

A MACRO SYSTEM FOR CALLING PROC GPLOT FOR GENERAL USAGE BUT ALSO HAVING A MAJOR APPLICATION FOR PROCS LIFETEST AND PHREG PLOTS

Hugh Geary, Trilogy Consulting Corporation, Waukegan, IL

Abstract

This system enables the user to specify titles, labels, footnotes, symbols, input and output data set names, annotations and various control variables in the arguments of a calling statement without having to write the title, footnote and axis statements, etc. for PROC GPLOT each time a graph needs to be produced. The system consists of a general macro for setting up the call to PROC GPLOT and a special purpose macro, called by the general macro, to compute the coordinates and annotations (which in our case are for the Lifetest/Phreg graphs but could be for any application.)

The application to PROCS LIFETEST and PHREG plots the survival data for one, two, three or more groups while depicting censoring. Extra annotations have been computed to display the number-at-risk at each tic mark along the horizontal axis. Also in the general macro a moveable window has been created in which pertinent statistics can be prominently displayed.

The chief advantages of this macro are the ease of creating a large number of similar graphs and the ease of building a new graph by drawing upon a number of refined techniques already embodied in the code.

System Description

This system of two macros, `g_plot3` and `g_point3`, has grown through usage, being continually generalized to include more features. Figure 1, on the next page, shows a Kaplan-Meier Survival graph. It was produced by using an output data set from Procs LIFETEST & PHREG. `G_point3` was written specifically for processing LIFETEST output. `G_point3` could be replaced by other macros to produce other types of graphs. The main program,

`G_plot3`, creates titles, footnotes, notes, a panel for the display of certain statistics and a means to raise and lower the horizontal axis. The display panel can be raised, lowered or split in half and printed in small print or large. It also has the call to PROCS GPLOT and Greplay, a call to the system to send the completed graph to a printer and another to delete an unneeded data set;

The program had the 3 added to its name when it was expanded to plot 3 or more graphs. It can plot 2 or just 1 graph.

For the Kaplan-Meier graphs, four are actually being plotted besides the dashed horizontal line across the middle (it was drawn by creating a simple graph.) Two of the four are for censored patients one for each drug and two are for the noncensored patients, one for each drug.

The Number of Patients at risk comes from Proc LIFETEST and is calculated by algorithm embodied in `G_point3.mac`.

Below is a graph, an example of output from this system, and the macro call that produced it.

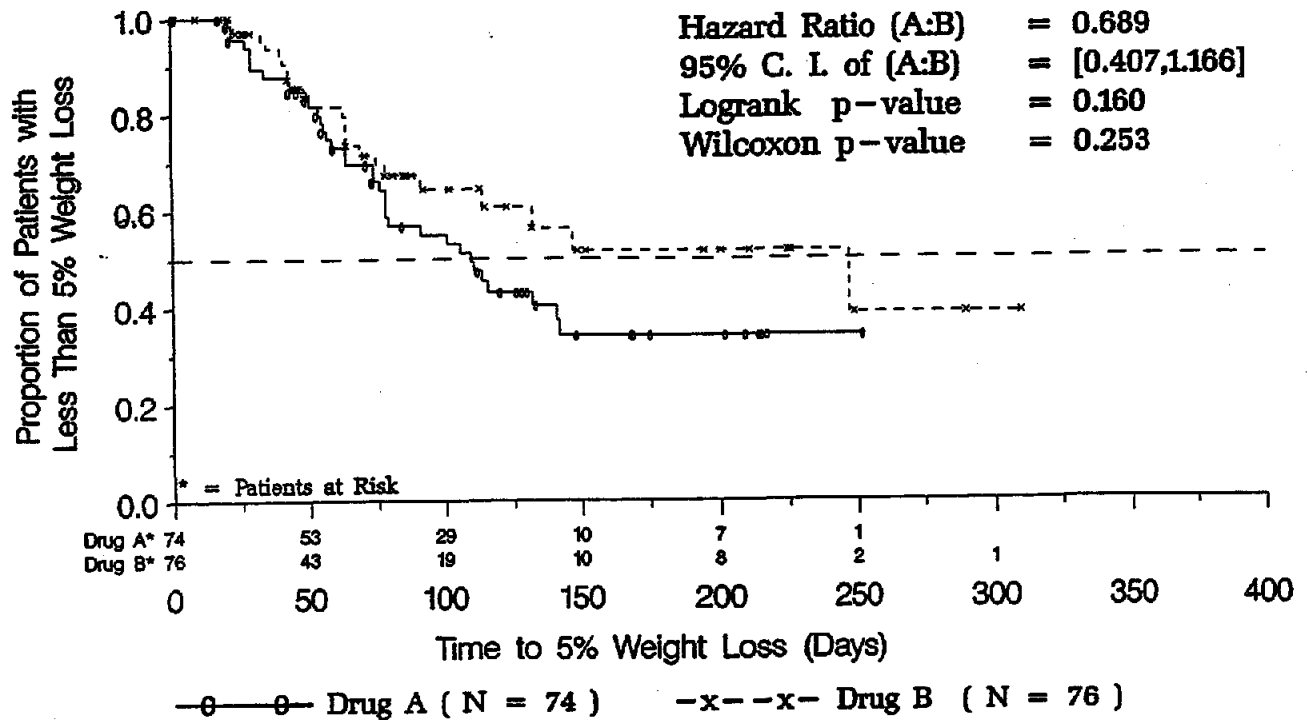
In the appendix is a listing of both macros interspersed with comments.

