

On customizing survival curves in SAS®

Michael J. Williams, ICON Clinical Research, San Diego, CA

ABSTRACT

In clinical research, survival analysis plays a major role, and the choice of visual tool is typically the survival plot. This paper will cover ways to enhance survival plots in SAS®. The primary focus will be on the Kaplan-Meier product-limit survival curve for a right-censored survival model. We will discuss the modification of the PROC LIFETEST graph template to customize Kaplan-Meier plots following a well-known approach by Warren Kuhfeld and Ying So. Another approach utilizes a combination of ODS OUTPUT statements for PROC LIFETEST or PROC PHREG, followed by DATA steps to create a dataset that can be graphed via PROC SGPLOT.

BASIC CONCEPTS

In this paper, we focus on “time to event” data. Examples of events are death, relapse, or recovery. Typically, the event is called a “failure” whether or not the event is actually negative. The absence of an event is “survival”. If a subject is lost to follow-up, then the subject is called “censored”.

The Survival Function is defined as $S(t) = \Pr(T > t)$, and the Failure Function is $F(t) = 1 - S(t)$.

KAPLAN-MEIER ESTIMATE

Let $0 < t_1 < \dots < t_D$ be distinct event times.

For each $j = 1, \dots, D$,

- Y_j = number of subjects at risk at time t_j
- d_j = number of events at time t_j
- $\Pr(\text{survival at time } t_j \mid \text{at risk at time } t_j) = (Y_j - d_j)/Y_j$

The Kaplan-Meier estimate of the survival function is

$$\hat{S}(t) = \prod_{t_j \leq t} (Y_j - d_j)/Y_j.$$

EXAMPLE DATA

SASHELP.BMT is a patient-level data set containing 137 records for bone marrow transplant patients in a study by Klein and Moeschberger (1997).

At the time of transplant, the patients were categorized into risk groups

1. ALL (acute lymphoblastic leukemia)
2. AML (acute myelocytic leukemia) – Low Risk
3. AML (acute myelocytic leukemia) – High Risk

The endpoint is disease-free survival time: time (in days) until death, relapse, or end of study. The SAS variables are

- T = disease-free survival time
- Status = event indicator: 1= event, 0= censored
- Group = risk group: 1= ALL, 2= AML-Low Risk, 3= AML-High Risk

The SASHELP.BMT dataset will be modified for use in this paper as follows.

```

proc format;
  invalue bmtnum  'ALL'=1  'AML-Low Risk'=2  'AML-High Risk'=3;
  value  bmtfmt  1='ALL'  2='AML-Low Risk'  3='AML-High Risk';
run;

data bmt(drop=g label="Bone Marrow Transplant Patients");
  set sashelp.bmt(rename=(group=g));
  Group = input(g, bmtnum.);
  format group bmtfmt.;
run;

```

PROC LIFETEST KM PLOT: DEFAULT GRAPH TEMPLATE

The code below produces a KM plot of the BMT data with the default graph template (see Figure 1):

```

proc lifetest data=bmt
  plots=survival(nocensor test
    atrisk(atrisktickonly maxlen=13)=0 to 2500 by 500);
  time T*Status(0);
  strata Group / order=internal;
run;

