Time to Event Analysis in the Pharmaceutical and Medical Device Industries

Helen M. Chmiel, Experis Corporation, Kalamazoo, Michigan
Evan L. Ritzema, Experis Corporation, Kalamazoo, Michigan

ABSTRACT

Time to event analysis, often referred to as survival analysis, is standard in many industries to evaluate clinically meaningful time to event outcomes, such as time to implant failure or death. In addition to this analysis being widespread, specific needs are varied based on the industry and require software with the ability to generate the required results as well being flexible and easy to use. Some examples of the current industry need for time to event analysis includes survival analysis for human subjects in a clinical trial environment, time to failure analysis in both cardiac stent and orthopedic device implantation, time to death toxicology profile for a dose level study in preclinical studies. With this great need for reliable measures to make risk/benefit assessment in the pharmaceutical and medical device industries there is a corresponding great need for software applications for analysts and clinicians. This presentation is an application which illustrates how SAS macro processing can be used with the statistical procedures of LIFETEST and PHREG to provide the flexibility needed for time to event analysis. This application further can be used to illustrate how SAS data step processing can be used for data preparation and graphics production in two different scenarios of time to event analyses. This presentation will be most valuable for experienced SAS programmers with exposure to macro processing.

INTRODUCTION

Time to event analysis, commonly referred to as survival analysis, is a standard in many industries and is used for a number of analysis needs. In a clinical trial, this analysis aids in the risk benefit assessment for any clinically relevant event, most often to meet objectives in providing reliable safety and efficacy for an experimental treatment. Survival analysis can be used in a comparison between a new treatment and a placebo or an active treatment in a randomized trial. In the orthopedic medical device industry treatment can include implants, procedures and/or systems. For a joint (i.e., knee or hip) replacement or spine device, longevity of placement is the usual assessment of efficacy and safety so time to failure would be the event of interest for a survival analysis. In the study of implanted cardiac devices survival analysis is used to analyze both death and any clinically meaningful measure where the event of interest is measured over a specified time interval. A particular example is revascularization of a vessel after initial stent implantation.

A common aim of survival analysis is to estimate causal or predictive models in which the risk of an event depends upon covariates. In addition to a treatment comparison, how to incorporate covariates into the analysis needs to be considered in exploring statistical methods.

The purpose of this investigation is to look closely at some of the considerations that should be used when implementing time to event analyses especially in the clinical trial application. Methods that can be implemented in a production environment, both easily understood as well as easy to access by analysts, will be explored. Additionally differential statistical treatment will be highlighted based on two different data scenarios. Each of these will be used to illustrate considerations and assumptions when using SAS to perform the statistical analysis.

This paper will specifically look at a method to assess the effect of treatment and to locate influential covariates. The data are simulated and expected similar to outcomes that might be found in the clinical areas of medical device implantation.

CLINICAL TRIAL ENVIRONMENT

Each environment where survival analysis is being done will have specific needs. In a clinical trial setting, right censoring is of interest since the trial will end and some of the patients will not have the event of interest over the time of the trial. In this scenario it will also be the case that the subject may have died or be lost to follow-up before the end of the trial. All standard methods of survival analysis have the assumption that censoring is random. Those censored at some time t_i should be representative of all subjects still alive at time t_i, $i = 1,2,…,n$, where i represents the underlying survival time and n is the number of observations in the sample of interest (Allison, 2010). For the purpose of this paper, considerations for the time to event analysis will have two emphases in the context of assessing treatment effect for a time to event analysis. First, considerations made in the analysis plan when baseline covariates are available, second, considerations made when using time-dependent covariates.

KAPLAN-MEIER MODEL

Time to event analysis was first studied by actuaries. The Kaplan-Meier (KM) estimator is the most widely used method for estimating survivor functions (Allison, 2010). This method, also known as the product-limit estimator, was used many years prior to 1958 at which time Kaplan and Meier showed that it was, in fact, the nonparametric maximum likelihood estimator. The Kaplan-Meier estimate of the survivorship function is the most frequently used estimator and it is the only one routinely calculated in software packages (Hosmer, Lameshow, and May, 2008).

