

# Memorandum

**To:** Dr. Ramon Leon  
**From:** Carrie McBrayer and Karin Kitchens  
**Date:** November 23, 2008  
**Re:** Primary Biliary Cirrhosis

## Executive Summary:

This data set is taken from the Mayo Clinic during a ten year interval that spans from 1974 to 1984. There are 312 patients that participated in the randomized trial in which some were given a placebo instead of the drug D-penicillamine. In addition to the survival data, there are 13 other explanatory variables in this clinical trial. To make the data set useful for this study, we were just concerned with the last observation for each patient. The purpose of the study is to determine if the drug D-penicillamine had any affect on survival time of the patients involved in the study. In addition, we are trying to find a model to predict survival time. To see a description of each variable, please look to the appendix.

Here is a summary of the number of events used in the test (Note: many observations were deleted because they had missing values for certain variables).

## Summary of the Number of Event and Censored Values

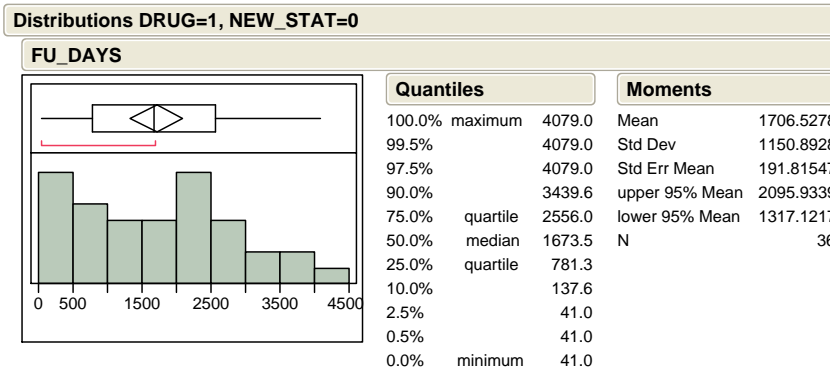
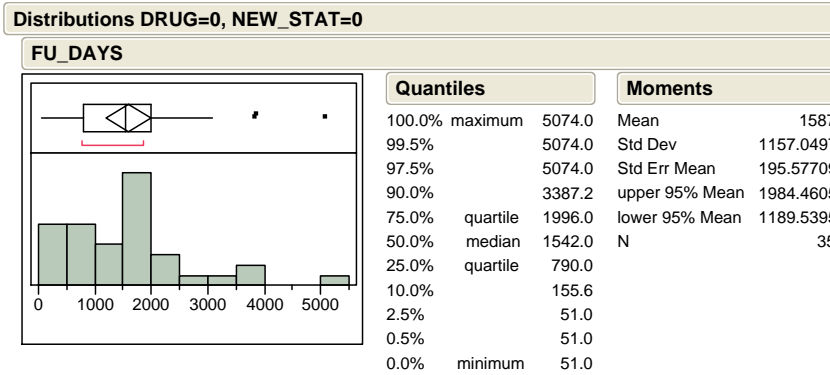
TOTAL	EVENT	CENSORED	PERCENT FAILED
185	71	114	38.83

**Results:**

*Descriptive Statistic Analysis:*

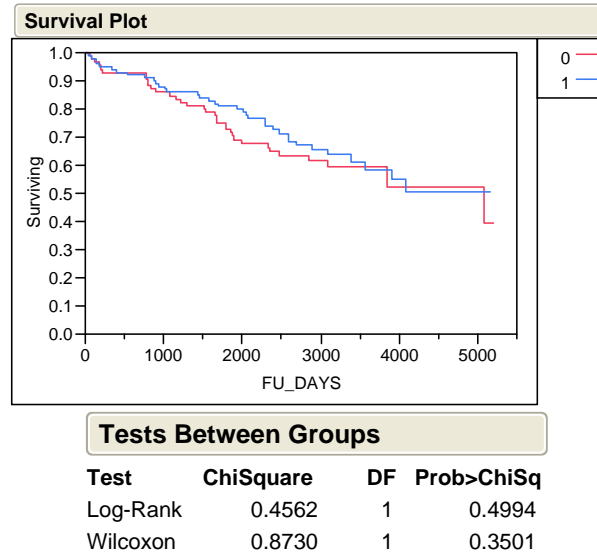
We did a simple analysis in JMP to see how the drug related to survival time. Below are two histograms grouped by drug and status. Our findings are summarized in the table below in terms of median time to failure:

	Drug Present	No Drug
Failure	1673.5	1542



Based on the table and plots above, it appears that the drug might not have a significant relationship to the median time to failure. Initially it appears that the drug extends the time to failure based on the median, but this analysis does not include any explanatory variables. Therefore, further analysis is needed.