```

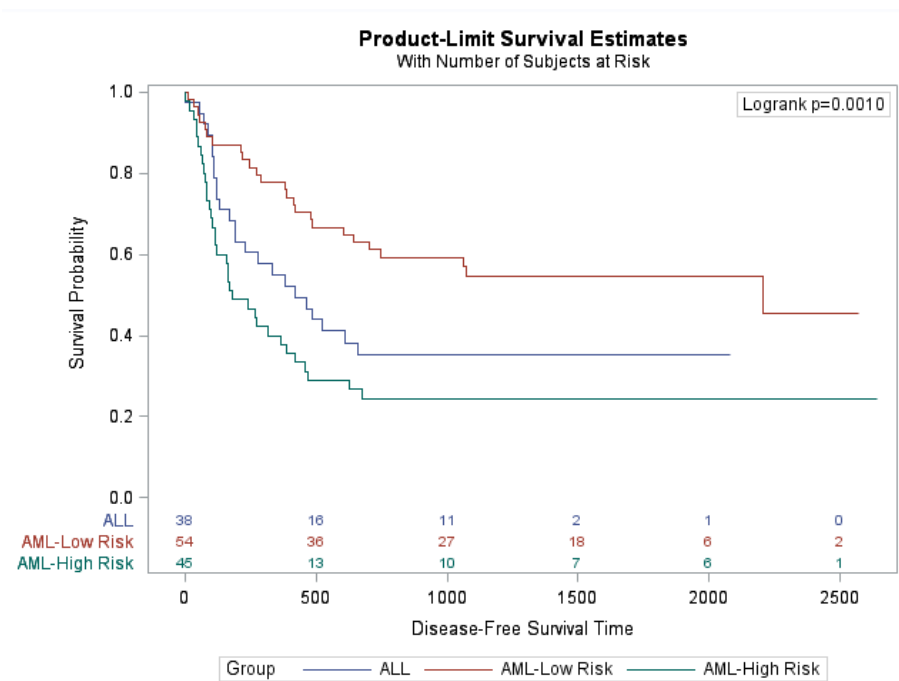


Figure 1. Kaplan-Meier curve with the default graph template.

PROC LIFETEST KM PLOT: MODIFIED GRAPH TEMPLATE

Macros are provided (to be saved as a .sas file) by SAS at

http://support.sas.com/documentation/onlinedoc/stat/ex_code/142/templft.html

The main macros are

- %ProvideSurvivalMacros
- %CompileSurvivalMacros

There are several macros and macro variables that you can modify within those macros to customize your template. The macros modify your user template until you restore the default. The Workflow for Graph Template Modification is given below.

- Run %include "<your path to the macro file>".
- Run %ProvideSurvivalMacros.
- Redefine macros and/or macro variables within %ProvideSurvivalMacros.
- Run %CompileSurvivalMacros.
- Run PROC LIFETEST.
- Restore the original template by running the following code:

```
proc template;
  delete Stat.Lifetest.Graphics.ProductLimitSurvival / store=sasuser.templat;
  delete Stat.Lifetest.Graphics.ProductLimitSurvival2 / store=sasuser.templat;
run;
```

THE CUSTOMIZED GRAPH

Once the template has been altered, we run the same PROC LIFETEST code as before. This time, the result is the customized graph (see Figure 2). The code that was implemented to alter the graph template is illustrated the following subsections.



Figure 2. Modified Kaplan-Meier curve obtained by modifying the graph template.

CHANGE TITLE

```
%macro StmtsBeginGraph;
  Entrytitle "Kaplan-Meier Plot"
  / textattrs=GraphValueText textattrs=(size=12pt weight=bold);
  Entrytitle "By Disease Group"
  / textattrs=GraphValueText textattrs=(size=10pt);
%mend StmtsBeginGraph;
```

CHANGE THE LEGEND AND INSET POSITION

```
%let InsetOpts = autoalign=(BottomLeft) Opaque=true
                 border=true BackgroundColor=GraphWalls:Color;
%let LegendOpts = autoalign=(TopRight) Opaque=true
                  title="Disease Group" location=inside across=1;
```

CHANGE LINE PATTERNS/COLORS

```
%let GraphOpts = attrpriority=none
                  DataLinePatterns=(Solid ShortDash MediumDash)
                  DataContrastColors=(red green blue)
                  DataColors=(red green blue);
%let StepOpts = lineattrs=(thickness=2.5);
```

ADD REFERENCE LINES AND SUMMARY TABLE

```
%macro StmtsTop;
  referenceline y=.75;
  referenceline y=.50;
  referenceline y=.25;
%mend;
%SurvivalSummaryTable
```

