PharmaSUG 2017 - Paper HT03

Survival 101 - Just Learning to Survive

Leanne Goldstein and Rebecca Ottesen, City of Hope, Duarte, CA

ABSTRACT

Analysis of time to event data is common in biostatistics and epidemiology but can be extended to a variety of settings such as engineering, economics and even sociology. While the statistical methodology behind time to event analysis can be quite complex and difficult to understand, the basic survival analysis is fairly easy to conduct and interpret. This workshop is designed to provide an introduction to time to event analyses, survival analysis and assumptions, appropriate graphics, building multivariable models, and dealing with time dependent covariates. The emphasis will be on applied survival analysis for beginners in the health sciences setting.

INTRODUCTION

In most statistics academic settings, Survival Analysis is a one to two semester course that is dense with methodology and notation. In this presentation we offer a simple way to approach time to event data which can be useful to those who have never taken a course in survival analysis or for those who need a bit of a refresher. This presentation begins with examining the data and deciding between parametric (Accelerated Failure Time) and semi-parametric (Cox Proportional Hazards) models. We will cover testing interactions and time-dependent covariates as well as how to verify that the model assumptions are met. We will also demonstrate useful graphing techniques for generating presentation or manuscript quality survival plots. This workshop uses the German Breast Cancer Study (GBCS) Data from Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition (Hosmer, Lemeshow, and May 2008) and evaluates covariates that are predictive of time to breast cancer recurrence. This GBCS data set can be found at the textbook site ftp://ftp.wiley.com/public/sci_tech_med/survival.

EXAMINING SURVIVAL DATA - KAPLAN-MEIER METHOD

First, we examine the data using PROC LIFETEST. This procedure can be used to compute survival estimates using the Kaplan-Meier method and the actuarial method. In our example, we study the `rectime` variable, which estimates for each patient, the number of days from date of diagnosis to an event, which in this disease free survival analysis is date of recurrence (or death), or number of days to date of last contact. If the date of last contact is used, and the patient has not had a recurrence during the time they were observed, then the patient's data is considered right censored. The Kaplan-Meier method will be used to estimate the cumulative proportion of patients surviving over time due to the censored observations not being organized in specific time intervals (more typical for the actuarial method).

```
PROC LIFETEST DATA = gbcs;
  TIME rectime*censrec(0);
RUN;
```

The first line of PROC LIFETEST includes a specification of the dataset. Next we use the TIME statement. In this statement, we write the time to recurrence (days) variable `rectime` and the censoring variable `censrec.` The `censrec` variable is coded 1 if a patient had an event and 0 when the patient is censored, i.e. has had no recurrence during the time observed. We put 0 in parentheses to indicate the value for the censored patients in the study. The output of PROC LIFETEST gives the following output shown in Output 1.

The LIFETEST Procedure

Product-Limit Survival Estimates

rectime		Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.00		1.0000	0	0	0	686
8.00	*	.	.	.	0	685
15.00	*	.	.	.	0	684
16.00	*	.	.	.	0	683
1806.00		0.5032	0.4968	0.0226	282	131
1807.00		0.4994	0.5006	0.0227	283	130
1807.00	*	.	.	.	283	129
1814.00		.	.	.	284	128
1814.00		0.4916	0.5084	0.0230	285	127
2612.00	*	.	.	.	299	1
2659.00	*	0.3428	0.6572	.	299	0

Note: The marked survival times are censored observations.

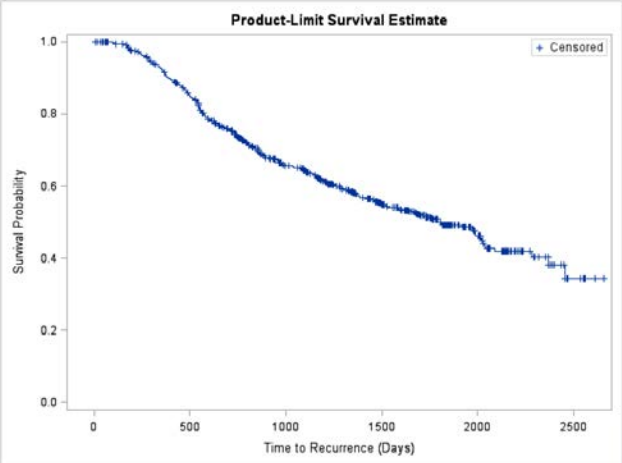
Summary Statistics for Time Variable rectime

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval		
		Transform	[Lower	Upper)
75	.	LOGLOG	2456.00	.
50	1807.00	LOGLOG	1528.00	2018.00
25	727.00	LOGLOG	600.00	799.00

Mean	Standard Error
1591.69	36.23

Note: The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.



Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
686	299	387	56.41

Output 1. Partial output from PROC LIFETEST

Output 1 shows the results from PROC LIFETEST. First, the output shows a portion of the Product-Limit Survival Estimates, or life table, produced by PROC LIFETEST. In this table, the first column lists time from 0 to the last observed time point: 2,659 days. The * indicates a patient who is censored and therefore their `rectime` is the last day that they were observed. Observations without the * had an event at their time point. The survival column shows the conditional probability of no recurrence given the patient has made it to the `rectime` point without recurrence. The failure column is the compliment of the Survival column or 1- Survival showing the probability of recurrence given the patient has not recurred yet. The Survival Standard Error provides an estimated standard error for the estimate in the survival column. The Number Failed column indicates the number of patients that have recurrence/death at the `rectime` day. The Number Left column gives the number of patients that are still being observed and had not yet had recurrence/death or were censored at the `rectime` time point.

How do we interpret these results? For example, at `rectime` day=1807.0, the probability of not having a recurrence/death (Survival) equals 0.4994. This is the day when the probability of failure is closest to and just greater than 50%, otherwise known as the median survival time. This is an important statistic for time to event data and is also reported in the Quartile Estimates Table for Percent = 50. Note that at `rectime` =1814, there are two patients who had an event which means that the Survival and Failure columns contain missing data for the first occurrence where `rectime` =1814 but data are filled in for the second occurrence where `rectime` = 1814. The last observation of this table, `rectime` = 2659 days, is censored.

The quartile estimates table lists the estimated number of days when approximately 25%, 50%, and 75% of the patients have had a recurrence. Note that there is no 75% point estimate because at the last observed time point `rectime` = 2659 days, failure = 0.6572 or probability of recurrence = 65.7%. This means that the failure percentage never reaches 75% and this point estimate cannot be provided. Under this table the mean is reported and a note is given that indicates the mean survival and standard error are underestimated. The skewed right tendency of censored time data demonstrates why median survival is more often requested that mean survival.

The plot of Survival Probability against Time to Recurrence is called a Kaplan-Meier Curve. From this curve, we can observe that with increasing number of days of time to recurrence, the survival probability decreases. The markers on the Kaplan-Meier Curve show when there were censored observations.

In the Summary of the Number of Censored and Uncensored Values Tables, we observe that 56.41% of the cohort was censored. This equates to 100% - 56.41% = 43.59% of the patients in this sample having a breast cancer recurrence or death.

UNIVARIATE SURVIVAL ANALYSES

After getting summary information about the disease free survival or time to recurrence for the cohort, we might wish to understand what covariates are predictive of time to breast cancer recurrence. Different techniques are used for evaluating categorical and continuous variables.

TESTING CATEGORICAL COVARIATES

For categorical variables we can look at the Kaplan-Meier curves, stratifying on the covariate of interest. Here we will look at `grade` to see if it is associated with time to recurrence.

```
PROC LIFETEST DATA = gbcs NOTABLE;
  TIME rectime*censrec(0);
  STRATA grade / TEST = LOGRANK ADJUST = SIDAK;
RUN;
```

In the first line of PROC LIFETEST we add the NOTABLE option. This suppresses the life table and summary table of quartiles. We add the STRATA statement with the covariate we wish to evaluate, `grade`. To quantifiably test the difference among the Kaplan-Meier curves, we use the log-rank test and

specify TEST=LOGRANK as an option in the STRATA statement. The `grade` covariate has three levels. Therefore, we wish to adjust the results of the log-rank test for multiple comparisons using the ADJUST=SIDAK option. The results of stratifying `rectime` by `grade` are show in Output 2.