COX REGRESSION MODEL

The Cox regression model was introduced by Sir David Cox who was the first to introduce the concept of adding covariates into the time to event model (Fisher & Lin, 1999). This model is also commonly referred to as the *Cox* proportional hazard model and has become the standard of analysis in clinical analysis for time to event outcomes. One of the many advantages to this methodology is that it is a semi-parametric model and as such allows the analyst to avoid some model checking and possible misspecification of distribution which is required for use in a parametric model. Cox proportional hazards model is a semi-parametric model, with a baseline hazard function, but one that does not need to be known in order to perform the analysis. In this analysis, one or more predictor variables, called covariates, can be used to understand the impact to or predict the status of the time to event variable.

SAS EXAMPLES

The use of the SAS macro, which was designed to provide the analyst with a tool to get survival analysis results, will now be explored. This macro was developed using SAS, Version 9.2. In addition to offering data analysis capabilities which were built into the macro being presented, the data step programming capabilities offered by SAS facilitates preparation for data for modeling. The advantages of using this macro are two-fold. For the statistician, this is a matter of convenience to have Proc LIFETEST, test for assumptions, and Proc PHREG all in the same place for execution of analyses. For the SAS programmer, it offers the ability to perform analysis with minimal supervision from the statistician once the sequence of analysis has been illustrated.

The developed macro has three major capabilities. First, the macro can be used to generate time to event survival rates using the Kaplan-Meier model and produce a graph of survival. Second, the macro has the capability to compare two treatment groups, for a particular time to event outcome. Third, the macro can be used to perform Cox Regression analysis, including a test for assumptions and use of covariates in the analysis.

The examples selected for presentation have the assumption that the statistical test of interest will assess treatment differences in the two study groups. Thus, the statistical hypothesis has been specified with the null hypothesis of survival estimates in treatment group are equal and the alternative is that they are not equal, that is, H_o: $S(t)_1 = S(t)_2$ vs H_1: $S(t)_1 \neq S(t)_2$, where $S(t)i$ is the estimated survival function for group i, i = 1,2. The testing level is assumed to be 0.05.

The macro was designed to use input parameters to drive the distinctions of usage. The parameters used in the macro are identified in **Table 1**.

Macro Parameter Name	Usage	Example
DSIN	Identify dataset of usage	DSIN = derived.survival
TYPEANAL	Specify the type of analysis {K_M, Graph, Tst_Assump, Cox}	TYPEANAL=K_M
RTFOUT	The name of the RTF file when TYPEANAL = Graph	RTFOUT=C:\OutStats.rtf
ANALVAR	Name of any variable, other than treatment, that will be used as a covariate	ANALVAR=age
BYVARS	Variables that might be used to perform multiple analyses	BYVARS=param
CENSORVAR	Name of the variable that contains the censor indicator	CENSORVAR=censor
CENSORVAL	The value of CENSORVAR that indicates a censored observation	CENSORVAL=0
EVENTVAR	Name of the variable that contains follow up or event time	EVENTVAR=fuptime
INTRACTNS	Name of variable that is used for interaction with time. Except for trt_time, this will be paired with the variable in ANALVAR	INTRACTNS=age_time
PATVAR	Name of the variable that contains subject identifier	PTVAR=pt
TRTVAR	Name of variable that contains treatment	TRTVAR=trt

**Table 1. Description of Macro Parameters**

KAPLAN-MEIER

Because the Kaplan-Meier model is non-parametric, there are few assumptions to check before usage. The only check is availability of key data and variables needed for analysis. Each subject needs to have a single time to event or time of either follow up or censoring which is associated with the outcome for the event of interest. For example, for each subject, if time to death is the outcome of interest then either time of death or time of last follow up needs to be known to determine censoring status for the subject. If treatment effect is of interest then the treatment group identifier has to be known.

In this application of the macro, the SAS procedure, LIFETEST, is being used to generate Kaplan-Meier (KM) estimates to produce a graph of survival curves as well as a statistical comparison between treatment groups. Proc LIFETEST was designed for univariate analysis of timed events. It also can also accommodate covariates, but does not produce estimates of parameters. **Macro Call 1** illustrates the macro parameter assignments which were used to produce a graph and generate test for treatment differences. Output 1 contains the resulting graphical output with the illustrated survival curves.