MODIFY P-VALUE INFO AND ATRISK TABLE

```
%macro pValue;
  if (PVALUE le .01) entry "Log Rank p " eval (PUT(PVALUE, PV42f.));
  else entry "Log Rank p = " eval (PUT(PVALUE, PV42f.));
  endif;
%mend;

proc format;
  value PV42f
    0-0.01 = "(*ESC*){unicode '2264'x} 0.01"
    0.01<-1 = [4.2];
run;

%let ClassOpts = class=CLASSATRISK;
```

X-AXIS OPTIONS

```
%let xOptions = label="Time from Enrollment (Days)"
                 shortlabel=XNAME
                 offsetmin=.075
                 labelattrs=(size=10pt weight=bold)
                 tickvalueattrs=(size=10pt)
                 linearopts=(viewmax=2700)
                 tickvalueelist=XTICKVALS
                 tickvaluefitpolicy=XTICKVALFITPOL);
```

Y-AXIS OPTIONS

```
%let yOptions = label="Disease-Free Survival"  
                shortlabel="Survival"  
                labelattrs=(size=10pt weight=bold)  
                tickvalueattrs=(size=8pt)  
                linearopts=(viewmin=.15 viewmax=1  
                tickvaluelist=(.2 .3 .4 .5 .6 .7 .8 .9 1)  
                tickvalueformat=percent6.0);
```

CUSTOMIZED PLOT VIA PROC SGPLOT

This time, use ODS OUTPUT datasets produced by PROC LIFETEST.

```
ods output SurvivalPlot=sp1 HomTests=pvall;  
  
proc lifetest data=bmt plots=survival  
    (test atrisk(atrisktickonly outside maxlen=13)=0 to 2500 by 500);  
    time T*Status(0);  
    strata Group / order=internal;  
run; quit;  
  
data _null_;  
    set pvall(where=(lowcase(test) eq "log-rank"));  
    call symputx("LRPV",put(ProbChiSq,6.4),"G");  
run;
```

As a warmup, we can recreate the default plot with the following code.

```
ods select all;  
title "Product-Limit Survival Estimates";  
title2 height=0.8 "With Number of Subjects at Risk";  
proc sgplot data=sp1;  
    step x=time y=survival / group=stratum name='s';  
    keylegend 's' / location=outside title="Group";  
    yaxis values=(.0 .2 .4 .6 .8 1);  
    inset "Log-Rank p = &LRPV." / position=top border;  
    xaxistable atrisk / x=tatrisk class=stratum colorgroup=stratum  
        location=outside;  
    scatter x=time y=censored / markerattrs=(symbol=plus) name='c';  
    scatter x=time y=censored / markerattrs=(symbol=plus) group=stratum;  
    keylegend 'c' / location=inside position=topright type=markersymbol;  
run;
```

The modified plot in Figure 3 is then created by the following code.

```
proc sgplot data=sp3 noautolegend;
  styleattrs DataLinePatterns=(Solid ShortDash MediumDash)
  DataContrastColors=(red green blue) DataColors=(red green blue)
  axisbreak=slantedright;

  step x=time y=survival / group=stratum name='s' lineattrs=(thickness=2);

  xaxis label="Time from Enrollment (Days)" labelattrs=(weight=bold);

  yaxis label="Survival Percentage" labelattrs=(weight=bold)
  values=(.0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1) valuesformat=percent6.0
  ranges=(-.01-.05 .18-1);
  refline .75 .50 .25 / transparency=.2 axis=Y;

  keylegend 's' / title="Disease Group" linelength=30 location=inside
  position=topright across=1;

  inset "Log-Rank p (*ESC*){unicode '2264'x} 0.01" / position=top border;

  xaxistable atrisk / x=tatrisk class=stratum colorgroup=stratum location=inside
  title="Number at risk"
  titleattrs=(family=arial style=italic weight=bold)
  valueattrs=(size=9);

  scatter x=time y=censored / markerattrs=(symbol=circlefilled) name='c';
  scatter x=time y=censored / markerattrs=(symbol=circlefilled) group=stratum;
  keylegend 'c' / location=inside position=bottom type=markersymbol;
run;
```

