

How to produce Forest Plots with Proc GPLOT and Annotate Facility

Laetitia Lemoine Le Halpert, Quintiles, Paris

ABSTRACT

The purpose of this paper is to describe how to produce Forest-Plots by using the SAS® procedure GPLOT and Annotate facility.

The Forest-Plot can be used for various analyses to display statistical data such as Odds-Ratios, Hazard Rates, Risk ratios and many other statistical data.

INTRODUCTION

There is no SAS procedure which can directly display a Forest plot. Using an annotate dataset containing all the elements of the graphics and then using a SAS® proc GPLOT allow to generate this graph.

This paper will be illustrated by an example of a Forest-Plot (see Section 4) produced to show the benefit of one medical treatment compared to another using Odds-Ratios. Odds-Ratios and 95%CIs will also be displayed on the same graph for several sub-groups of patients.

The forest-plot allows the reader to quickly assess if an odds-ratio is significant by showing the value of the OR and its 95%CI. An odds-ratio is significant when the line joining the confidence limits does not cross the vertical axis, i.e. if the 95%CI does not contain the value 0.

Four steps are needed to produce the Forest-Plot.

- 1- Preparation of the main SAS® dataset
- 2- Creation of annotate dataset containing all information to be displayed in the plot
This part will explain how and in which areas to put text (label of variables, of sub-groups, number of events in each sub-group, odds-ratios and 95%CIs, draw horizontal and vertical lines, how to put ticks on the horizontal axis since a logarithmic scale has to be used etc.).
- 3- Procedure GPLOT
- 4- Forest-plot results

1-PREPARATION OF THE MAIN SAS® DATASET

The example used is data from 1000 patients suffering from cancer, two treatment groups X and Y, several variables such as the Best Overall Response to treatment, and sub-groups variables (age, sex, ethnic origin, region).

Frequencies, common OR and CI values are obtained by performing a PROC FREQ between the response variable and treatment group, for each sub-group variable.

```
ods output commonrelrisks=commonOR crosstabfreqs=freqs ;  
proc freq data=table;  
tables trtgrp*response /all ;  
by subgroup ;  
run;
```

PhUSE 2008

Resulting Dataset 'FINAL'

VAR	CAT	N	N1	N2	NBEV1	NBEV2	OR	ORLCL	ORUCL	NTEXT	NBEV	ORCI
1	10	1091	547	544	279	231	1.408	1.109	1.789	(547 vs 544)	279 vs 231	1.408 [1.109, 1.789]
2	9	691	343	348	183	139	1.710	1.265	2.310	(343 vs 348)	183 vs 139	1.710 [1.265, 2.310]
2	8	399	203	196	96	92	1.011	0.679	1.505	(203 vs 196)	96 vs 92	1.011 [0.679, 1.505]
3	7	668	344	324	175	145	1.277	0.941	1.733	(344 vs 324)	175 vs 145	1.277 [0.941, 1.733]
3	6	423	203	220	104	86	1.591	1.079	2.347	(203 vs 220)	104 vs 86	1.591 [1.079, 2.347]
4	5	938	469	469	238	197	1.419	1.096	1.837	(469 vs 469)	238 vs 197	1.419 [1.096, 1.837]
4	4	153	78	75	41	34	1.336	0.707	2.527	(78 vs 75)	41 vs 34	1.336 [0.707, 2.527]
5	3	476	236	240	119	114	1.130	0.789	1.621	(236 vs 240)	119 vs 114	1.130 [0.789, 1.621]
5	2	372	188	184	100	64	2.113	1.387	3.218	(188 vs 184)	100 vs 64	2.113 [1.387, 3.218]
5	1	243	123	120	60	53	1.205	0.729	1.994	(123 vs 120)	60 vs 53	1.205 [0.729, 1.994]

DESCRIPTION OF VARIABLES

VAR : 1 TO 5 , all subjects, Age , Sex, Ethnic Origin, Region

CAT : 1 TO 10, subgroups

N: : Total number of patients

N1 : Number of patients in Treatment X group

N2 : Number of patients in Treatment Y group

NBEV1: Number of events in Treatment X group

NBEV2: Number of events in Treatment Y group

OR : Odds-ratio

ORLCL : Lower limit of odds-ratio

ORUCL : Upper limit of odds-ratio

NTEXT : Number of patients in Treatment X group versus Number of patients in Treatment Y group

NBEV : Number of events in Treatment X group versus Number of events in Treatment Y group