For more information/comments contact:

Hugh Geary
Trilogy Consulting Corporation
630 Morning Street
Worthington, OH 43085
(614) 885-3164

*SAS is a trademark or registered trademark of SAS Institute Inc. in the USA and other countries.

Figure 1
Time to Weight Loss of 5% from Baseline
Drug A vs. Drug B Patients



The Macro Call Producing the Above Plot

<pre> %g_plots (QL_wt1, disk11:[geary.epi],QL_wt1, days2, col, tr, Figure 1, Time to Weight Loss of 5% from Baseline, Drug A vs. Drug B Patients, '', Drug A:Drug B, 0.689, 0.407, 1.166, = 0.160, = 0.253, Drug A,74, Drug B,76,, %str(Proportion of Patients with), %str(Less Than 5% Weight Loss), Time to 5% Weight Loss (Days), 400, 50, print=no, box=BP, up=-4, N_a_R=zz); </pre>	The output graph data set name Directory and data set for input file Name of the Lifetest "time" variable Name of the Lifetest "group" variable Currently unused parameter Title1 Title2 Title3 Title4,Title5 & Title6 The following parameters are for display in thr statistics "box": Character string Hazard Ratio value Confidence interval bounds Logrank & Wilcoxon p-values Footnote: drug name & number of patients Footnote: drug name & number of patients Vertical axis label Vertical axis label (2nd line) Horizontal axis label Horizontal axis end point Horizontal axis tic mark spacing Print a listing of the input to Proc Plot the box can be in big print, regular print or omitted depending upon whether BP, RP or NO is given here up=n moves the "box" up n units The Number at Risk is printed is zz is given but is not printed otherwise
---	---

APPENDIX

```
*G PLOT3.mac Executes Proc GPLOT, calls G Point3;
*****
** Program      G_PLOT3.mac
** Purpose      Create Kaplan-Meier Survival Graphs
**              The calling program provides titles,
**              statistics and the name of a
**              Lifetest output file containing
**              "At Risk" info & co-ordinates
** Programmer    Hugh Geary
** Date          10July92, Feb94, 20May94, 14Oct94;
*****
%include '[geary.adr]g_point3.mac';
```

```
%macro g_plot3 (out, dir, in, time, group, TR,
titl1, titl2, titl3, titl4, titl5, titl6, l_hr, hr,
cil, ci2, lr, gw, l_n1, n1, l_n2, n2, l_n3, n3,
csno, labell, labella, label2, x_max, x_incr,
print=no, y_max=1, y_incr=.2, N a R=, box=RP, up=0,
FET=, u1=, v1=, u2=, v2=, X0=0, E a R1=10000,
E a R2=0, Rn1=, Rn2=, Rn3=, X_minor=(N=1), l_n4=,
n4=., l_n5=, n5=., Rn4=, Rn5=, u3=, v3=, u4=, v4=,
u5=, v5=);
```

* Note: &label2 is being used in footnotel
instead of in the x-axis label;

* Note: Group can be the value 0, the value 1,
or the name of a variable which takes
on the values 0 or 1;

```
FILENAME GRAPH "%out..GSF";
GOPTIONS DEVICE=TEKLN01 DISPLAY NOPROMPT GSFNAME=
GRAPH GSFMODE=REPLACE vpos=33 hpos=71;
*
an option here could be gunit=pct;
LIBNAME IN "%dir.";
```

```
%g_points (&in,final,AN)
```

```
*proc print data=&in.&in;
* title "%in (Output from PHREG)";
*proc print data=final;
* title 'Final (Input to Proc GPlot)';
*proc print data=AN;
* title 'An - Annotation Input';
```