Macro Call 1:

```
%SurvivalAnalysis (DSIn     = Survival,
                   TypeAnal = Graph,
                   RTFOut   = %str(<path for .rtf file>),
                   AnalVar  = ,
                   ByVars   = ,
                   CensorVar = var1,
                   CensorVal = 0,
                   EventVar  = var1day,
                   IntrActns = ,
                   PatVar   = Pt,
                   TrtVar   = Trt);
```
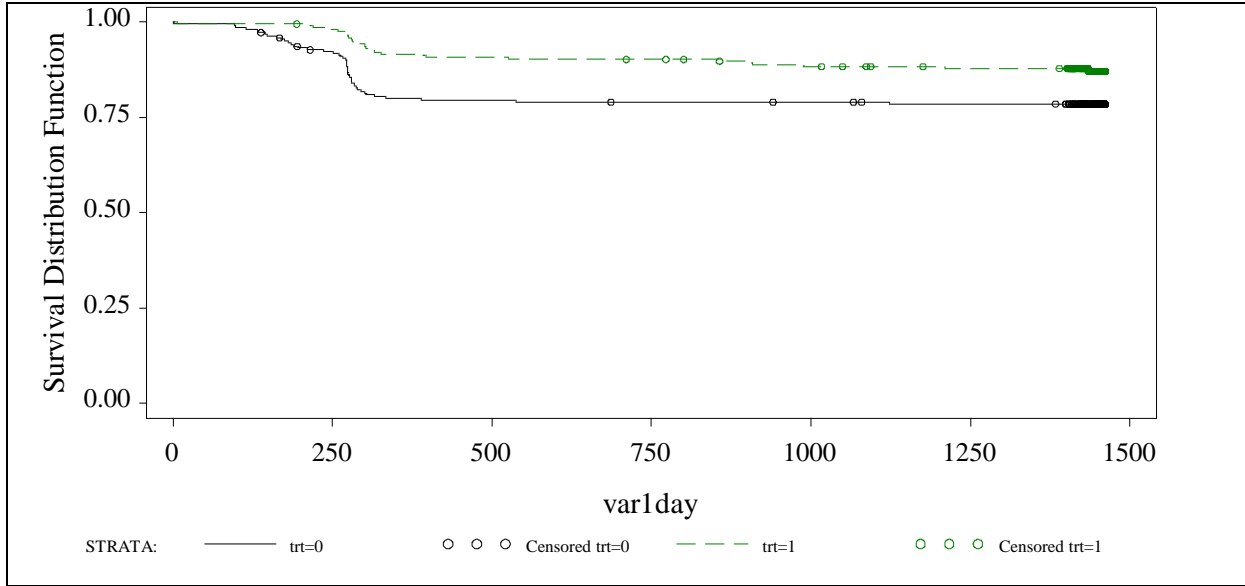
Output 1. Survival Graph by Treatment

Macro Call 1 also produces statistical results to test for treatment effect, from which a report can be generated. The results of this testing indicate, in all tests, that the two treatments are statistically different (p < 0.01) at most levels of testing. Output 2 illustrates these results. Proc LIFETEST provides three different tests, the Log Rank, the Wilcoxon, and the likelihood-ratio test (-2Log(LR)). Each of these tests use a different test statistic which is calculated with varying weights attached to observations. The generalized Wilcoxon test uses weights equal to the number at risk so it puts relatively more weight on differences between survival functions at earlier times. The log-rank test uses weights equal to 1 so this test will place more emphasis at observations that occur at later points in time. The likelihood-ratio test (-2log(LR)) requires the assumption of the hazard function is constant. For this example, all tests results in supporting the same conclusion.

```
            Survival Testing for Treatment Difference


     Obs       Test       Chi Square      DF   P-Value
      1      Log-Rank        6.9787         1    0.0082
      2      Wilcoxon        8.2630         1    0.0040
      3      -2Log(LR)       7.1771         1    0.0074
```