ORCI : Odds-ratio and Confidence interval

PROC FORMAT

```
proc format ;
VALUE VAR      1 ="All Subjects"
                2 ="Age"
                3 ="Sex"
                4 ="Ethnic Origin"
                5 ="Region"

;
VALUE CAT      10 ="All Subjects"
               9  "< 65 years"
               8  ">= 65 years"
               7  "Male"
               6  "Female"
               5  "Caucasian"
               4  "Non caucasian"
               3  "Western Europe"
               2  "Eastern Europe"
               1  "Other";

run;
quit;

data final ;
set final ;
format var var. cat cat. ;
      logOR=log (OR) ;
      logORLCL=log (ORLCL) ;
      logORUCL=log (ORUCL) ;
run ;
```

2- CREATION OF THE ANNOTATE DATASET

Annotate SAS Facility will be used to display text as well to draw lines in the required area of the graph. Some macros variables will be defined at first to fix the size and format of text.

* Initialization of annotate macros ;

```
%annomac ;
```

* Macro-variables for size and font of text :

```
%let size=1.2 ;
%let tsize=1.5 ;
%let ftitle='Arial' ;
%let ftext ='Arial' ;
```

* Macro for ticks on the horizontal axis ;

```
%macro tick(val=, VALlbl=_NO_, fontVAL=simplex, origin=0, tickl=-0.8, ticks=0.15) ;
```

* Area (Data): Absolute, values

```
xsys='2';ysys='2';
```

```
%move(&val,&origin);
```

* Area (Procedure output area): Relative, %.

```
xsys='B';ysys='B';
```

* Draw each tick from origin.

```
%draw(&origin,&tickl.,black,1,&ticks.);
```

`%cntl2txt;` (Copies the values of the internal coordinates (XLAST, YLAST) to the text coordinate (XLSTT, YLSTT).

```
%if "&VALlbl" eq "blank" %then %let val=" ";
%else %if "&VALlbl" ne "_NO_" %then %let val="&VALlbl.";
%else %let val="&val.";
%label(0,0,&val,black,0,0,1.2,&fontval,E);
%mend tick;
```

```
proc sort data=final;
    by var descending cat;
run ;
```

* Annotate dataset ;

```
DATA anno;
    length text catnobs varc. $100. ;
    SET final ;
    by var descending cat catgrpc;
    catnobs=compbl(put(cat,catg.)!!' '!!ntext);
    varc=trim(left(put(var,var.)));
```

1- HOW TO PUT TITLE AT THE TOP OF THE GRAPH AND TEXT AT THE BOTTOM OF THE GRAPH

* Area (Graphic output area): Absolute, %.

```
%system(3,3);
```

Draw the text by using the SAS function `%label` and the two arrows by using the SAS function `%line`. Several variables will be created in the annotate dataset: function variable, positional variables and attribute variables.

```
%label(2,88,'Subgroup (number of patients in Treatment X vs Treatment Y)'
,black,0,0,&size.,&ftext.,6) ;
%label(63,88,'Number of responses'
,black,0,0,&size.,&ftext.,6) ;
```

PhUSE 2008

```
%label(60,86.5,'Treatment X vs Treatment Y' ,black,0,0,&size.,&ftext.,6) ;
%label(83,88,'OR [95% CI]' ,black,0,0,&size.,&ftext.,6) ;

%label(20,3.5,'No Benefit under Treatment X' ,black,0,0,&size.,&ftext.,6) ;
%line(30,1.8,50,1.8,black,1,0.5) ;
%line(30,1.8,32,2.1,black,1,0.5) ;
%line(30,1.8,32,1.5,black,1,0.5) ;

%label(65,3.5,'Benefit under Treatment X' ,black,0,0,&size.,&ftext.,6) ;
%line(54,1.8,74,1.8,black,1,0.5) ;
%line(74,1.8,72,2.1,black,1,0.5) ;
%line(74,1.8,72,1.5,black,1,0.5) ;

%label(54,3.7,'OR and 95% CI ' ,black,0,0,&size.,&ftext.,5) ;
```

2- HOW TO PUT TICKS ON THE HORIZONTAL AXIS USING A LOGARITHMIC SCALE

In this example, each tick will be drawn from OR=-2.99 to 2.99 (=log(0.05) to log(20))

```
%dclanno;
hsys='3';

%tick(val=-2.99,VALlbl=0.05 ,fontVAL=&ftext);
%tick(val=-2.81,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-2.66,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-2.53,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-2.40,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-2.30,VALlbl=0.1 ,fontVAL=&ftext);
%tick(val=-1.61,VALlbl=0.2 ,fontVAL=&ftext);
%tick(val=-1.20,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-0.92,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-0.69,VALlbl=0.5 ,fontVAL=&ftext);
%tick(val=-0.51,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-0.36,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-0.22,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val=-0.11,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 0 ,VALlbl=1 ,fontVAL=&ftext);
%tick(val= 0.69,VALlbl=2 ,fontVAL=&ftext);
%tick(val= 1.1 ,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 1.39,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 1.61,VALlbl=5 ,fontVAL=&ftext);
%tick(val= 1.79,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 1.95,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 2.08,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 2.2 ,VALlbl=blank,fontVAL=&ftext,tickl=-0.5,ticks=0.15);
%tick(val= 2.3 ,VALlbl=10 ,fontVAL=&ftext);
%tick(val= 3.0 ,VALlbl=20 ,fontVAL=&ftext);
```

