.

Point 5: Truncating HR x-axis to 4 and adding left arrow for the CI which is larger than 4.

FUNNEL PLOT:

Funnel plots were first introduced in 1984 for presenting results of meta-analysis. They are increasingly used to compare studies in order to check for publication bias and identify outliers. They are scatter plots of the effect estimates from individual studies against some measure of each study's size or precision. A symmetrical funnel shape plot would give an indication of an appropriately sampled dataset whereas an asymmetrical funnel plot would imply possible publication bias or heterogeneity between studies as shown in below Example Figure. Caution must be taken when interpreting funnel plots as a clear definition of precision, and effect in constructing the funnel plot may affect the shape of the plot, leading to discrepancies. There are several reasons that can lead to asymmetry of the funnel plot, including selection bias, publication bias, true heterogeneity, data irregularities. An example of funnel plot asymmetry could be seen in smaller studies with a higher proportion of patients in the high-risk groups. The patients in the high-risk groups may tend to have a higher response rate to the study treatment arm compared with the patients in the normal or low risk groups. This could lead to funnel plot asymmetry because of risk differences with variation in the number of patients in each group, rather than bias.



Above Figure is for Illustrations of funnel plot asymmetry. Log of the risk ratios was plotted against the standard error of the risk ratio of each study to identify asymmetry in the distribution of trials. Gaps in the funnel plot suggest potential publication bias. Figure courtesy of Ritchie et al.

CONCLUSION

Cancer remains a leading cause of mortality and it is a global issue that affects many lives across geography and all demographics. Visualization and interpretation of cancer data in the context of clinical trials is crucial in order to understand the disease and the potential treatment effect. Both Forest plots and the Funnel plots can be used for graphical representations of treatment effect in oncology trials. Both plots can be generated by GTL and ODS Statistical Graphics procedures. Forest plots Helps determine behaviors of different subgroups within a larger dataset. But subject to error if there are only small number of data points within subgroup analysis resulting in false interpretation. Funnel plots are scatter plots of the effect estimates that can give an indication of heterogeneity. But Shape of the plot is dependent on number of patients recruited in different risk groups.

REFERENCES

1. Tufte ER. The visual display of quantitative information. Cheshire, CT: Graphics Press; 1986.
2. Pocock SJ, Travison TG, Wruck LM. Figures in clinical trial reports: current practice & scope for improvement. Trials. 2007;8:36.
3. Kaplan EL MP. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53:457–481.
4. Chen TT. Statistical issues and challenges in immuno-oncology. J Immunother Cancer. 2013;1:18.
5. Ritchie ML, Romanuk TN. A meta-analysis of probiotic efficacy for gastrointestinal diseases. PloS One. 2012;7(4):e34938.
6. Complex Clinical Graphs using SAS, Sanjay Matange. SAS Institute, Inc.