### Comparison of Survival Probability with and without Drug



Based on the survival plot, it appears that there is not a difference in the survival rate when a patient is given the drug or given the placebo. This is also indicated by the large p-values in the Tests Between Groups, which tests for homogeneity between survival times.

Before any conclusions can be made, we need further analysis that includes explanatory variables to prove or disprove this relationship. In this study, there were 13 variables thought to contribute to survival time of the patients in the clinical trial. However, it was unknown if each of these variables provided any useful information for modeling survival time of patients. It would be appropriate to use the proportional hazard model to fit this data since we are dealing with clinical trial data. To find the appropriate model, we used a SAS program to do a stepwise regression to find the best fit proportional hazard model for this data.

### Proportional Hazard Assumptions

For the proportional hazard model there are no assumptions made about the shape of the underlying hazard function. However, there are two assumptions needed. The model must specify a multiplicative relationship between the underlying hazard function and the log-linear function of the covariates. This is also known as the proportionality assumption. It is assumed that given two observations with different values for the independent variables, the ratio of the hazard functions for those two observations does not depend on time. The second assumption is that effect parameters multiply hazard. If the proportional hazards assumption holds (or is assumed) then it is possible to estimate the effect parameters without any consideration of the hazard function.

Before the assumptions can be checked, a proportional hazard model must be chosen. The assumptions analysis is included after the model selection.

This is the Cox proportional hazard model. The proportional hazards model is represented by

$$h(t) = [h_0(t)] e^{(b_1X_1 + b_2X_2 + \dots + b_kX_k)}$$

### Model Selection Process:

#### **Model Fit Statistics**

	FULL	REDUCED
AIC	603.753	592.788

We compared the AIC value for the full model (all 13 variables) to the reduced model (5 variables). One can see that the AIC is smaller for our reduced model, which is what we expected.

### Model Significance

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	95.0182	5	<.0001
Wald	87.4482	5	<.0001

Based on the likelihood ratio and Wald test, we have a statistically significant relationship between the time in days and the model.

### Proportional Hazard Model Selected Through Stepwise Regression

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq
Age_in_days_	1	0.0001137	0.0000347	10.7155	0.0011
Sex	1	-1.06800	0.37336	8.1824	0.0042
alkaline_phosphatase	1	0.00101	0.0001693	35.8030	<.0001
bilirubin	1	0.05918	0.02135	7.6832	0.0056
edema	1	1.55789	0.35313	19.4632	<.0001

Given these parameter estimates, one can use the proportional hazard model to predict survival time:

$$h(t) = [h_0(t)] e^{(0.0001137 * X_1 - 1.06800 * X_2 + 0.00101 * X_3 + 0.05918 * X_4 + 1.55789 * X_5)}$$

It is interesting to note the explanatory variables chosen, and even more importantly, what was not chosen. Drug does not seem to be statistically significant for this model. However, age of the patient, gender, level of serum bilirubin, presence of edema, and level of alkaline phosphatase are all considered to be statistically significant.

We used SAS (see SAS program in Appendix) to check the proportional hazard assumption. This was accomplished by finding the correlation between Schoenfeld residuals for a particular covariate and ranking the individual failure times. If the proportional hazard assumption is met, then the correlation between residuals and ranked times should be near 0.