3- HOW TO DISPLAY THE SUB-GROUPS AND VARIABLES LABELS, HOW TO DRAW LINES BETWEEN LOWER AND UPPER CONFIDENCE LIMITS

*** To put sub-groups and variables labels ;**

Area :

```
%system(3,2);

if var^=1 then do ;
    if first.var then do;
        %label(5,cat+0.5,varc,black,0,0,&size.,&ftext.,6) ;
    end;
end ;

if ^(type=2 & cat in(8)) then do ;
    %label(10,cat,catnobs,black,0,0,&size.,&ftext.,6) ;
end ;
```

Math Font has to be used to put the symbol “ $\geq$ ” and not “ $\geq$ ” .

PhUSE 2008

```
else if (type=2 & cat=8) then do;
    %label(10,cat,'M',black,0,0,&size.,'Math',6) ;
    %label(12,cat,'65 years',black,0,0,&size.,&ftext.,6) ;
    %label(20,cat,ntext,black,0,0,&size.,&ftext.,6) ;
end ;
```

*** To draw a line between the Lower and Upper limit ;**

Area= Data, values

```
%system(2,2) ;
if OR^=. then do ;
    %line(logORLCL,cat,logORUCL,cat,black,1,0.5);
    %line(logORLCL,cat-0.05,logORLCL,cat+0.05,black,1,0.5);
    %line(logORUCL,cat-0.05,logORUCL,cat+0.05,black,1,0.5);
end ;
```

*** Put the number of events (ev1 vs ev2) and OR [CI] for each sub-group;**

```
%system(3,2);

%label(70,cat,nbev,black,0,0,&size.,&ftext.,6) ;
%label(82,cat,ORCI,black,0,0,&size.,&ftext.,6) ;
run ;
```

3- PROC GPLOT

The Annotate macro facility will be used to display text as well as to draw lines in the required area of the graph. Some macros variables will be defined at first to fix the size and format of text.

*** Initialization of annotate macros ;**

```
%let vsize=27.7 cm;
%let hsize=20.5 cm;
%let papersize=A4 ;

options orientation=portrait papersize="&papersize.";

title h=0.5 cm "options landscape goptions rotate=portrait" ;

goptions reset=all
    gunit=pct
    noborder
    htitle=2
    htext=&size.
    cback=w
    ftitle=&ftitle.
    ftext=&ftext.
    offshadow=(2,2)
    rotate=portrait
    hsize=&hsize.
    vsize=&vsize.
    autofeed;
```

*** Set-up Axes and Plots**

```
AXIS1 MINOR=NONE
      MAJOR=NONE
      ORDER=(0 TO 10 BY 1)
      LABEL=NONE
      value=NONE
      offset=(0,7);

AXIS2 MINOR=NONE
      MAJOR=NONE
      ORDER=(-4.38 to 4.38)
      LABEL=(h=3 c=white ' .')
      VALUE=none
      WIDTH=1 offset=(30,30) ;
```

PhUSE 2008

* Create symbol for odds-ratio

```
symbol11 i=none value=dot color=black ;

filename gr "&Flst_root.\Figure_X.emf ";

goptions gsfname=gr
          device=emf
          gsfmode=replace;
```