CONTACT INFORMATION

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

Yellareddy Badduri
Quartesian
+91-963-246-4567
yella.badduri@quartesian.com

APPENDIX

TEMPLATE CODE FOR FOREST PLOT

Step 1:

```
proc template;
      define statgraph Forest;
      dynamic  _bandcolor _headercolor _subgroupcolor;
      begingraph/border=FALSE designwidth=30cm designheight=15cm;

layout lattice / columns=6 columnweights=(0.31 0.16 0.10 0.10 0.14 0.19)
rowdatarange=union;
          rowAxes;
          rowAxis/reverse=true display=none /*offsetmin=0*/;
          endRowAxes;

sidebar / align=top;
layout lattice/rows=2 columns=6 columnweights=(0.31 0.16 0.10 0.10 0.14 0.19)
  backgroundcolor=_headercolor opaque=true;
  entry textattrs=(size=8 weight=bold) halign=left "Category";
  entry textattrs=(size=8 weight=bold) "";
  entry textattrs=(size=8 weight=bold) halign=center "Pt:N + F vs F";
  entry textattrs=(size=8 weight=bold) halign=center "Event";
  entry textattrs=(size=8 weight=bold) halign=center "mOS";
  entry textattrs=(size=8 weight=bold) halign=center "HR CI";
  endlayout;
  endsidebar;


    *** First column-Subgroup ****;
layout overlay/ walldisplay=none xaxisopts=(display=none
                  linearopts=(viewmin=0 viewmax=25) /*offsetmin=0*/);
     %let refbandattrs=lineattrs=(thickness=7 color=_bandcolor);
      referenceline y=ref/ &refbandattrs;
highlowplot y=obsid low=zero high=zero/highlabel=heading
lineattrs=(thickness=0) labelattrs=(size=7 weight=bold) ;
highlowplot y=obsid low=zero high=one /highlabel=col7 lineattrs=(thickness=0)
labelattrs=(size=7);
endlayout;

*** Second Column-Graph ***;
layout overlay/xaxisopts=(label=("HR") linearopts=(viewmax=4
tickvaluepriority=true tickvaluelist=( %cmpres(&txt1) %cmpres(&txt2)
%cmpres(&txt3) %cmpres(&txt4) %cmpres(&txt5) %cmpres(&txt6)  )))
 walldisplay=none;

referenceline y=ref/ &refbandattrs;
highlowplot y=obsid low=HRLowerCL high=HRUpperCL /highcap=capsymbol;
scatterplot y=obsid x=HazardRatio/markerattrs=(symbol=squarefilled);
referenceline x=1;
endlayout;

*** Third column- Pt:A vs B*************;
layout overlay/ x2axisopts=(display=none) walldisplay=none;
%let refbandattrs1=lineattrs=(thickness=7 color=_bandcolor);
```

```
referenceline y=ref/ &refbandattrs1;
scatterplot y=obsid x=trt/markercharacter=col2 xaxis=x2;
endlayout;

*** Fourth column- Event********************;
layout overlay/ x2axisopts=(display=none) walldisplay=none;
%let refbandattrs1=lineattrs=(thickness=7 color=_bandcolor);
referenceline y=ref/ &refbandattrs1;
scatterplot y=obsid x=evt/markercharacter=col3 xaxis=x2;
endlayout;

*** Fifth column- mOS***;
layout overlay/ x2axisopts=(display=none) walldisplay=none;
%let refbandattrs1=lineattrs=(thickness=7 color=_bandcolor);
referenceline y=ref/ &refbandattrs1;
scatterplot y=obsid x=os/markercharacter=col4 xaxis=x2;
endlayout;

*** 6th column-Hazardratio and 95 % CI***;
layout overlay/ x2axisopts=(display=none) walldisplay=none;
     %let refbandattrs1=lineattrs=(thickness=7 color=_bandcolor);
      referenceline y=ref/ &refbandattrs1;
      scatterplot y=obsid x=hr/markercharacter=col5 xaxis=x2;
    endlayout;

  endlayout;
endgraph;
end;
run;
```

Step 2:

```
*** Render ForestPlot ***;
proc sgrender data=fin4(where=(grpx=&i)) template=forest&i;
      dynamic _bandcolor='white' _headercolor='white';
run;
```

SGPLOT CODE FOR FOREST PLOT

```
proc sgplot data= forest  dattrmap=attrmap noautolegend nocycleattrs nowall;
  ******* To remove box around plot ****;
      styleattrs  axisextent=data ;
  ****** To add reference line *****;
      refline 1 / axis=x ;
  ****** estimates and CIs ******;
      scatter y=obs x=mean /  markerattrs=(symbol=squarefilled);
      highlow y=obs low=low high=high;
      refline 1 / axis=x; text x=xl y=obsid text=text ;
  ******** Adding row labels ********** ;
      yaxistable subgroup /   location=inside position=left textgroup=level
      textgroupid=text indentweight=indentWt ;
  ***** adding yaxis table at right ****;
      yaxistable col3 col4 col5 col6/ location=inside position=right ;
  ******** Add banding to every other group ********** ;
      yaxis reverse display=none  reverse display=none offsetmin=0
```

```
    colorbands=odd colorbandsattrs=(transparency=1);
***** cleaner axis *****;
 xaxis  display=(nolabel) ;
***** text above xaxis *****;
 text x=x1 y=record text=text / position=bottom contributeoffsets=none
    strip ;
****** text above x2axis ****;
 scatter y=obs x=mean /  markerattrs=(size=0)      x2axis ;
 x2axis  label='Hazard Ratio' display=(noline noticks novalues)       ;

 run;
```