The null hypothesis for these tests is  $H_0: \rho = 0$ . We expect to have large p-values for these tests in order for the proportional hazard assumption to be met. The following SAS output includes the p-values for these tests for each variable.

```

                                The CORR Procedure
1 With Variables:      timerank
3   Variables:      rage_in_days_      ralkaline_phosphatase      rbilirubin

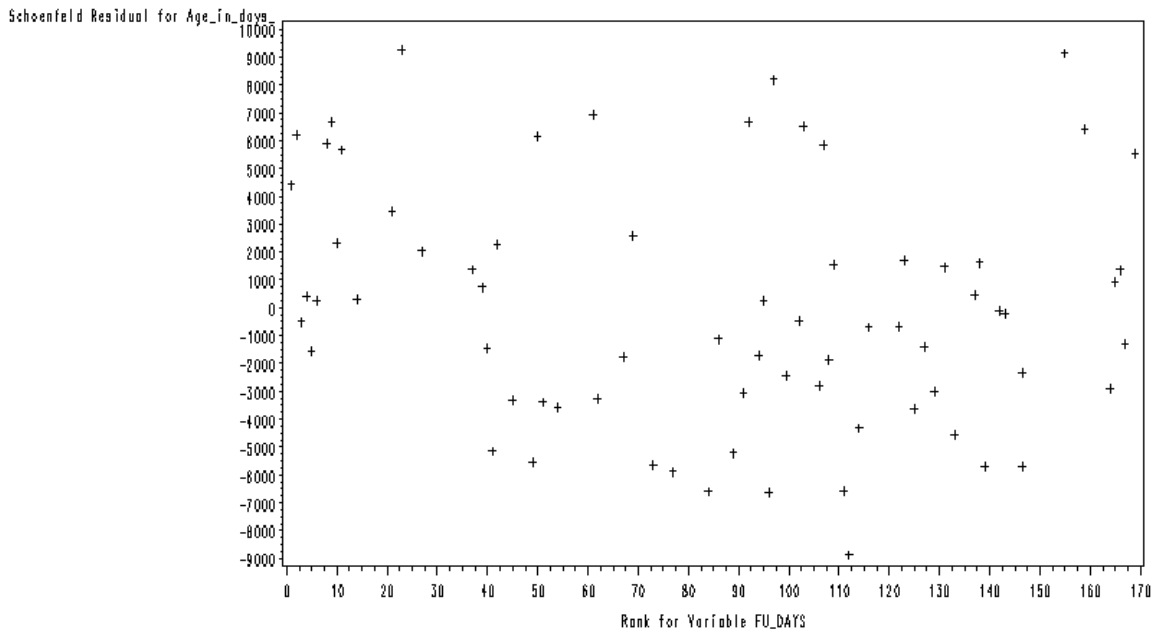
                                Pearson Correlation Coefficients
                                Prob > |r| under H0: Rho=0
                                Number of Observations
                                rage_in_      ralkaline_      rbilirubin
                                days_      phosphatase
timerank      -0.18903      0.19148      -0.06027
Rank for Variable FU_DAYS 0.1144      0.1097      0.6176
                                71      71      71

```

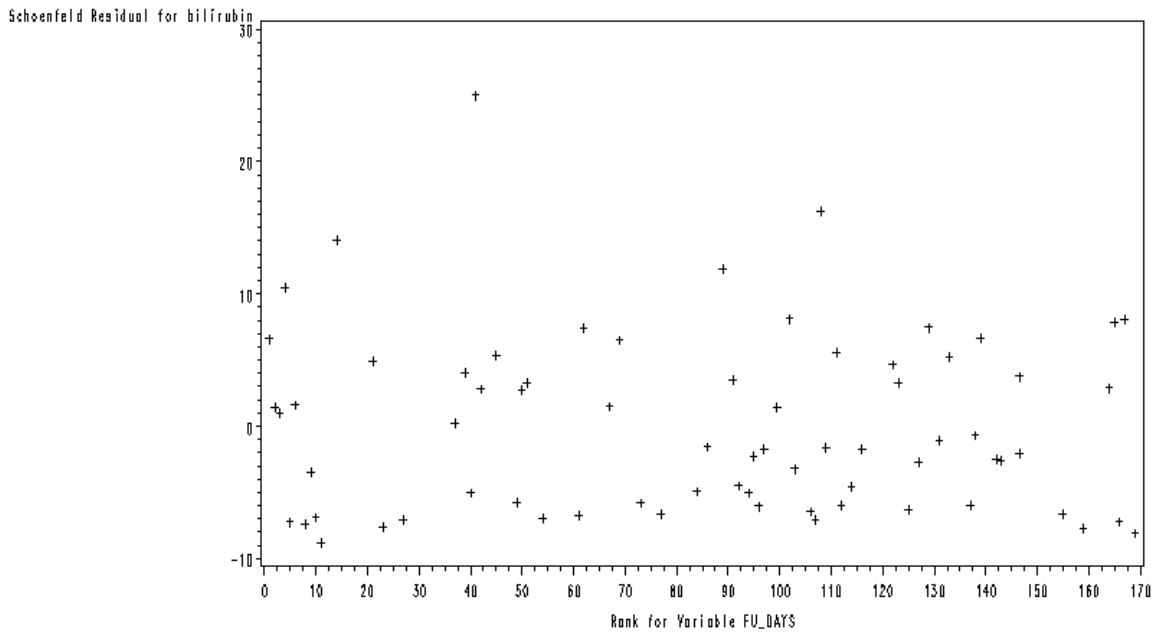
The p-values for these tests are all greater than .05. Therefore, we don't have evidence to conclude that the correlations between the ranked times and the explanatory variables' residuals are different from 0. This means that the proportional hazard assumption holds.