Output 2. Report generated by LIFETEST to test for treatment differences

COX REGRESSION

Assumptions for Cox Regression model

Use of the Cox Regression model requires a check for assumptions. This model is very powerful for analyzing survival outcomes with covariates in the analysis. The goal of the estimated regression model is to produce a summary presentation and interpretation of what factors have influenced survival or the event of interest, that is, estimated hazard ratio (HR) and/or estimated covariate adjusted survivorship functions. Hosmer et al (2008) noted that the appeal of the Cox model is the ease of interpretation by use of a relative risk ratio.

Cox regression is implemented in SAS using Proc PHREG. Proc PHREG uses Cox's partial likelihood method to estimate regression models with censored data (Hosmer et al, 2008). This model allows for an estimation method for time-dependent covariates. Proc PHREG can have both continuous-time and discrete-time data variables in the model. Only Proc PHREG has the special capability for analysis on repeated events like arrests, hospitalization, cardiac events, or orthopedic implant revisions.

Allison (2008) noted that the proportional hazards model assumes that the hazard ratio of two groups being tested is independent of time. Hazard is a characteristic of individuals, not populations or samples. It can be thought of as a dimensional quantity that has the form number of events per interval of time, sometimes referred to as a rate. This

assumption requires that any covariates added to the model be tested to check if there is any dependency between it and time. This can be done by using an interaction of time with the covariate of interest to see if the interaction term is significant in the model (i.e. H_o: $b_i = 0$, where b_i is the coefficient associated with the covariate x_i in the term $b_i x_i$ in the model, and i= 1, n indicates the number of covariates entered into the model.). If H_o is rejected, it can be concluded that Cox's proportional hazard model is not appropriate for the given covariate.

A log cumulative hazard plot can also be used to diagnose the validity of the proportional hazards assumption. This diagnosis procedure was built into this SAS macro. Previously, Proc LIFETEST was used to produce a plot of the same data that was just used in the K-M survival plot, see **Macro Call 1** and graph in Output 1. In the second call, Proc LIFETEST will be used to illustrate one method to check the hazards ratio assumption.

To test the assumptions of proportional hazards, a log(log) graph was produced in Proc LIFETEST. The details of the macro call are illustrated in **Macro Call 2** and the output graph is contained in Output 3. The graph indicates non-proportional hazards thus indicating Cox method is not appropriate for use of the variable treatment in the model.

Macro Call 2:

```
%SurvivalAnalysis (DSIn       = Survival,
                    TypeAnal  = Tst_Assump,
                    RTFOut    = %str(<local path and .rtf name for graph>),
                    AnalVar   = ,
                    ByVars    = ,
                    CensorVar = var1,
                    CensorVal = 0,
                    EventVar  = var1Day,
                    IntrActns = ,
                    PatVar    = Pt,
                    TrtVar    = Trt
                   );
```
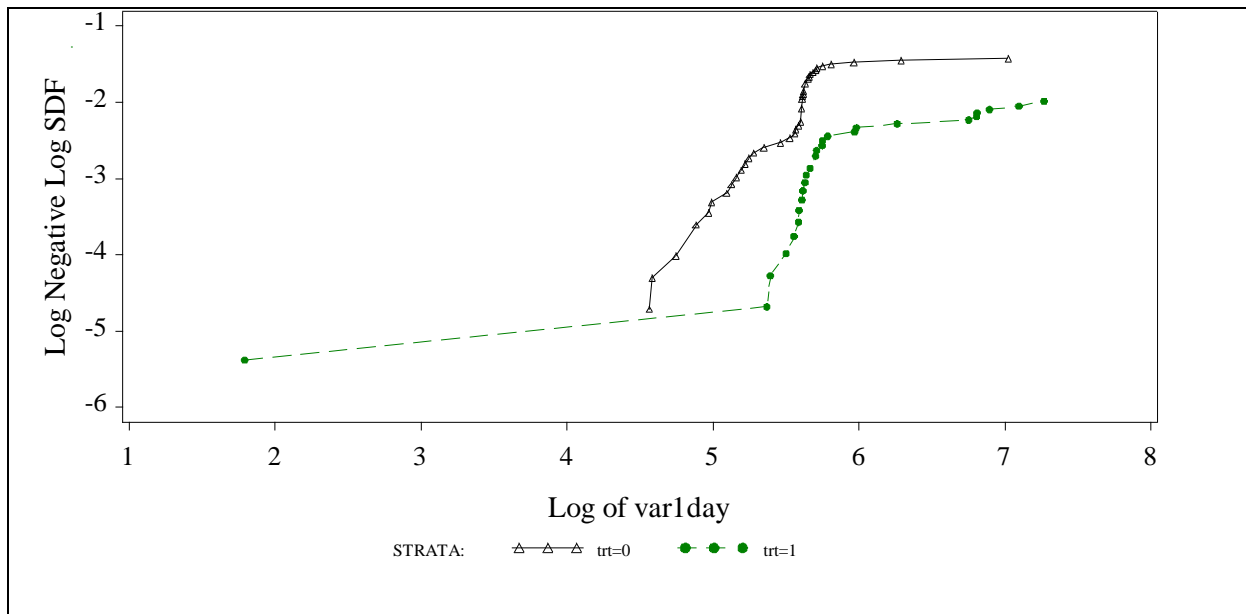