* Set-up titles and footnotes

```
*title1 j=1 h=&tsize. "Title 1 ;
*title2 j=1 h=&tsize. "Title 2" ;
*title3 j=1 h=&tsize. "Title 3";
*title4 ;
title5 j=1 h=&tsize. "Figure XXX: Odds-Ratios for Best Overall Response by subgroup
- xxx Population";
footnote1 h=1 " ";

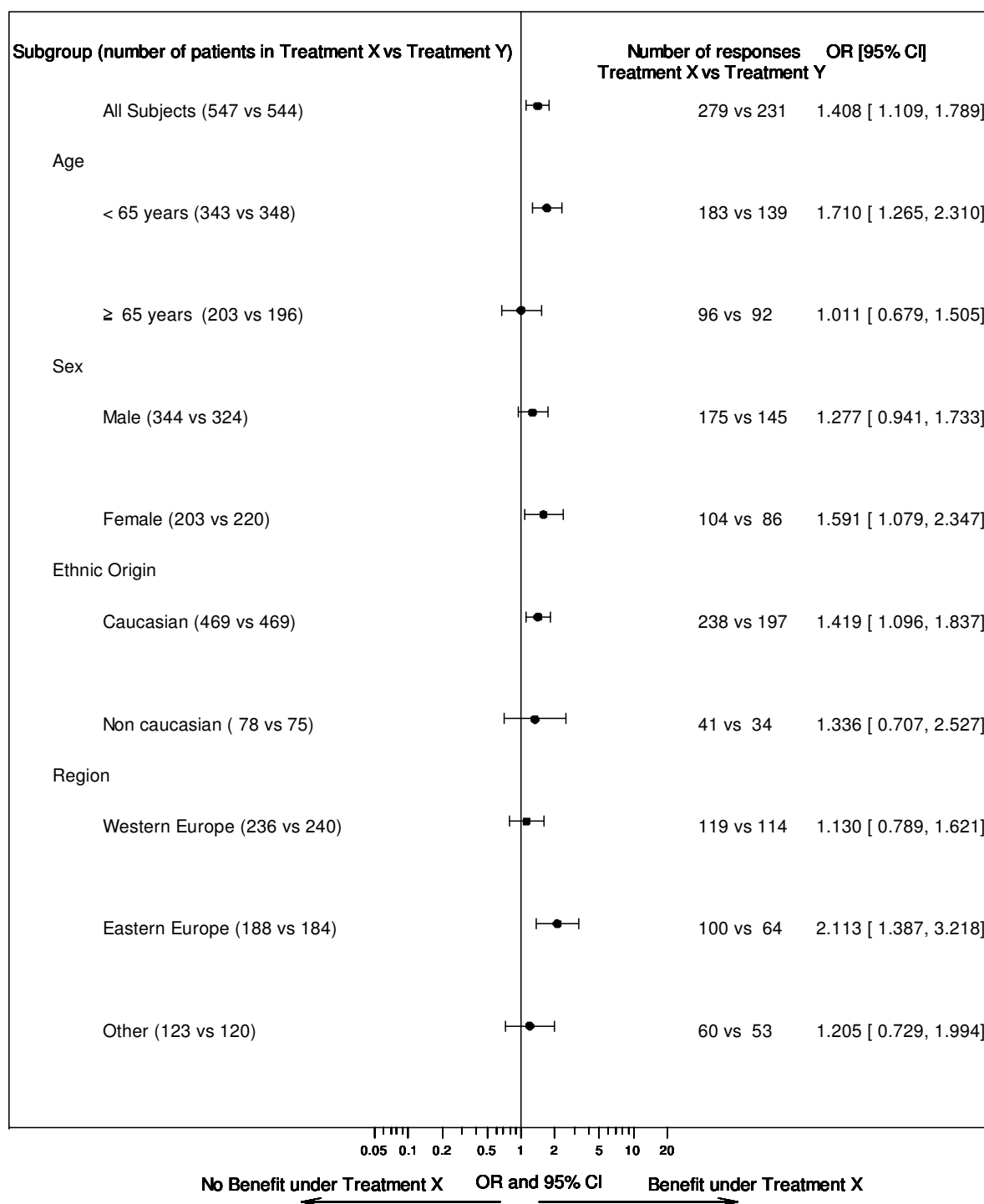
proc gplot data=final;
    plot cat*logOR/ VAXIS=AXIS1
                   HAXIS=AXIS2
                   HREF=0
                   NOLEGEND
                   ANNOTATE=anno;

run ;

QUIT;
```

4- FOREST-PLOT RESULTS

Figure XXX: Odds-Ratios for Best Overall Response by subgroup



CONCLUSION

Building a Forest-plot with the Annotate facility is not an easy task, fortunately the macro facility helps a lot and use of a logarithmic scale enables the display of extreme values for 95% confidence interval values.

The figure obtained allows the reader to visualize which odds-ratio is significant and hence determine which of the two treatments is better overall and also within each sub-group.

REFERENCES

SAS online V9.1.3

"How to annotate Graphics" PhUSE 2006 – Sandrine Stepien, Quintiles, Sidney, Australia

CONTACT INFORMATION (HEADER 1)

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

Laetitia Lemoine Le Halpert
Quintiles Paris
3-5 rue Maurice Ravel
92594 Levallois-Perret Cedex
Work Phone: +33 (0) 1 41 27 73 03
Fax: +33 (0) 1 41 27 72 72
Email: laetitia.lehalpert@quintiles.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.

Brand and product names are trademarks of their respective companies.