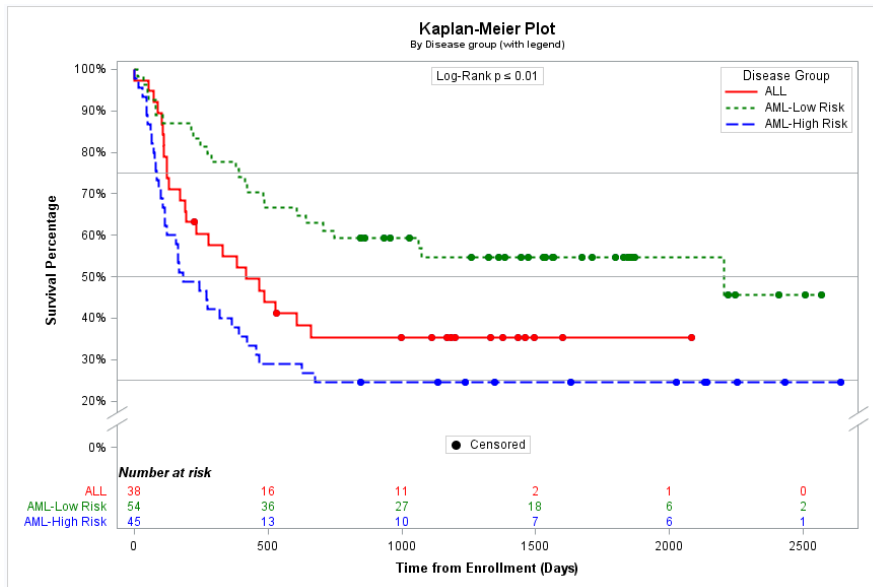


Figure 3. Customized KM plot obtained via PROC SGPLOT (with legend).

If we omit the KEYLEGEND statement, and revise the STEP statement, then we have curve labels without a legend. See Figure 4.

```
step x=time y=survival / group=stratum name='s' lineattrs=(thickness=2)
  curvelabel curvelabelattrs=(family=arial size=9 style=italic weight=normal)
  curvelabelpos=end splitchar='-';
```

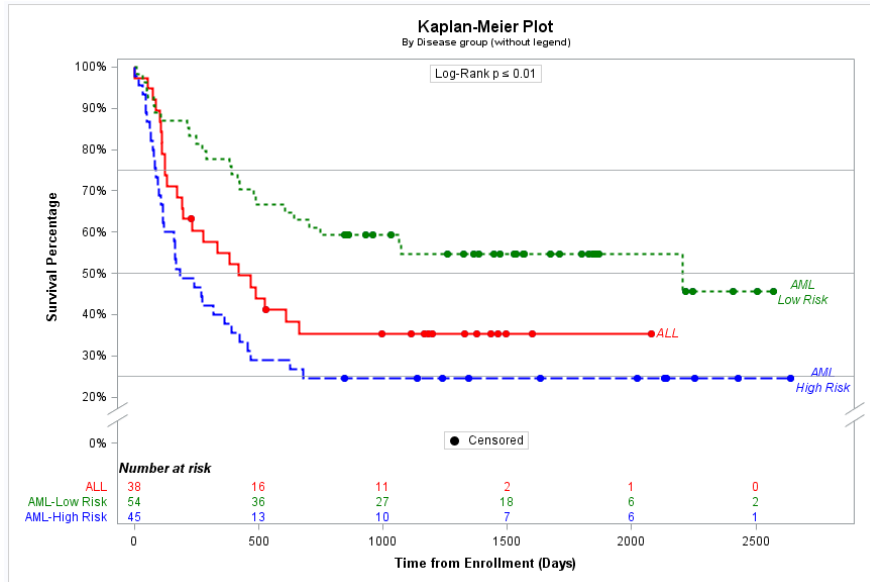


Figure 4. Customized KM plot obtained via PROC SGPLOT (with curve labels).