Output 3. Proc LIFETEST used to generate KM curve of log … test of assumptions.

A second method to check for the proportional hazards assumption is to use an interaction term, of time with the variable of interest, in the Cox Regression model. Proc PHREG will be used with an interaction term of treatment with time. **Macro Call 3** illustrates the details of the call to the macro, using the INTRACTNS parameter, indicating the interaction of interest. The results of this macro call produced a SAS output report, which can be converted to a report using ODS, see Output 4. This testing of the interaction term produces the same result as was evidenced in Output 3, that is, the treatment variable does not meet the HR assumption of proportional hazards, p = 0.0129. Thus, the testing that was done in Proc LIFETEST will suffice for the treatment difference hypothesis. Covariate analysis for this

data will not be pursued. The null hypothesis is rejected and a statistically significant treatment difference is inferred.

Macro Call 3:

```
%SurvivalAnalysis (DSIn      = Survival,
                   TypeAnal  = Cox,
                   RTFOut    = ,
                   AnalVar   = trt,
                   ByVars    = ,
                   CensorVar = var1,
                   CensorVal = 0,
                   EventVar  = var1day,
                   IntrActns = trt_time,
                   PatVar    = Pt,
                   TrtVar    = Trt);
```

```
                              The PHREG Procedure

                     Analysis of Maximum Likelihood Estimates

                    Parameter    Standard                           Hazard
Parameter   DF      Estimate        Error   Chi-Square  Pr > ChiSq   Ratio  Label

trt          1      -1.80134      0.50790      12.5786      0.0004    0.165  Treatment(0-Trt0,1-Trt1)
Trt_Time     1       0.00358      0.00144       6.1875      0.0129    1.004
```

Output 4. Output from proc PHREG to test for HR assumption

Covariates in the Cox Model

To determine use of covariates in a particular model, the HR assumption will be assessed for each covariate. For the first set of data, it was illustrated that the time to event did not meet the HR assumptions thus this same dataset cannot be used in a Cox model with covariates. A different dataset will be used to illustrate use of covariate analysis in the Cox Model. A check for the HR assumption for a variable of age will be done and then used as a covariate in the model that contains treatment. The new set of data also has two treatment groups of interest and the parameter of interest is time to revision in an orthopedic clinical trial.

Macro Call 4 is made to check the HR assumption for the variable age. Output from that check is contained in Output 5. A similar check is assumed for variable treatment which resulted in HR assumption being met.