* Below we create the "box" for the display of the
Hazard Ratio, etc. statistics with variations in
the size of the print available;

* RP-Regular Print, BP-Big Print,
BWP-Big Wide Print, NO-No Stats Printed;

```
data _null_;
if "%upcase(&box)"="RP" %then %do;
h=.7; x1=49.5; x2=64;
y1=76.75+&up; y2=74.5+&up; y3=72.25+&up;
y4=70+&up; y5=67.75+&up; %end;
%else if "%upcase(&box)"="BP" %then %do;
h=1.; x1=39.5; x2=60;
y1=74+&up; y2=70+&up; y3=66+&up;
y4=62+&up; y5=58+&up; %end;
%else if "%upcase(&box)"="BWP" %then %do;
h=1.; x1=43; x2=62;
y1=78; y2=74; y3=70; y4=66; y5=62; %end;
call symput ('h',put(h,3.1));
call symput ('x1',put(x1,5.2));
call symput ('x2',put(x2,5.2));
call symput ('y1',put(y1,5.2));
call symput ('y2',put(y2,5.2));
call symput ('y3',put(y3,5.2));
call symput ('y4',put(y4,5.2));
call symput ('y5',put(y5,5.2));
run;
```

* NEXT create the title placements and do it so
that if a title line is blanked out, the ones
below it will be moved up to fill its space;

```
title f=zapf h=1 "atitl1";
title2 f=zapf h=1 "atitl2";
TITLE3 F=ZAPF H=1 "atitl3";
TITLE4 F=ZAPF H=1 "atitl4";
title5 f=zapf H=1 atitl5;
title6 f=zapf H=1 %str(atitl6);
* It wants quotes on stitl unless it is null;
```

* Title numbers 7 to 10 are used for the "box";

* If fewer than 5 of the 6 allotted titles are;
* used, then the title numbers used for the ;
* Statistics Box are changed from 7-10 to ;
* smaller numbers;

* When fewer titles are created, the plot ;
* can expand over the additional unused space;

```
%do; %let tn7=7; %let tn8=8; %let tn9=9;
%let tn10=10; %end;
%if "%atitl4"="" & "%atitl5"="" & "%atitl6"="" %then
%do; %let tn7=6; %let tn8=7; %let tn9=8;
%let tn10=9; %end;
%if "%atitl3"="" & "%atitl4"="" & "%atitl5"="" &
"%atitl6"="" %then
%do; %let tn7=5; %let tn8=6; %let tn9=7;
%let tn10=8; %end;
```

```
%if "&hr"="" %then %do;
%let tn10=&tn9; %let tn9=&tn8; %let tn8=&tn7;
%let y5=&y4; %let y4=&y3; %let y3=&y2;
%let y2=&y1; %end;
```

```
run;
```

```
%if "&cil"="" %then %do;
%let tn10=&tn9; %let tn9=&tn8;
%let y5=&y4; %let y4=&y3; %let y3=&y2; %end;
```

```
%if "%upcase(&box)"="NO" %then %do;
TITLE&tn7 F=ZAPF H=&h. M=(&x1,&y1. pct)
"Hazard Ratio (&l hr.)"
M=(&x2,+0) "=" &hr";
TITLE&tn8 F=ZAPF H=&h. M=(&x1,&y2. pct)
"95% C.I. of (&l hr.)"
M=(&x2,+0) "=" [&ci1,&ci2]";
TITLE&tn9 F=ZAPF H=&h. M=(&x1,&y3. pct)
'Logrank p-value' M=(&x2,+0) "&lr";
```

```
TITLE&tn10 F=ZAPF H=&h. M=(&x1,&y4. pct)
'Wilcoxon p-value' M=(&x2,+0) "&gw";
%end;
```

```
%if &N a R= %then %do;
%let ANNO AN=;
%let ax2_val=(F=SWISSL);
%let ax2_lab=none; *(F=SWISSL "&label2");
%end;
%else %do;
%let ANNO AN = %str(ANNO=AN);
%let ax2_val=none;
%let ax2_lab=none;
%end;
%let len_X_ax = %eval(&x_max - &X0);
```