TAKING IT FURTHER

With more work, we can get add features such as

- quartile survival with confidence limits (from PROC LIFETEST),
- unadjusted hazard ratios with confidence limits (from PROC PHREG),
- markers for intersection points with the reference lines.

See Figure 5. The main idea is that more rows can be added to SP1 by merging in the following ODS OUTPUT datasets

- QUANTILES from PROC LIFETEST,
- HAZARDRATIOS from PROC PHREG.

The PROC PHREG code that produces the unadjusted hazard ratios is given below. Here we set “AML-Low Risk” (group=2) as the reference group.

```
proc phreg data=bmt;
  class group(ref='2') / param=ref;
  model t*status(0) = group / ties=breslow;
  hazardratio group / diff=ref;
run;
```

In PROC SGPLOT, use a YAXISTABLE statement to include the new data. For brevity, the details are omitted.

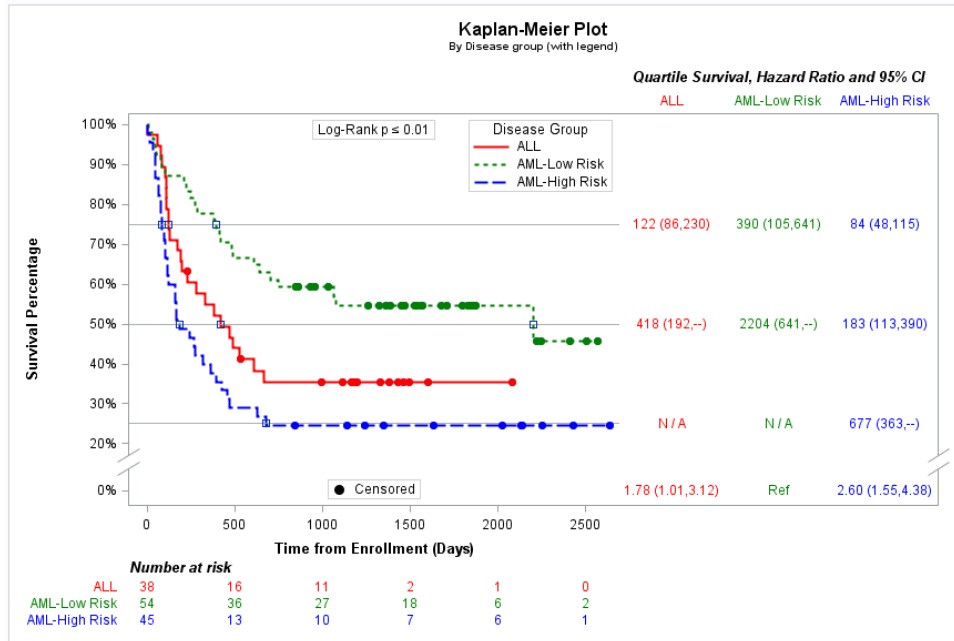


Figure 5. Kaplan-Meier plot including quartile survival times with confidence limits and unadjusted hazard ratios with confidence limits. Also, intersection points of curves with the quartile reference lines are marked.

RECOMMENDED READING

SAS/STAT 14.3 User Guide Chapter 23. Customizing the Kaplan-Meier Survival Plot.

SAS 9.4M5 documentation for PROC SGPLOT.

Sanjay Mantange, "Survival Plot", Graphically Speaking blog on SAS.com, February 9, 2014.

CONTACT INFORMATION

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

Name: Michael J. Williams
 Enterprise: ICON Clinical Research
 Address: 3636 Nobel Drive, Suite 340
 City, State ZIP: San Diego, CA 92122
 E-mail: michael.williams@iconplc.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.

Other brand and product names are trademarks of their respective companies.