The LIFETEST Procedure

Summary of the Number of Censored and Uncensored Values

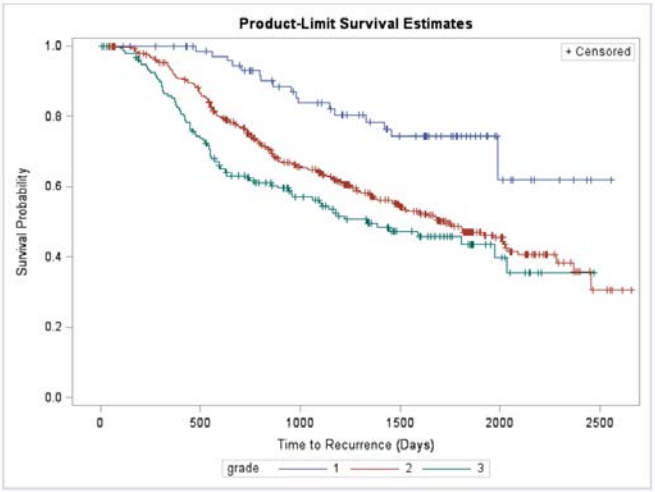
Stratum	grade	Total	Failed	Censored	Percent Censored
1	1	81	18	63	77.78
2	2	444	202	242	54.50
3	3	161	79	82	50.93
Total		686	299	387	56.41

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	21.0944	2	<.0001

Adjustment for Multiple Comparisons for the Logrank Test

Strata Comparison			p-Values	
grade	grade	Chi-Square	Raw	Sidak
1	2	4.9210	0.0265	0.0775
1	3	19.9609	<.0001	<.0001
2	3	1.4387	0.2303	0.5441



Output 2. PROC LIFETEST partial output for using STRATA grade

The three Kaplan-Meier Curve plots in Output 2 allow us to evaluate the association of time to recurrence `rectime` with the categorical covariate `grade`. In the first summary table of the output, we can observe the number of failed and the number of censored at each level of the strata. The Test of Equality over

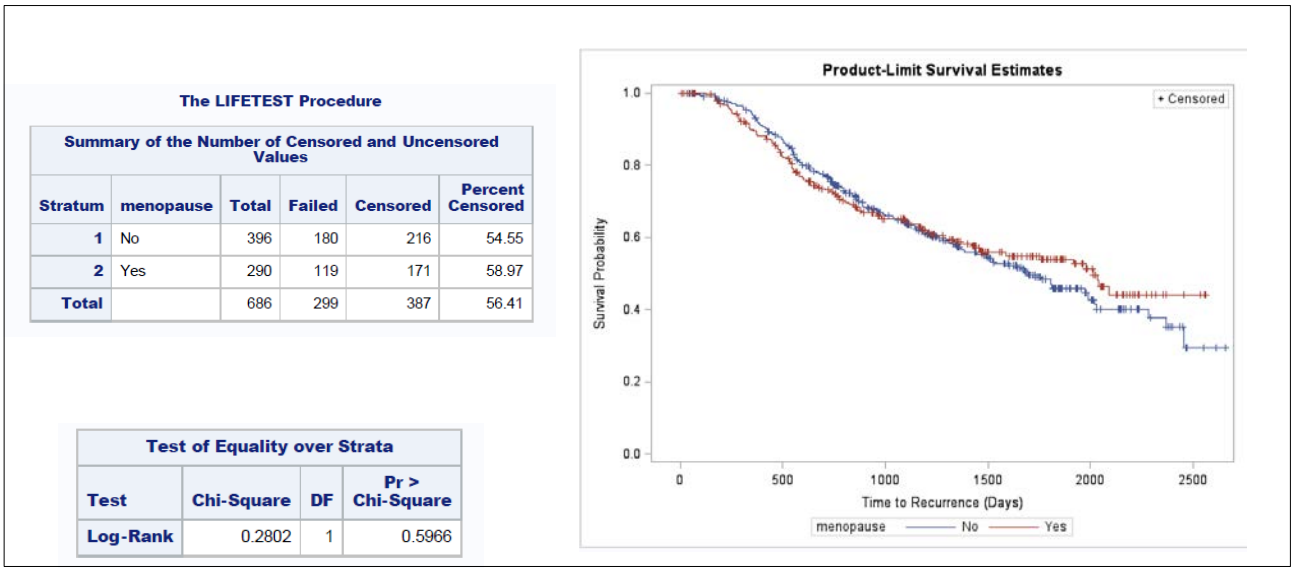
Strata table has the results of the log-rank test. With a p-value<0.05, the log-rank test is statistically significant indicating that time to recurrence `rectime` is significantly associated with `grade`. In the Adjustment for Multiple Comparisons for the log-rank test, the Sidak p-values show that there is a statistically significant difference between `grade` levels 1 and 3, but not between `grade` levels 2 and 3 nor between `grade` levels 1 and 2. The plot of Kaplan-Meier curves demonstrates the differences in `rectime` among the `grade` levels. When there is good separation among the curves then there is a significant difference in time to recurrence for the different levels of the covariate. When there is little separation or overlap among the curves then the categorical covariate is not associated with time to recurrence. In this figure, we see overlap between the red line (grade 2) and green line (grade 3) which indicates that they are not significantly different. The Kaplan-Meier curves for `grade` levels 1 (blue) and 2 (red) appear further apart for the most part as well as the curves for grade levels 1(blue) and 3 (green). However, the Sidak multiple comparison results indicate that levels 1 and 3 are significantly different but levels 1 and 2 are not. This demonstrates that it is not enough to look at the plots and overall Log-Rank test for evaluating the strata, we need to also look at the multiple comparison results.

In another example, we show the results of testing the `menopause` (yes/no) covariate. Here we don't need the ADJUST=SIDAK option because there are only two levels of the covariate therefore adjusting for multiple comparisons in unnecessary. We also create a Yes/No format for menopause which makes the plot easier to interpret.

```
PROC FORMAT;
  VALUE ynf
  1='No'
  2='Yes';

PROC LIFETEST DATA = gbcs NOTABLE;
  TIME rectime*censrec(0);
  STRATA menopause / TEST = LOGRANK;
  FORMAT menopause ynf.;
RUN;
```

The output for the stratifying `rectime` by `menopause` is in Output 3.



The LIFETEST Procedure

Summary of the Number of Censored and Uncensored Values

Stratum	menopause	Total	Failed	Censored	Percent Censored
1	No	396	180	216	54.55
2	Yes	290	119	171	58.97
Total		686	299	387	56.41

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	0.2802	1	0.5966

Output 3. PROC LIFETEST partial output for using STRATA menopause

Output 3 shows the results of stratifying the `rectime` Kaplan-Meier curves by the covariate `menopause`. The first summary table shows the number of failed and censored in each of the two strata for `menopause`. The percent censor rates are similar in the two levels of `menopause`. The Tests of Equality over Strata for `menopause` show a non-significant Log-Rank test p-value, greater than 0.05. Therefore, there is no statistically significant difference in the `rectime` Kaplan-Meier curves when stratified by `menopause`. This non-significance is also demonstrated in the plot which shows a great amount of overlap in the `menopause` no (blue) and yes (red) Kaplan-Meier curves.

Suppose now we want to evaluate the association of continuous covariates with `rectime`. Since the Kaplan-Meier curves cannot be stratified by continuous covariates, we have two possible solutions. First a continuous variable could be defined into meaningful groups and then tested as above for a categorical variable. Second we could evaluate the continuous covariate in the Cox Proportional Hazard Model.

COX PROPORTIONAL HAZARD MODEL

The Cox Proportional Hazard Model is a semi-parametric model often used to for time to event data which can be used when the underlying distribution of the time to event data is unknown -- often the case with health sciences data. In a Cox Proportional Hazard Model, the hazard, or rate of failure, is modeled as a function of the linear combination of the covariates. Similar to logistic regression which looks at the exponentiated covariate coefficient as an odds ratio, the Cox Proportional Hazard Model speaks of the exponentiated covariate coefficient as a hazards ratio. In order to use a covariate in the Cox Proportional Hazards Model, we need to make sure that two assumptions are satisfied. When the linearity assumption is satisfied, we assume that there is a linear association between the covariate and the Cox Proportional Hazards Model outcome. And for the proportionality assumption, we assume that over time hazards ratio stay the same.

TESTING CONTINUOUS COVARIATES

Since we cannot use Kaplan-Meier plots and PROC LIFETEST, we instead use PROC PHREG and evaluate the significance of a continuous covariate in the Cox Proportional Hazards Model. In this section, we also evaluate that the linearity assumption and proportional hazards assumption are held.

The following code evaluates the univariate association of tumor size (mm), the variable named `size`, with time to recurrence `rectime`.