* For cases of footnote/no footnote, At Risk/
no At Risk use the following to Raise the AXES
by 2 lines to provide "At Risk" Annotation space;
* Footnote2 will be generated <=> u1,v1,u2,v2
are not used and conversely;

```
* When a footnote is too long, use h=.9 to shorten
it (also need m=(0,0));
```

```
%if &N a R= & sul= %then
%do; %let up=+4; %let dn=-3; %end;
%if &N a R= & sul= %then
%do; %let up=+2; %let dn=-1.75; %end;
%if &N a R= & sul= %then
%do; %let up=+.4; %let dn=+0; %end;
%if &N a R= & sul= %then
%do; %let up=+1.4; %let dn=+0; %end;
```

```
run;
Footnote1 f=swissl h=1 J=C ' ' m=(+0,&up) " "
m=(+0,&dn) "&label2";
```

```
* Using a "jump" to provide room for the At Risk
annotations;
```

```
%if &ul= %then %do;
Footnote2 f=zapf J=R
draw=(2.8,.95,9.6,.95,2.8,.90,9.6,.90)
' ' m=(+0,+2) m=(+0,-.18)
h=.85 " 0 0 &l n1 ( N = &n1 )"
" -x-x- &l n2 ( N = &n2 )"
m=(+1.2,+3) h=.20 " - - - "
h=.85 m=(-4.25,-.45) " * *"
m=(+0,+1.5) " &l n3 ( N = &n3 )";
%end;
```

```
%if "&label2"=" " %then %let just=;
%else %let just=;%unquote(%str(J=C));
```

```
data _null_;
half=.5*&y_max;
call symput('midline',put(half,7.2));
call symput
('nam','y'||trim(left(put(int(time()),5))));
* NOTE: nam, the name of the GOUT= dataset must
not contain any underscores _;
```

```
PROC GPLOT DATA=FINAL GOUT=&nam.;
%if "&sfet"=" " %then %do;
NOTE F=ZAPF H=&h. M=(&x1,&y5. pct)
'FET p-value' M=(&x2,+0) "&sfet"; %end;
%if &N a R= %then %do;
NOTE F=ZAPF H=.6
M=(10.5,8.8) '* = Patients at Risk'; %end;
```

```
*G POINT3.mac Called by G_plot3 for computations;
*****
** Program G_point3.mac
** Purpose To create the input data set for
** PROC GPLOT (point coordinates),
** compute the Number of Patients at
** Risk from LIFETEST output and
** and create the Annotation data set
** used on the /VAXIS statement;
** Programmer Hugh Geary
** Date July 1995
*****
```

```
* macro g_points, when passed survival data,
computes the co-ordinates for the
censored (a line graph) and the
non-censored (a points only graph) patient
graphs for each of two drugs (4 graphs) and a
5th graph, a horizontal mid-point line;
* G_points also computes the Number at Risk data
and its display coordinates.;
```

```
%if &ul= %then %do;
NOTE F=ZAPF H=.7 M=(&ul,&v1) "&l n1"
M=(&ul,&v2) "&l n2"; %end;
```

```
AXIS1 VALUE = (F=SWISSL)
LABEL = (F=SWISSL A=90 R=0 "&label1"
%unquote(&just) %unquote("&label2"))
ORDER = 0 to &y_max by &y_incr
MINOR = (N=1)
OFFSET = (0);
AXIS2 VALUE = &x2_val
LABEL = &x2_lab
ORDER = 0 to &len_X_ax by &x_incr
MINOR = &x_minor
OFFSET = (0);
```

```
symbol1 v=none i=join line=1;
symbol2 v=0 i=none h=.4;
symbol3 v=none i=join line=2;
symbol4 v=X i=none h=.4;
symbol5 v=none i=join line=33 width=2.0;
symbol6 v=Star i=none h=.4; * 3 line asterisk;
symbol7 v=none i=join line=43 width=2.0;
symbol8 v=$ i=none h=.4; * diamond shape;
* more symbol statements may be added here;
```

```
plot survival*&time=plotline
/vaxis=axis1 haxis=axis2 vref=&midline
lvref=20 kanno_an nologend;
```

```
PROC GREPLAY IGOUT=&nam. NOFS;
TC PAT1;
TDEF ATEMP1 DES = 'GRAPHING TEMPLATE'
1/LLX=0 ULX=0 URX=100 LRX=100 SCALEX=.925
LLY=0 ULY=100 URY=100 LRY=0 SCALEY=.925;
TEMPLATE ATEMP1;
TREPLAY 1 : 1;
```

```
run;
x "delete &out..gsf;-1"; ** works fine now
with some system parameters reset;
quit;** the quit enables the next command to work;
** without quit we used to get
"- dataset currently in use";
x "print/que=Ion2/notify/noflag/nofeed &out..gsf;";
%mend g_plot3;
```

```
%macro g_points (inn,outt,AN);
```

```
data &inn; set in.&inn;
* y_max is 1 for proportions, 100 for percents;
survival=survival*&y_max;
&time =&time<&x_max;
* truncating any large values;
* x_max is also used in the axis2 statement and to
generate the midway line;
if &time>&x0 then &time=&time - &x0 +.1;
if &time=0 & lag(&time)=0 then &time=.05;
* When the second time value is also zero, it
needs to be given a slightly positive value;
```

```
proc sort data=&inn; by &group;
```

```
data &outt (keep=&group plotline &time survival);
set &inn; by &group;
```

```
* Create the even numbered plotlines. They consist
of just the points labeled by the symbols such as,
0,x,* etc.;
```

```

if censor=1 then do;
  plotline=2*%group;
  output;
end;

* Create the unlabeled plotlines: 1,3,...;
* They consist of the vertical & horz. lines;

plotline=2*%group - 1;
curr=survival;
prev=lag(survival);

* Next output twice except for the first point of
each odd numbered plotline;
if ~first.%group then do;
  survival=prev;
  output;
end;
survival=curr;
output;

*proc print data=%outt;* title "%outt";

*****;
*The following code to calculate 'Number of
Patients at Risk' uses an output data set
from LIFETEST;

proc sort data=%inn
  out=xx2 (keep=atime %group _censor_ survival);
  where %time>.;
  by %group %time;

data for %AN (keep=%group t00 num_left);
set xx2_end=last; by %group;
retain x_incr 0 beta t00 num_left;

* Here the sur time points to be used in the
annotation coordinate display along the bottom
of the plot are selected;

if first.%group then do;
  %do ii = 1 %to 5;
    if %group=%ii then beta=%ann%ii;
  %end;
  num_left=beta;
  x_incr=%eval(%x_incr);
  t00=%x0; * the 1st annotated major tic mark;
end; * is placed directly below the Y axis;

do while (t00<%time + .1 & num_left>0 &
  t00<=%x_max);

* Outputting annotations at every major tic mark;
output;

* Computing the next annotated major tic mark;
t00=t00 + x_incr;
if %E_a_R1 + x_incr/2 <= t00 <= %E_a_R2 + x_incr
  then t00=t00 - x_incr/2;

* Extra annotations are added at the midpoints
from E_a_R1 thru E_a_R2;
end;

num_left=beta*survival - _censor_;
if survival>0 then beta=num_left/survival;

*proc print data=for_%AN;* title "for_%AN";