We can also check the proportional hazard assumption with plots of the residuals vs. ranked time for each explanatory variable. We expect a random scatter in these plots to ensure that the assumption has been met.

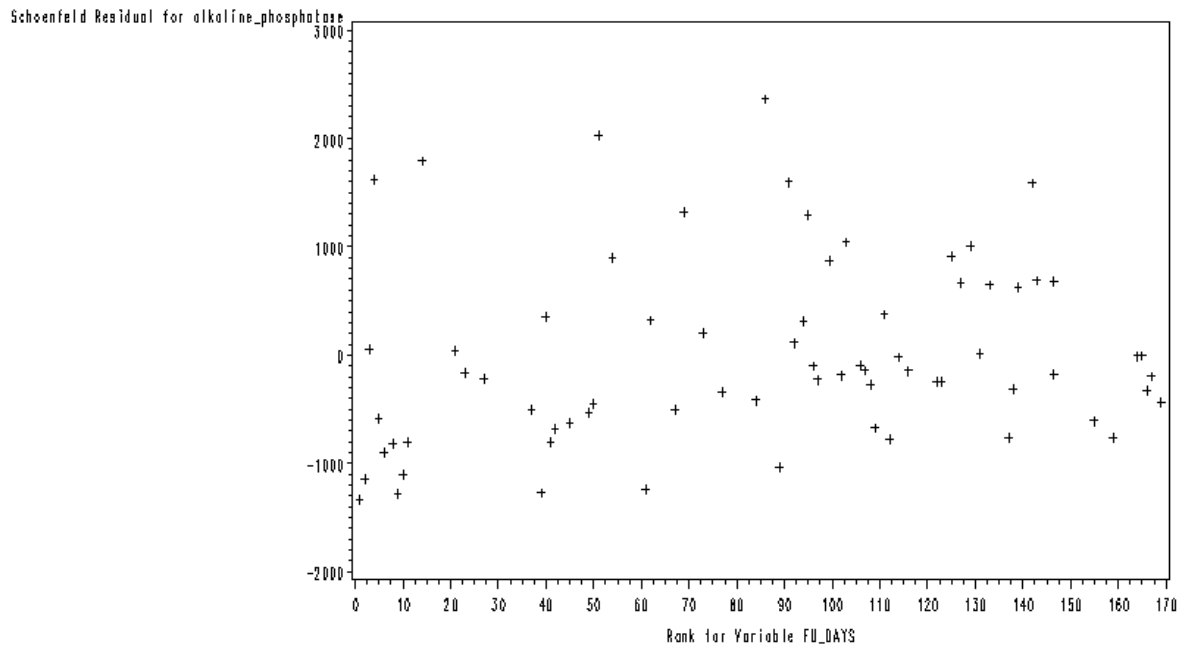
Shoenfeld residuals (for age in days) plotted against failure time



Shoenfeld residuals (for bilirubin) plotted against failure time



### Schoenfeld residuals (for alkaline phosphatase) plotted against failure time



There does seem to be a random scatter in each of these plots which reconfirms that the proportional hazard assumption holds.