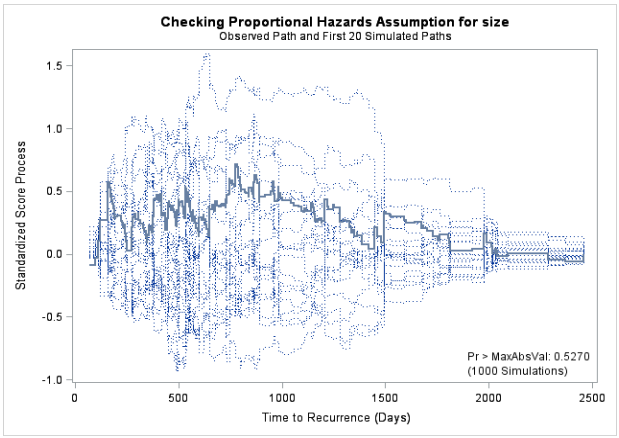
```
PROC PHREG DATA = gbcs;
  MODEL rectime*censrec(0) = size;
  ASSESS VAR = (size) PH / RESAMPLE SEED = 12345;
RUN;
```

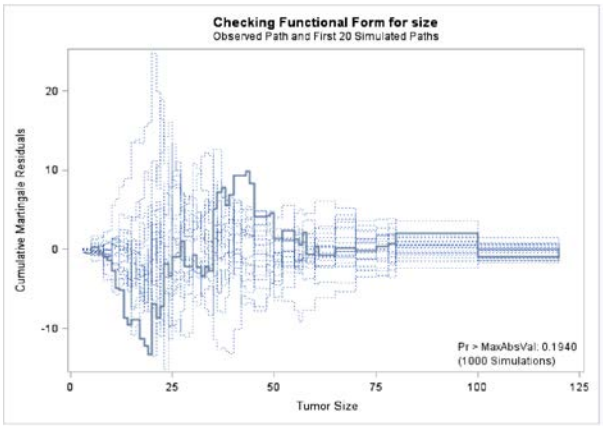
Note that the statements for PROC PHREG are very similar to those used PROC LIFETEST. In PROC PHREG, instead of a TIME statement, a MODEL statement is used. And instead of specifying the covariate in a STRATA statement, the covariate, `size`, is written in the MODEL statement after the equal sign. There is an additional ASSESS statement used in this procedure. This ASSESS statement checks for any major issues with the linearity assumption (called functional form) and also for the proportional hazard assumption -- that the covariate is linearly associated with the hazard for `rectime` and that the hazard ratio does not change with `rectime`. To check the linearity assumption we specify the covariate name VAR =`(size)` in the ASSESS statement. To check the proportional hazards assumption, we specify PH in the ASSESS statement. This ASSESS statement generates two plots: a plot of the cumulative Martingale residuals against the covariate to test the linear assumption and the plot of the standardized score process (a Martingale residual transform) against time to recurrence `rectime` to check the proportional hazards assumption. Details about Martingale residuals or standardizes score processes are not covered in this paper. We are mainly interested in the interpretation of these plots. For further information please refer to (Lin, Wei, & Ying, 1993). The RESAMPLE option in the ASSESS statement generates a p-value for the maximum residual in each of the plots. Since this plot is based on

simulations using a random number generator, we set the SEED to 12345 in order to be able to reproduce the similar plots at a later time point. The partial results of the PROC PHREG for the continuous covariate `size` are provided in Output 4.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
size	1	0.01484	0.00351	17.9019	<.0001	1.015	Tumor Size





Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
size	0.7106	1000	12345	0.5270

Supremum Test for Functional Form

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
size	13.3571	1000	12345	0.1940

Output 4. Partial Output from PROC PHREG Univariate Test for Size

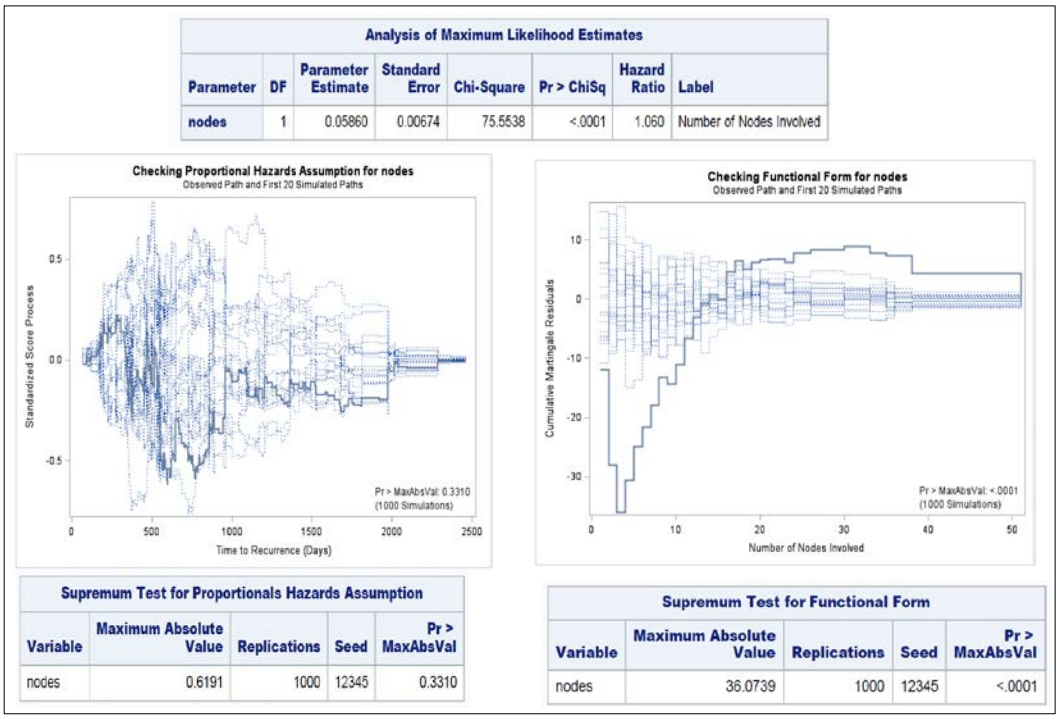
In Output 4, the Analysis of Maximum Likelihood Estimates table provides a Hazard Ratio and p-value for the `size` covariate. The Hazard Ratio is the exponentiated parameter estimate, or exp(Parameter Estimate). It is best to think of a hazard as the risk of experiencing an event at a certain time point given an observed time period. A hazard in this German Breast Cancer study would be the risk of having a recurrence for example 50 days after diagnosis. Hazard Ratio compares the risk for different levels of the covariate. For example, the hazard ratio for `size` is 1.015 means that for every +1 increase in mm of tumor size, the risk of having a recurrence multiplies by 1.015 or increases by 1.5%. Next, we look at the Cumulative Martingale Residual Plot. What we are looking for in this plot is that the solid line does not deviate too much from the middle of the plot -- that the absolute value of any residual is not much larger than 0. We see that the largest absolute residual value occurs with tumor size close to 25. From the output of the Supremum Test for Functional Form, we see that this Maximum Absolute Value = 13.3751 and the test has a p-value of 0.1940 which is not significant. Since the p-value is not significant, this deviation is not too large, and we can conclude that the linearity assumption of the size variable holds. Similarly, in the plot of the Standardized Score Process against time to recurrence, the largest deviation occurs around 800 days. The Supremum Test for Proportional Hazards Assumption tells us that the maximum absolute value = 0.7106 and this test has a p-value of 0.5270 which is non-significant. The non-significance of this test tells us that the proportional hazards assumption holds as well.

In another example, we test a continuous covariate, number of nodes involved (`node`) in the Cox Proportional Hazards model.

```
PROC PHREG DATA = gbcs;
  MODEL rectime*censrec(0) = nodes;
  ASSESS VAR = (nodes) PH / RESAMPLE SEED = 12345;
```

```
RUN;
```

The results of testing nodes in Cox Proportional Hazards Model are in Output 5.



Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
nodes	1	0.05860	0.00674	75.5538	<.0001	1.060	Number of Nodes Involved

Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
nodes	0.6191	1000	12345	0.3310

Supremum Test for Functional Form

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
nodes	36.0739	1000	12345	<.0001

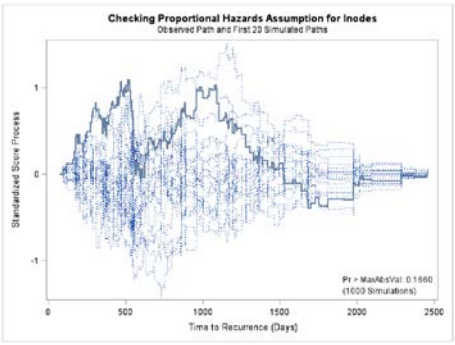
Output 5. Partial Output of PROC PHREG Univariate test for nodes

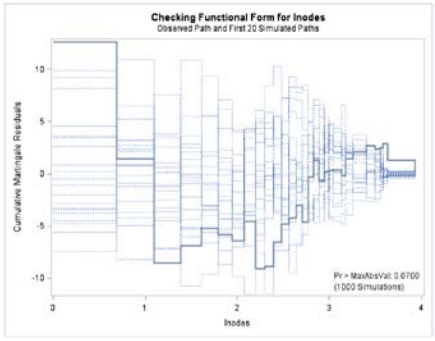
In Output 5, we observe that the Hazard Ratio for nodes is equal to 1.06 and is statistically significant with a p-value<0.0001. However in the plot Checking Functional Form for linearity of nodes there is a residual with maximum absolute value of 36.0739 and the test is statistically significant with a p-value <0.0001. This result indicates that the linearity assumption is violated. In the plot Checking Proportional Hazards for nodes, the residual with the maximum absolute value is 0.6191 with a non-significant p-value=0.3310. Therefore the proportional hazards assumption for nodes still holds. There are a couple options in how to deal with the violation of the linearity assumption. One option is to do a transform on the variable so that the continuous variable can still be used in the model. The nodes (number of positive lymph nodes) variable ranges from 1 to 51, therefore using a log transform of nodes seems reasonable and constructing this variable (lnodes) can be done within the PROC PHREG code.

```
PROC PHREG DATA = gbcs;
  MODEL rectime*censrec(0) = lnodes;
  ASSESS VAR = (lnodes) PH / RESAMPLE SEED = 12345;
   lnodes = log(nodes);
RUN;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
lnodes	1	0.54294	0.06334	73.4664	<.0001	1.721





Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
lnodes	1.0901	1000	12345	0.1660

Supremum Test for Functional Form

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
lnodes	12.6069	1000	12345	0.0700

Output 6. Partial Output of PROC PHREG for `lnodes`

Output 6 shows the partial output of PROC PHREG for `lnodes`, the log-transform of the `nodes` variable. The output shows us that this transform helped. Both the proportional hazards assumption and linearity assumption are met as demonstrated by the non-significant p-values in the Supremum tests. Also, the hazard ratio of lnodes is highly significant with a p-value <0.0001. One of the challenges of using a log-transformed variable is the interpretation. Now the hazard ratio=1.721 can be interpreted as a 72% increased risk of recurrence with every +1 increase in log nodes. This might not be so useful to reviewers.

Another possible solution to the issue of the linearity assumption violation is to convert the continuous variable into a categorical variable. This can easily be achieved by formatting the `nodes` variable. Converting the continuous variable to a clinically meaningful categorical variable might be preferable because it can be challenging to interpret log transformed variables.

```
PROC FORMAT;
  VALUE nodes
  1-3 = 'N1'
  4-9 = 'N2'
  10-HIGH = 'N3';
RUN;

PROC PHREG DATA = gbcs;
  CLASS nodes(REF = 'N1') / PARAM = REF;
  MODEL rectime*censrec(0) = nodes;
  ASSESS PH / RESAMPLE SEED = 12345;
  FORMAT nodes nodes.;
RUN;
```

Using PROC FORMAT, we divide nodes into three categories: N1, N2, and N3. Then in PROC PHREG, we add a CLASS statement for the categorical `nodes`. We can indicate the preferred reference level in the parentheses (REF='N1') and the option PARAM=REF specifies that we want the hazards ratio to be given using reference coding, i.e. the hazards of N2 and N3 will be compared to N1 for the hazards ratios.

Note we remove VAR = (nodes) from the ASSESS statement because we do not have to check the linearity assumption if nodes is a categorical covariate instead of a continuous one. However, we still check the proportional assumption for categorical covariates. Output not shown.

MODEL BUILDING

After conducting the univariate analyses to see which covariates are associated with time to recurrence. We might wish to build a model which simultaneously looks at multiple covariates.

STEPWISE MODEL SELECTION

Once we have a set of covariates that can be included in the model, how do we select which variables to include in our model? One way to do this would be a stepwise selection. Using stepwise selection, we create a model that includes all significant covariates. In this example, we use an entry (SLE) and stay (SLS) criteria of 0.05 – covariates with p-values <.05 will be included in the model. We can include all categorical variables in the CLASS statement and select the formatted reference level in parentheses. We include in this model, all categorical and continuous variables including the following which were not shown in univariate testing above: `age` – age at diagnosis, `menopause` (yes/no) – Menopause Status, `hormone` (yes/no) – Hormone Therapy, `prog_recp` – Number of Progesterone Receptors, and `estrg_recp` - Number of Estrogen receptors. For ease of interpretation, we create Yes/No formats for `menopause` and `hormone` and positive/negative formats for `estrg_recp` and `proc_recp`.

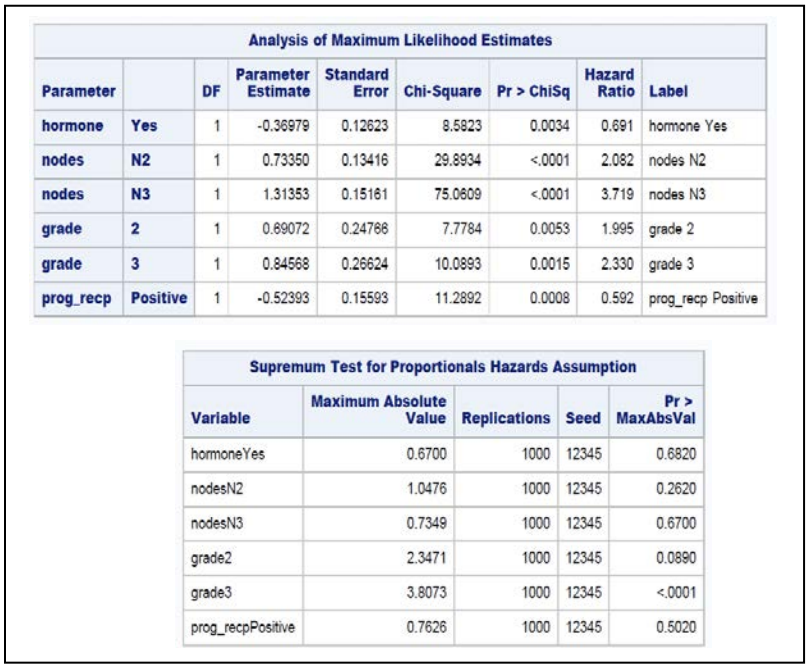
```
PROC FORMAT;
  VALUE nodes
  1-3 = 'N1'
  4-9 = 'N2'
  10-HIGH = 'N3';

  VALUE ynf
  1 = 'No'
  2 = 'Yes';

  VALUE posneg
  0 = 'Negative'
  1-HIGH = 'Positive';
RUN;