```

```

*****;
data %AN (keep=x y text function size xsys
  ysys style);
set for %AN end=last; by %group;
length function $8 text $16;
retain y;

function='LABEL'; size=0.6; xsys='4';
ysys='4'; style='swiss1';

* This data step creates a # At Risk Line for each
graph (or group or strata). Below, "text"
contains the label for the # At Risk Line;

if first.%group then do;
  * Get the label for the # At Risk Line;
  %do ii = 1 %to 5;
    if %group=%ii then do;

      text="%scan(%&1 %ann%ii,1,','))" || '**';
      y=8.1 - .8*%ii; * i.e. y= 7.3, 6.5 & 5.7;

      %if %&Rn%ii= %then %do;

* Revise num_left, the # that are still at risk
at the point t00=%x0, if necessary;
* Rn%ii has the revised value;

      %let n%ii=%&Rn%ii;
      num_left=%&Rn%ii;
      %end;
    end;
  %end;

  ltext=1 + max(length("%scan(%&1 %n1,1,','))",
    length("%scan(%&1 %n2,1,','))",
    length("%scan(%&1 %n3,1,','))",
    length("%scan(%&1 %n4,1,','))");

  x=10 - size*(ltext + 1 +
    (%length(%&n1)<>%length(%&n2)<>%length(%&n3)))/2;

* Output the At Risk label - once for each group;
output;
end;

* It is useful to know that the Annotate routine
will center "text" in its char10 field;

x = 10 + 59.5*(t00 - %x0)/%eval(%x_max - %x0);
text= trim(left(put(num_left,5)));

output; * Outputting the actual # At Risk;

* The Annotated points are all multiples of x_incr.
Create the VALUE labels for the X-Axis Tic Marks.
The regular values were SUPPRESSED to make room
for the "At Risk" annotations;

if last then do;
  size=1.0;
  %do ii=%x0 to %x_max by %x_incr;
    x=10 + 59.5*(ii - %x0)/%eval(%x_max - %x0);
    y=4.5;
    text=trim(left(put(ii,5)));
    output;
  end;
end;

run;
*proc print data=%an; title "%AN";
%mend %points;

```