### Risk Ratios:

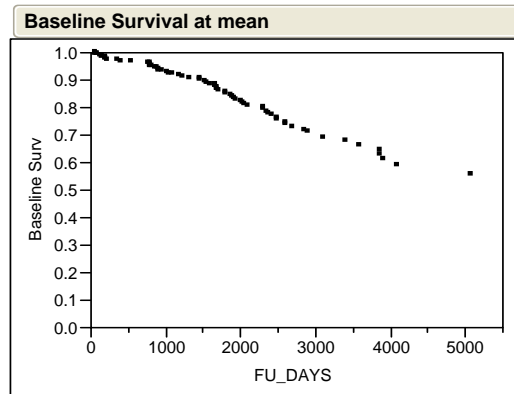
#### Unit Risk Ratios

Per unit change in regressor

Term	Risk Ratio	Lower CL	Upper CL	Reciprocal
AGE__IN_	1.000114	1.000045	1.000181	0.9998863
SEX	0.343696	0.170376	0.744953	2.9095488
EDEMA	4.748801	2.382055	9.523894	0.2105795
BILIRUBI	1.060968	1.01638	1.10528	0.9425356
ALKALINE	1.001013	1.000679	1.001345	0.9989875

The hazard or risk ratio is the effect of an explanatory variable on the hazard of an event. The hazard ratio is the effect on the hazard rate of a difference, such as being gender, as estimated by the regression model. For two individuals who differ only in the relevant membership (i.e. gender) their predicted log-hazard will differ additively by the relevant parameter estimate. Therefore their predicted hazard rate will differ by  $e^{\beta}$ . The estimate can be considered a hazard ratio. To put it more simply, it is the ratio between the predicted hazard for a member of one group and that for a member of the other group, holding everything else constant. When it is in terms of continuous explanatory variables, it is interpreted as a unit difference. As one can see from the output, one unit increase in edema while one is holding everything else constant the hazard function increases by 4.7488 units.

### Baseline Survival Plot



The baseline survival plot plots the number of days against the mean probability of survival. The average probability of survival decreases overtime, as one would expect with this type of data.

### Conclusion

Our initial objective for this analysis was to try to find a relationship between survival time and the presence of the drug penicillamine. After fitting a proportional hazard model to this data, we discovered that penicillamine is not statistically significant in the model. Instead, we found 5 other explanatory variables that sufficiently predicted survival time.

## Appendix A:

Variables included in the full model:

- **Case Number**
- **Number of days between registration and the earlier of death, transplantation, or study analysis time**
- **Status:** 0=alive, 1=transplanted, 2=dead
- **Drug:** 1= D-penicillamine, 0=placebo
- **Age in days,** at registration
- **Sex:** 0=male, 1=female
- **Day:** (number of days between enrollment and this visit date, remaining values on the line of data refer to this visit)
- **Presence of ascites:** 0=no 1=yes  
It is an accumulation of fluid in the peritoneal cavity.
- **Presence of hepatomegaly** 0=no 1=yes  
It is the condition of having an enlarged liver.
- **Presence of spiders** 0=no 1=yes
- **Presence of edema** 0=no edema and no diuretic therapy for edema; .5 = edema present without diuretics, or edema resolved by diuretics; 1 = edema despite diuretic therapy. It is the abnormal accumulation of fluid beneath the skin.
- **Serum bilirubin** in mg/dl:  
Serum bilirubin is often considered to be a true test of liver function, as it reflects the liver's ability to take up, process, and secrete bilirubin into the bile.
- **Serum cholesterol** in mg/dl
- **Albumin** in gm/dl  
Serum albumin is the most abundant blood plasma protein and is produced in the liver and forms a large proportion of all plasma protein.
- **Alkaline phosphatase** in U/liter  
It is a bacteria found in the body.
- **SGOT** in U/ml (serum glutamic-oxaloacetic transaminase, the enzyme name has subsequently changed to "ALT" in the medical literature) An enzyme that is normally present in liver and heart cells. SGOT is released into blood when the liver or heart is damaged. The blood SGOT levels are thus elevated with liver damage or with an insult to the heart
- **Platelets** per cubic ml / 1000
- **Prothrombin** time in seconds  
This is used to determine clotting time in seconds of blood.
- **Histologic stage of disease**

## Appendix B:

### SAS Program Used:

```
/* ----- PRIMARY BILIARY CIRRHOSIS ANALYSIS -----  
-----*/  
  
/***** Performing a statistical test for the PH assumption  
******/
```

```
/* The following SAS program makes use of the Biliary Cirrhosis data  
set to demonstrate how to perform a statistical test of the  
proportional hazard assumption for a given covariate. This is  
accomplished by finding the correlation between the Schoenfeld  
residuals for a particular covariate and the ranking of individual  
failure times. If the proportional hazard assumption is met, then the  
correlation should be near zero. The p-value for testing this  
correlation can be obtained from proc corr (or proc reg). The  
Schoenfeld residuals for a given model can be saved in a SAS dataset  
using proc phreg. The ranking of events by failure time can be saved in  
a SAS dataset using proc ranked. The null hypothesis is that the PH  
assumption is not violated.
```

Note that the null hypothesis is never proven with a statistical test. The most that may be said is that there is not enough evidence to reject the null. A p-value is driven by sample size. A gross violation of the null assumption may not be statistically significant if the sample is very small. Conversely, a slight violation of the null assumption may be highly significant if the sample is very large.