PROC PHREG DATA = gbcs;
  CLASS nodes(REF = 'N1') grade(REF = '1') hormone(REF = 'No')
        menopause(REF = 'No') prog_recp(REF = 'Negative')
        estrg_recp(REF = 'Negative') / PARAM = REF;
  MODEL rectime*censrec(0) = menopause hormone age size nodes grade
        prog_recp estrg_recp / SELECTION = STEPWISE SLS = .05 SLE = .05;
  ASSESS VAR = (age size) PH / RESAMPLE SEED = 12345;
  FORMAT nodes nodes. menopause hormone ynf. prog_recp estrg_recp posneg.;
RUN;
```

Below is the Analyses of Maximum Likelihood Estimates table of the final model from the Stepwise Selection, and the Supremum Test for Proportional Hazards Assumption table from the covariates that were included in the final model.



Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
hormone	Yes	1	-0.36979	0.12623	8.5823	0.0034	0.691	hormone Yes
nodes	N2	1	0.73350	0.13416	29.8934	<.0001	2.082	nodes N2
nodes	N3	1	1.31353	0.15161	75.0609	<.0001	3.719	nodes N3
grade	2	1	0.69072	0.24786	7.7784	0.0053	1.995	grade 2
grade	3	1	0.84568	0.26624	10.0893	0.0015	2.330	grade 3
prog_recp	Positive	1	-0.52393	0.15593	11.2892	0.0008	0.592	prog_recp Positive

Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
hormoneYes	0.6700	1000	12345	0.6820
nodesN2	1.0476	1000	12345	0.2620
nodesN3	0.7349	1000	12345	0.6700
grade2	2.3471	1000	12345	0.0890
grade3	3.8073	1000	12345	<.0001
prog_recpPositive	0.7626	1000	12345	0.5020

Output 7. Partial Output from PROC PHREG Stepwise Selection Model

The final Cox-Proportional Hazards Model from the Stepwise selection method includes `hormones`, `nodes`, `prog_recp`, and `grade`. All remaining covariates have p-values <0.05, and they are all categorical. In interpreting the hazard ratios of these covariates, we need to be sure to mention that these are adjusted hazard ratios. For example, the hormone hazard ratio is equal to 0.691. This can be interpreted as the risk of recurrence for women taking hormone treatment is (1-0.691)*100% or 29.9% less for women not taking hormone treatment, adjusting for grade, node, and progesterone receptor status. The other covariates can be interpreted similarly with mention of adjusting in the statement. The Supremum Test for Proportional Hazards Assumption table shows that all p-values are greater than 0.05 with the exception of grade 3. When the proportional-hazards assumption is not met there are three different ways of handling the covariate that violates the assumption: creating a time dependent covariate, stratification, and using a parametric model with PROC LIFEREG. Here, we discuss the first two possibilities.

TIME-DEPENDENT COVARIATES

To convert `grade` into a time dependent covariate, we multiply `grade` by log(`rectime`) and include this term in the model. We can create this interaction term directly in PROC PHREG without having to use an extra DATA step. However since `grade` is a categorical covariate as defined previously in the CLASS statement, PROC PHREG will not interpret the interaction correctly unless we create dummy variables. Therefore, we eliminate grade from the CLASS statement, create dummy variables for the two levels of grade (grade=2, 3) that are not the reference group (grade=1), interact these dummy variables with log(`rectime`), and include these interactions and dummy variables in the MODEL statement. The original grade variable is no longer included in the CLASS or MODEL statements.

```
PROC PHREG DATA = gbcs;
  CLASS nodes(REF = 'N1') hormone(REF = 'No') prog_recp(REF = 'Negative') /
        PARAM = REF;
  MODEL rectime*censrec(0) = hormone nodes prog_recp grade2 grade3 grade2t
grade3t / RL;
  ASSESS PH / RESAMPLE SEED = 12345;
   grade2 = (grade = 2);
```

```
    grade3 = (grade = 3);
    grade2t = grade2*log(rectime);
    grade3t = grade3*log(rectime);
    grade: TEST grade2, grade3, grade2t, grade3t;
  FORMAT nodes nodes. hormone ynf. prog_recp posneg.;
RUN;
```

Directly below the MODEL statement, we create dummy variables `grade2` and grade3, and time-dependent covariates `grade2t` and `grade3t`. All four of these covariates are included in the model in place of `grade`. We can test the significance of all grade terms simultaneously in the TEST statement. The model also includes the significant terms from the stepwise selection: `nodes` (categorical), `prog_recp`, and `hormone`. In this code we also include the RL option in our model statement which produces the confidence intervals for the hazard ratio, these are estimates that are often needed when reporting results. No ASSESS statement is used in this model since we are including time-dependent covariates. The partial output of this PROC PHREG is shown below.

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
hormone	Yes	1	-0.37538	0.12620	8.8482	0.0029	0.687	0.536	0.880
nodes	N2	1	0.73719	0.13451	30.0350	<.0001	2.090	1.606	2.721
nodes	N3	1	1.29898	0.15193	73.0976	<.0001	3.666	2.722	4.937
prog_recp	Positive	1	-0.52132	0.15604	11.1611	0.0008	0.594	0.437	0.806
grade2		1	5.97043	3.11535	3.6728	0.0553	391.672	0.873	175676.5
grade3		1	9.89775	3.29000	9.0507	0.0026	19885.56	31.484	12560005
grade2t		1	-0.78012	0.45290	2.9670	0.0850	0.458	0.189	1.114
grade3t		1	-1.37437	0.48397	8.0643	0.0045	0.253	0.098	0.653

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
grade	19.3999	4	0.0007

Output 8. Partial Output of PROC PHREG including Time-Dependent Grade

The Analysis of Maximum Likelihood Estimates Table in Output 8 contains the hazard ratios for all model terms. All terms are statistically significant except `grade2` and `grade2t`. Because we are including interactions for `grade` in our model, we don't need to evaluate the p-values of the main effects for `grade`, we will focus on the interaction terms. We will still keep `grade2` and `grade2t` in the model since the test for all grade terms in the Linear Hypotheses Testing Results is statistically significant and indicates that the time dependent version of grade is important overall. We need to include all levels (dummy variables) in the model for the grade term to be represented properly. The 95% Hazard Ratio Confident Limits are included in the Analysis of Maximum Likelihood Estimates Table. Statistically significant covariates have confidence intervals which are below or above 1, but never include 1. Now we interpret the hazard ratio of `grade` differently at every time point. For example, for a given `rectime = 100` days, the risk of recurrence for a patient with grade 2 = exp(5.97043 - 0.78012*log(100)) = 10.78 times the risk of recurrence for a patient with grade 1, adjusting for nodes, hormone, and progesterone receptor status.

It should be noted that for ease of interpretation, another way to create time-dependent covariates is by interacting with a meaningful time dependent dummy variable instead of log (time). For example, we can create a dummy variable t100 equal to 0 when `rectime` <100 days and 1 when `rectime` $\geq$ 100 days. This way we will have two hazards ratio, one for before 100 days and one for 100 days and greater. This structuring of a time dependent covariate is much easier to interpret.

Stratification

Another method for handling the grade covariate not satisfying the proportional hazards assumption is to

remove the grade covariate from the MODEL statement and put it into a STRATA statement. The MODEL statement then only includes `nodes`, `hormones`, and `prog_recp`. Grade will automatically be treated as categorical so it can also be removed from the CLASS statement.

```
PROC PHREG DATA = gbcs;
  CLASS nodes(REF = 'N1') hormone(REF = 'No') prog_recp(REF = 'Negative') /
        PARAM = REF;
  MODEL rectime*censrec(0) = hormone nodes prog_recp / RL;
  STRATA grade;
  FORMAT nodes nodes.  hormone ynf. prog_recp posneg.;
RUN;
```

The Partial Output of the PROC PHREG stratified by `grade` is shown below.

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
hormone	Yes	1	-0.37253	0.12628	8.7021	0.0032	0.689	0.538	0.882
nodes	N2	1	0.73987	0.13480	30.1265	<.0001	2.096	1.609	2.729
nodes	N3	1	1.29831	0.15230	72.6723	<.0001	3.663	2.718	4.937
prog_recp	Positive	1	-0.52704	0.15640	11.3556	0.0008	0.590	0.434	0.802

Output 9. Partial Output from PROC PHREG for Cox Proportional Hazards Model stratified by grade

We compare the hazard ratios in Output 9 to those in Output 7. We expect the covariates to be somewhat different in the two model outputs. We see that the covariates change using stratification, but only slightly. The choice in using stratification or including time-dependent covariates in the model depends on the importance of the covariate that does not meet that proportional hazards assumption. If the variable is clinically important to the time to event outcome then we may prefer to use time-dependent covariates over stratification since with the use of stratification no coefficient is provided for grade.

CREATING SURVIVAL PLOTS

We've shown how to create a basic survival plot with the use of ODS GRAPHICS ON in the LIFETEST procedure. These graphics are the defaults that come with the procedure and are fine for analysis but may not be presentation or publication ready. Creating good looking survival plots is still easy to do with various options as well as other methods for producing graphs.

There are many ways to produce graphics in SAS® and the choice of which method to use depends on the users background and the desired flexibility. The two methods that we will cover are:

1) Using ODS graphics is easy and the plots that are produced come with the standard graphics for the appropriate analysis. The user can create editable point and click graphics. This method is limited in that changes can only be made with certain options.
2) The ODS Graphics Designer tool is a click and point interface that makes it very easy to customize graphics and will also save complicated template coding. The user need to know more about what graphic is appropriate for the analysis, and depending on the analysis the statistical estimates sometimes need to be output to a data set before the graph can be produced.

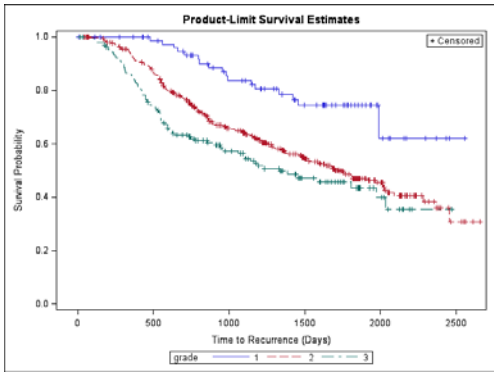


Figure 1. Simple survival plot

There are certainly other ways to make nice graphics in SAS such as utilizing the SG procedures or using the Graph Template Language (GTL). However our aim is to cover simple ways to conduct survival analysis and make corresponding graphics so we will focus on just the methods mentioned above. While using the SG procedures is far easier than GTL, the statements used to create the graph are much like the work that would be done with the click and point interface of ODS graphics designer.

ODS GRAPHICS

An additional way to create the basic survival plot with LIFETEST is to add the PLOTS= option to the PROC LIFETEST statement. This is important because it is where we will add options that will enhance our survival plot. There are other plots that can be made with the PLOT= option such as the diagnostic plots that come standard with LIFETEST.