Macro Call 4:

```
%SurvivalAnalysis (DSIn      = survival_ortho,
                   TypeAnal  = Cox,
                   RTFOut    = ,
                   AnalVar   = age,
                   ByVars    = ,
                   CensorVar = rev_yn,
                   CensorVal = 0,
                   EventVar  = daysfu,
                   IntrActns = age_time,
                   PatVar    = id,
                   TrtVar    = Trt
                  );
```

```
                              The PHREG Procedure

                     Analysis of Maximum Likelihood Estimates

                        Parameter    Standard                           Hazard
        Parameter   DF   Estimate        Error   Chi-Square   Pr > ChiSq   Ratio
```

trt	1	-0.79957	0.44841	3.1795	0.0746	0.450
AGE	1	-0.02280	0.02357	0.9359	0.3333	0.977
age_time	1	-0.0000108	0.0000193	0.3137	0.5754	1.000

Output 5. Output from proc PHREG to test for interaction of age and age*time, check HR assumption

Since the interaction of age and time was not significant, p = 0.5754, age is allowed for use in the model as a covariate, and further to check for the coefficients in the model. The interaction term will be removed from the list of covariates considered in this analysis.

Macro Call 5, not illustrated, will be just like **Macro Call 4** except that no interaction is being tested. Only treatment and age will be included in the model. Output 6 illustrates the output associated with this call. In these results, the partial likelihood ratio test, the Wald test, or the Score test can be used to assess the significance of any coefficient. The standard error can be used to construct confidence intervals if needed. Note that age did result in being a significant covariate in this analysis once the interaction term was removed from the model. Conclusion about treatment difference is the same for both models, that is, there was not sufficient evidence, based on this data, to reject the null hypothesis resulting in the conclusion that there were no treatment differences, p = 0.0748. Finally, the Hazard Ratio can aid in interpretation of the differences between the treatments. The Hazard Ratio value of .45 indicates that the hazard for those undergoing treatment 1 is about 45% of the hazard for those undergoing treatment 2, Even though the difference in treatment was not statistically significant, treatment 2 is a little more than twice as likely to have failure in the implant.

```
                         The PHREG Procedure

                  Analysis of Maximum Likelihood Estimates

                      Parameter     Standard                              Hazard
      Parameter   DF   Estimate        Error    Chi-Square    Pr > ChiSq    Ratio

      trt          1   -0.79829      0.44806        3.1743        0.0748    0.450
      AGE          1   -0.03300      0.01449        5.1875        0.0228    0.968
```

Output 6. Output from proc PHREG to test for significance of covariates

CONCLUSIONS

The purpose of this investigation was to outline considerations when implementing time to event analyses in a clinical trial application. A SAS Macro was designed and used to illustrate a method that could be applied in a production environment. The macro was easy to use and applied to the process that the analyst uses when implementing survival analysis. Two different data scenarios were used to illustrate considerations and assumptions for using SAS to aid in the statistical analysis. Often an analyst is called to design models which can assist in determining treatment differences as well as how that treatment risk might depend on covariates. This macro was designed to help with assessment of treatment differences as well as addition of covariates which might be considered in exploring the impact of treatment.

REFERENCES

Allison, Paul D. 2010. *Survival Analysis Using SAS: A Practical Guide, Second Edition*. Cary, NC. SAS Institute, Inc.

Fisher, LD and Lin DY. 1999. "Time-dependent covariates in the Cox proportional-hazards regression model." *Annu Rev Public Health 20.* 145-157.

Hosmer, D., Lameshow, S. and May, S. 2008. *Applied Survival Analysis: Regression modeling of time to event data*. Hoboken, NJ. John Wiley & Sons, Inc.

Nardi A. and Schemper M. 2003. "Comparing Cox and parametric models in clinical studies." *Statistics in Medicine 22*. 3597-3610.

ACKNOWLEDGMENTS

RECOMMENDED READING

- Survival Analysis using SAS: A Practical Guide

- SAS/STAT User's Guide

CONTACT INFORMATION

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

Name: Helen Chmiel or Evan Ritzema
Enterprise: Experis
Address: 5220 Lovers Lane, Suite 200
City, State ZIP: Kalamazoo, MI 49002
Work Phone: 269-553-5183 (Chmiel) and 269-553-5168 (Ritzema)
E-mail: helen.chmiel@experis.com or evan.ritzema@experis.com
Web: www.experis.com