First run the full model. The output statement creates a SAS data set, the out= option defines an output dataset, and the ressch= statement is followed by user defined variables names so that the output dataset contains the Schoenfeld residuals. Note that the order of the names match up with the order of the independent variables in the model statement. The actual variable names are arbitrary.

```
*/ */  
  
/* Stepwise Model */  
Proc PHReg data=perm.Pbc2;  
    model fu_days*new_status(0)= albumin age_in_days_ chol drug  
        SGOT sex alkaline_phosphatase asictes bilirubin edema  
        hepatomegaly platelets prothrombin spiders stage  
    / selection=stepwise slentry=0.25  
      slstay=0.05 details;  
    output out=resid  
      ressch=ralbumin rage_in_days_ rchol rdrug rSGOT rsex  
        ralkaline_phosphatase rasictes rbilirubin redema  
        rhepatomegaly rplatelets rprothrombin rspiders  
        rstage;  
  
Run;  
  
Proc Print data=resid;  
Run;
```

```

/* Create a SAS data set that deletes censored observations. We just
want a correlation with the order of events.  */
Data events;
    set resid;
    if new_status=1;
Run;

/* Use proc rank to create a data set with the ranking of failure
times. The user supplies the name of the output data set (out=). The
variable to be ranked is fu_days (the time variable). The rank
statement precedes a user-defined variable name for the rankings of
failure times. The user-defined names are arbitrary.  */

Proc Rank data=events out=ranked ties=mean;
    var fu_days;
    ranks timerank;
Run;

Proc Print data=ranked;
Run;

/* Proc corr is used to get the correlations between the ranked failure
time variable(called TIMERANK in this example) and the variables
containing the Schoenfeld residuals of age_in_days, sex,
alkaline_phosphatase, bilirubin, AND edema (called rage_in_days_, rsex,
ralkaline_phosphatase, rbilirubin, and redema respectively in this
example). The nosimple option suppresses the printing of summary
statistics. If the proportional hazard assumption is met for a
particular covariate, then the correlation should be near zero. The p-
value for testing the correlation of zero is the p-value for the
testing of the proportional hazard assumption  */

Proc Corr data=ranked nosimple;
    with timerank;
    var rage_in_days_ rsex ralkaline_phosphatase rbilirubin redema;
Run;

/* The same p-values can be obtained with proc reg. */
Proc Reg data=ranked;
    model timerank=rage_in_days_;
Run;

/* Proc gplot can be used to plot the residuals against the ranked
failure times.  */
Proc gplot data=ranked;
    plot rage_in_days_*timerank;
    title 'Shoenfeld residuals (for age in days) plotted against
failure time';
Run;

Proc gplot data=ranked;
    plot rsex*timerank;
    title 'Shoenfeld residuals (for sex) plotted against failure
time';
Run;

```

```
Proc gplot data=ranked;
  plot ralkaline_phosphatase*timerank
  title 'Shoenfeld residuals (for alkaline phosphatase)
        plotted against failure time';
Run;

Proc gplot data=ranked;
  plot rbilirubin*timerank;
  title 'Shoenfeld residuals (for bilirubin) plotted against failure
        time';
Run;

Proc gplot data=ranked;
  plot redema*timerank;
  title 'Shoenfeld residuals (for edema) plotted against failure
        time';
Run;
```

References:

Data came from:

Statlive from Dr. Leon's webpage and can be found in appendix D of Fleming and Harrington, *Counting Processes and Survival Analysis*.

Website used:

<http://www.med.uio.no/imb/stat/forskerkursene/metode/metode2/survival%5B1%5D.%20analysis%20with%20SAS%20program.sas.txt>

Which was largely based on:

F. Harrel & K. Lee, proceedings of the eleventh annual SAS user's group international, 1986, pages 823 - 828.