```
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS = SURVIVAL;
  TIME rectime*censrec(0);
  STRATA grade;
RUN;
ODS GRAPHICS OFF;
```

Typically when creating a presentation graphic for survival analysis we would want to include a table of the patients at risk for each time point in the plot. This can be done by adding the ATRISK= option to the survival plot. The graphics options for LIFETEST are nested options within the plot option.

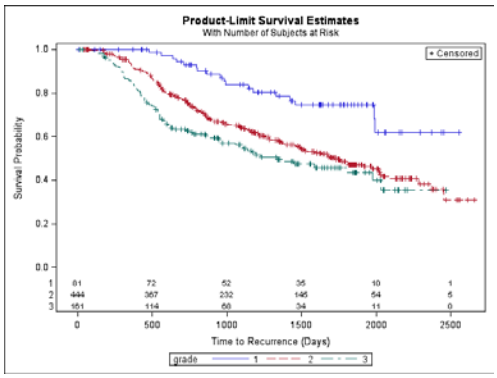


Figure 2. At risk table

```
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS =
     SURVIVAL(ATRISK = 0 to 2500 by
500);
  TIME rectime*censrec(0);
  STRATA grade;
RUN;
ODS GRAPHICS OFF;
```

We may wish to add labels to the legend in the plot and this can be achieved easily with formats.

```
PROC FORMAT;
  VALUE grd
  1 = 'Low'
  2 = 'Intermediate'
  3 = 'High';
RUN;
```

```
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS =
     SURVIVAL(ATRISK = 0 to 2500 by 500);
  TIME rectime*censrec(0);
  STRATA grade;
  FORMAT grade grd.;
RUN;
```

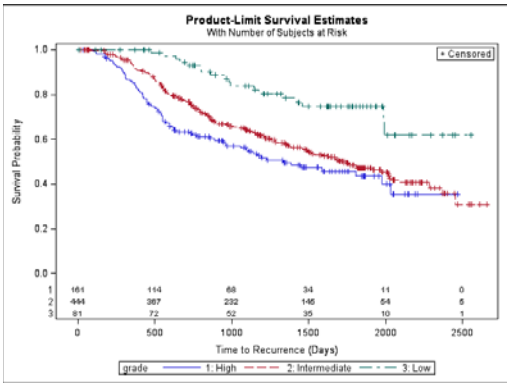


Figure 3. At risk table with labels

```
ODS GRAPHICS OFF;
```

Adding more options to the survival plot creates a plot with confidence bands per strata group, adds the Log-Rank test statistic to the plot and suppresses the censoring tick marks.

```
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS =
     SURVIVAL(ATRISK = 0 to 2500 by 500
     CB = HW TEST NOCENSOR);
  TIME rectime*censrec(0);
  STRATA grade;
  FORMAT grade grd.;
RUN;
ODS GRAPHICS OFF;
```

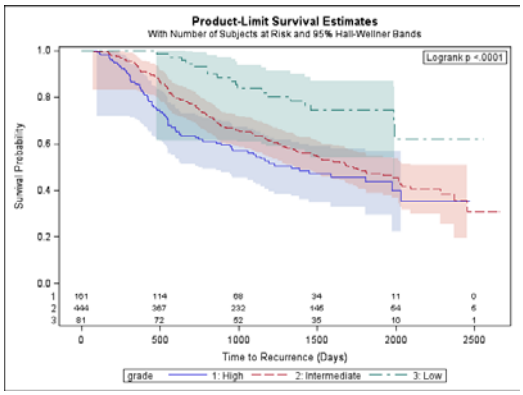


Figure 4. 95% confidence bands

Style options control the colors and style of the plots. These can be changed to any of the SAS supplied style templates using the ODS HTML statement. Note that the style is changed in the environment where the graphic is produced. If the graphic is sent to the listing window then the style would be modified with the ODS LISTING statement.

```
ODS HTML STYLE = journal;
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS =
     SURVIVAL(ATRISK = 0 to 2500 by 500
     CB = HW TEST NOCENSOR);
  TIME rectime*censrec(0);
  STRATA grade;
  FORMAT grade grd.;
RUN;
ODS GRAPHICS OFF;
```

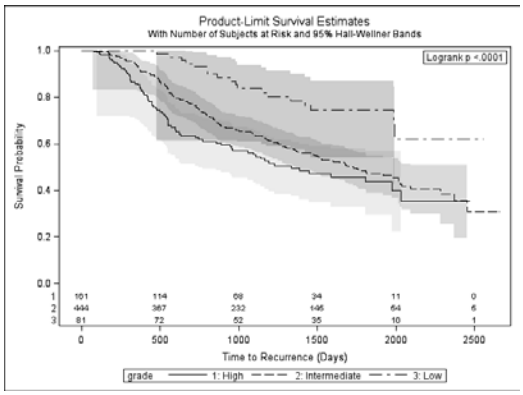


Figure 5. Journal style

One perk of working with ODS graphics through coding in the program editor is that the ODS Graphics Editor can be invoked to create graphics that can be edited by hand. There are some limitations of what the Graphics Editor can and can't do, however it can be very useful for quick edits or insets. One disadvantage is that these edits are made by hand and cannot be replicated in the coding.

```
ODS LISTING SGE = ON;
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS =
     SURVIVAL(ATRISK = 0 to 2500
     by 500 CB = HW TEST
     NOCENSOR);
  TIME rectime*censrec(0);
```

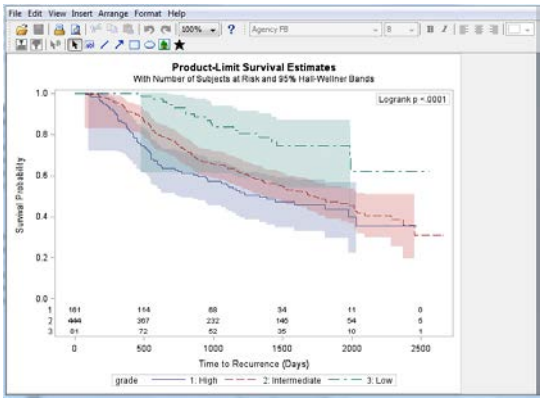


Figure 6. ODS Graphics Editor Window

```
  STRATA grade;
  FORMAT grade grd.;
RUN;
ODS GRAPHICS OFF;
ODS LISTING SGE = OFF;
```

In this environment many custom changes can be made to the plot such as modifying the title and axis labels, changing the attributes of lines and confidence bands, adding grid lines to axes, relocating the legend, inserting text and symbols, and more. These files can be saved as ODS Graphic Editor files which means that they can be opened later for further editing, or they can be saved as png files.

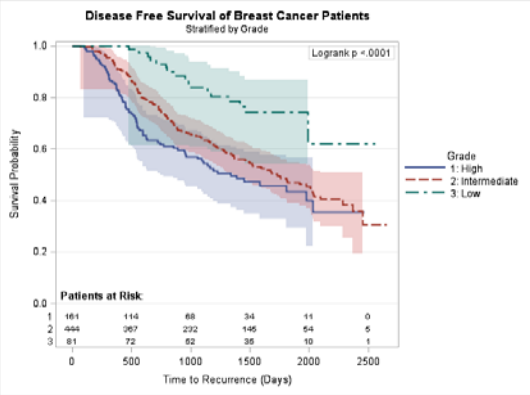


Figure 7. Custom survival plot edited with the Graphics Editor

ODS GRAPHICS DESIGNER

Creating plots with the ODS Graphics Designer is similar in concept to creating plots with SG procedures. These types of plots are made in layers, and additionally some plot types can be overlaid with each other while others cannot. In the case of survival analysis the survival probabilities and corresponding statistics have to be calculated with LIFETEST and then output to a data set before using ODS Graphics Designer to make survival plots to identify what components of the output can be made into data sets the ODS TRACE statement can be used with PROC LIFETEST. This will identify all of the sections of the LIFETEST output in the log.

```
ODS TRACE ON;
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS = SURVIVAL(ATRISK = 0 to 2500 by 500
     CB = HW TEST NOCENSOR);
  TIME rectime*censrec(0);
  STRATA grade;
  FORMAT grade grd.;
RUN;
ODS GRAPHICS OFF;
ODS TRACE OFF;
```

```
Output Added:
-------------
Name:       SurvivalPlot
Label:      Survival Curves
Template:   Stat.Lifetest.Graphics.ProductLimitSurvival
Path:       Lifetest.SurvivalPlot
-------------
```

The data we would like to capture in a data set relates to the survival plot so we will use an ODS OUTPUT statement to capture this information into a SAS data set. The data that is contained in the data set has everything that is needed to create a plot with ODS Graphics Designer. Note that the `tAtRisk` variable has the total at risk patients for the given time points on the plot, and is stored as a redundant row in the data set. This is fine because it doesn't not impact the calculation of the point estimates or confidence intervals, and this `tAtRisk` data will only be used to make the at risk table on the plot.

```
ODS OUTPUT Survivalplot =
SurvivalPlotData;
ODS GRAPHICS ON;
PROC LIFETEST DATA = gbcs PLOTS
= SURVIVAL(ATRISK = 0 to 2500
by 500
     CB = HW TEST NOCENSOR);
  TIME rectime*censrec(0);
  STRATA grade;
  FORMAT grade grd.;
RUN;
ODS GRAPHICS OFF;
```

Obs	HW_UCL	HW_LCL	Time	Survival	AtRisk	Event	Censored	sAtRisk	Stratum	StratumNum	HW_UCL1	HW_LCL1	HW_UCL2	HW_LCL2	HW_UCL3	HW_LCL3
1	.	.	0	1.00000	161	0	.	.	1: High	1	.	.	.	.	.	.
2	.	.	0	.	161	.	.	0	1: High	1	.	.	.	.	.	.
3	.	.	15	.	161	0	1.00000	.	1: High	1	.	.	.	.	.	.
4	.	.	17	.	160	0	1.00000	.	1: High	1	.	.	.	.	.	.
5	.	.	29	.	159	0	1.00000	.	1: High	1	.	.	.	.	.	.
6	1.00000	0.72127	98	0.99367	158	1	.	.	1: High	1	1.00000	0.72127	.	.	.	.
7	1.00000	0.72127	113	0.98734	157	1	.	.	1: High	1	1.00000	0.72127	.	.	.	.
8	0.99994	0.72127	120	0.98101	156	1	.	.	1: High	1	0.99994	0.72127	.	.	.	.
9	0.99897	0.72127	177	0.96835	155	2	.	.	1: High	1	0.99897	0.72127	.	.	.	.
10	.	.	177	.	155	0	0.96835	.	1: High	1	.	.	.	.	.	.
11	0.99781	0.72127	180	0.96198	152	1	.	.	1: High	1	0.99781	0.72127	.	.	.	.
12	.	.	186	.	151	0	0.96198	.	1: High	1	.	.	.	.	.	.
13	0.99615	0.72127	195	0.95557	150	1	.	.	1: High	1	0.99615	0.72127	.	.	.	.
14	0.99402	0.72127	205	0.94916	149	1	.	.	1: High	1	0.99402	0.72127	.	.	.	.
15	0.99146	0.72127	227	0.94274	148	1	.	.	1: High	1	0.99146	0.72127	.	.	.	.

Output 10. Print of the survival plot data

```
PROC PRINT DATA=SurvivalPlotData (OBS = 15);
RUN;
```

To launch the designer, go to Tools – ODS Graphics Designer. It might take a minute to launch the application and if there is a message about connecting to the server, just click retry. The Designer is set-up with a graph gallery where custom graphs can be selected from, including user created plots. The left pane has all the plot layers and inset options that can be used to define a plot.

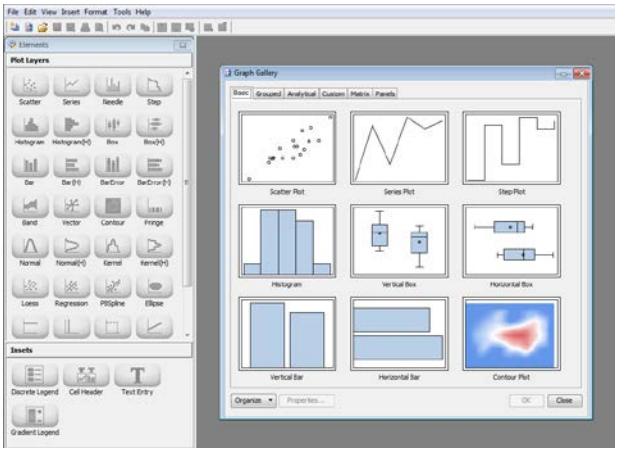


Figure 8. ODS Graphics Designer window

1) To create a basic survival plot.
   a. Go to File – New Blank Graph.
   b. Drag the Step plot from the Plot Layers box over to the new graph.
   c. A window named Assign Data will pop-up. Choose the WORK library and the SurvivalPlotData data set.
   d. On the Plot Variables tab choose `Time` as the X variable, `Survival` as the Y variable and `Stratum` as the Group variable. Click OK.
2) Modify axes
   a. Right click on the y-axis and choose Axis Properties.
   b. Click the advanced tab. Check Custom Axis Range, set the Min = 0 and Max = 1.
   c. Change the Axis dropdown to X.
   d. Click the Label tab. Change the label font to 12 point. Click OK.
3) Add titles.
   a. Go to Insert Title. Modify title as appropriate.
   b. Add a second title. Go to Insert Title. Modify title as appropriate.
   c. Right click on the second title and choose Title Properties. Choose a smaller font and different font style. Click OK.

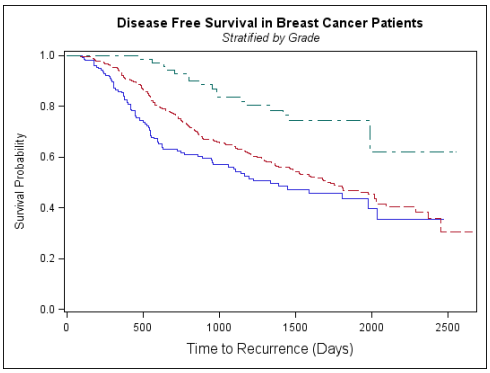


Figure 9. Simple survival plot

4) Add an at risk table.
   a. Add a cell to the graph. Right click in the plot and choose Add a Row.
   b. Drag a StackBlock plot to the new row.
   c. Using the same data set, set `TAtRisk` as the X variable, `AtRisk` as the Block variable and `Stratum` as the Group variable. Click OK.
5) Modify Block Plot properties.
   a. Right click on the Block Plot and choose Plot Properties.
   b. On the display tab un-check Fill and Outline, and check Label.
   c. On the Value tab set the horizontal alignment to Start. Click OK. Note if the numbers look cutoff try refreshing by minimizing the graph window and bringing it back up.

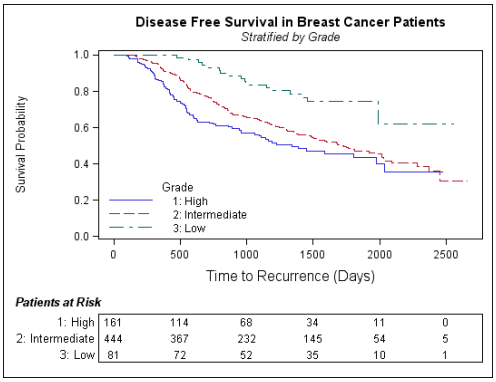


Figure 10. At risk table and inside legend

6) Unify axes.
   a. Right click on the lower X axis and select Axis Properties.
   b. At the bottom of the Axes tab change the Data Range to Union.
   c. Uncheck Label, Value, and Ticks. Click OK.
7) Resize the at risk table.
   a. Click the at risk table. Position the cursor at the top of the at risk table and drag the top part of the table down to resize the window.
8) Add a cell header.
   a. Drag Cell Header to the Block Plot. Add a header label.
   b. Right click the header and select Cell Header Properties. Position text to the left, modify font as appropriate. Resize the at risk table if necessary.
9) Add a legend.
   a. Drag Discrete Legend to the bottom of the survival plot.
   b. Right click and select Legend Properties. Add a legend title of Grade, un-check outline, and position on Bottom Left.
10) Viewing template code and saving
   a. From the toolbar select View – code. This is the PROC TEMPLATE code that creates the plot. It can be copied and run directly in the SAS program editor which means that it can be saved as code and modified as needed, for example in a macro.
   b. From the toolbar select File – Save as. This will allow saving the plot as a SGD file which can be opened in the ODS Graphics Designer later for further editing. Or the graph can be saved as many other file types including PNG, JPEG, and more.
   c. From the toolbar select File – Save in Graph Gallery. This will allow saving as a custom made plot in the graph gallery to be used later with ODS Graphics Designer.

At this point we might wish to add confidence bands to the plot. This can be achieved with the Designer by tweaking the output data set. When the SurvivalPlotData data set was created it added the lower and upper confidence band limits to the data set as `HW_LCL` and `HW_UCL`. A band plot can be added as a layer on the survival plot in Designer, however there is no grouping option for the strata. Creating new variables for the limits for each strata allow us to add bands to our plot.

```
DATA SurvivalPlotData; SET SurvivalPlotData;
  IF StratumNum = 1 THEN DO;
   HW_UCL1 = HW_UCL; HW_LCL1 = HW_LCL;
  END;
  ELSE IF StratumNum = 2 THEN DO;
   HW_UCL2 = HW_UCL; HW_LCL2 = HW_LCL;
  END;
  ELSE IF StratumNum = 3 THEN DO;
   HW_UCL3 = HW_UCL; HW_LCL3 = HW_LCL;
  END;
```

```
RUN;
```

11) Add confidence bands to the survival plot.
    a. Drag a Band plot to the survival plot. Choose `Time` as the X variable, `HW_UCL1` as the Limit Upper variable, and `HW_LCL1` as the Limit Lower variable. Click OK.
    b. Repeat step 11a) two more times to add confidence bands for the 2nd and 3rd strata.
    c. Right click on a band plot and choose Plot Properties. Change the color of the band to match the color of the line, and modify the transparency to 85%.
    d. Using the Plot drop down at the top, repeat step 11c) for the other two strata. Click OK.
12) Modify the legend and y axis.
    a. Right click on the legend and choose Legend Contents. Uncheck band, band2 and band3. Click OK.
    b. (Note this step is needed if an outline reappears on the legend) Right click on the legend and choose Legend Properties. Check and uncheck the Outline box. Click OK.
    c. Right click on the y axis and choose Cell Properties. Check Grid. Click OK.

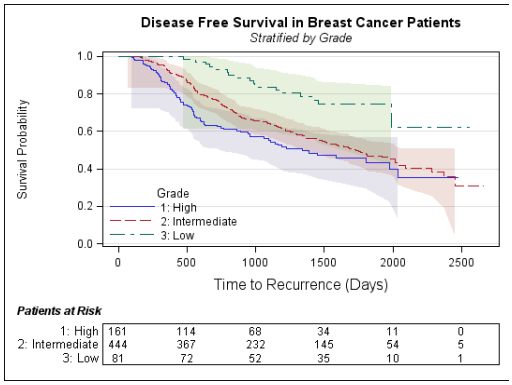


Figure 11. Addition of 95% confidence bands

ADJUSTED SURVIVAL PLOTS

Survival plots can also be made in PHREG by using the PLOTS= option. There are a couple of important points to note. First if there are strata in the adjusted model PHREG will return one plot per strata. Ideally we would like these to appear on one plot. A modification can be made to the PLOTS= option to overlay the strata.

```
ODS GRAPHICS ON;
PROC PHREG DATA = gbcs PLOTS(OVERLAY = BYROW) = (SURVIVAL);
  CLASS nodes (REF = 'N1') hormone(REF = 'No') prog_recp(REF = 'Negative');
  MODEL rectime*censrec(0) = hormone nodes prog_recp / RL;
  STRATA grade;
  FORMAT nodes nodes. hormone ynf. prog_recp posneg.;
RUN;
ODS GRAPHICS OFF;
```

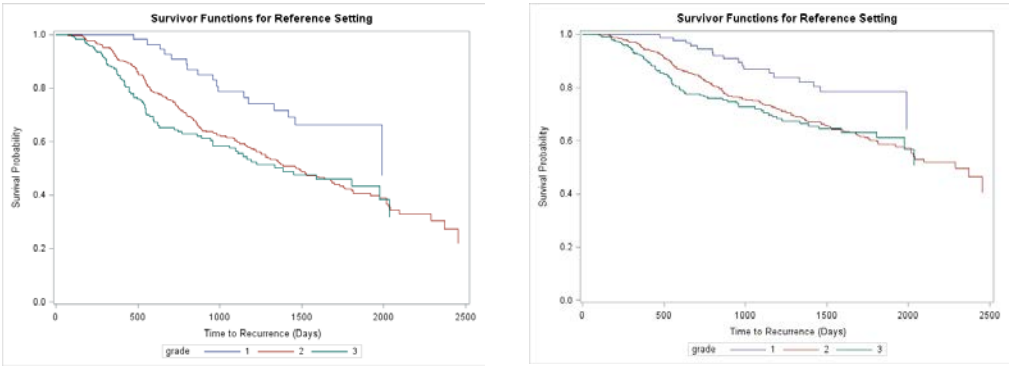


Figure 12. Adjusted survival plots in PHREG for Negative and Positive `prog_recp`

The plots that are generated are adjusted survival plots for each of the covariates in the model. This means that the plot is showing the survival probabilities for the reference groups of categorical variables and the means of continuous predictors. In the code and figure above the reference group for `prog_recp` was set to Negative. Modifying the reference group for `prog_recp` to Positive shows the change in the adjusted plot. Note that this method of producing overlaid survival plots does not work in SAS version 9.3 TS1M0 and TS1M1. There is a work around by way of using the BASELINE statement with a COVARIATES= referring to a data set with the means and reference groups, as shown in SAS/STAT PHREG 9.3 Example 66.8: 'Survivor Function Estimates for Specific Covariate Values'.

CONCLUSION

This paper serves to provide an introduction to survival data analyses using SAS v 9.3 and pays special attention to the ODS graphing capabilities. There are many more topics that can be explored in-depth with respect to survival analyses, especially time-varying parameters in PROC PHREG, model goodness of fit assessment, Martingale Residuals, and parametric regression using PROC LIFEREG. However, this paper is a useful reference in how to approach a survival analyses in SAS without being overwhelmed by notation.

REFERENCES

Hosmer, D.W. and Lemeshow, S. and May, S. (2008).Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition, John Wiley and Sons Inc., New York, NY.

Lin D, Wei L, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika 1993;80(3):557–572.

Matange, Sanjay. SAS Global Forum 2012. *Quick Results with SAS ODS Graphics Designer*. Available at: http://support.sas.com/resources/papers/proceedings12/153-2012.pdf.

Matange, Sanjay and Heath, Dan. November 2011. *Statistical Graphics Procedures by Example*. Cary, North Carolina: SAS Institute Inc.

SAS Support. “Example 66.8 Survivor Function Estimates for Specific Covariate Values”. September 2013. Available at: http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug_phreg_sect053.htm.

RECOMMENDED READING

For other good tutorials on Survival Analyses and more examples, the UCLA Statistical Consulting Group website http://www.ats.ucla.edu/stat/sas/topics/survival.htm is a great resource.

For more in-depth information on survival plots including GTL see ‘Creating and Customizing the Kaplan-Meier Survival Plot in PROC LIFETEST’ by Warren F. Kuhfeld and Ying So, SAS Institute Inc. Available at: http://support.sas.com/resources/papers/proceedings13/427-2013.pdf.

CONTACT INFORMATION

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

Rebecca Ottesen
City of Hope
1500 East Duarte Road
Duarte, CA 91010
Email: rottesen@coh.org