Survival Analysis of Diabetic Retinopathy Patients

Sriharsha Addepalli, Jaeyun Lee, Lukas Pokhrel

12/01/2018

Introduction/Research Interest

We decided to do an analysis on the survival of laser treatment for Diabetic Retinopathy patients. We found our dataset in the survival package of R[1] [2]. Our dataset contains 197 patients that were a random sample of patients with Diabetic Retinopathy. We are measuring the survival rate until vision loss, which is decided when the vision drops below an acuity of 5/200, after the treatment was done. Our time until failure is measured in months. Additionally, each patient was randomized to see which eye they get laser treatment on and the other eye was treated as a control, so treatment and control are not independent in this dataset. Survival time of the eye vision was measured for both the treatment eye and the control eye. Furthermore, there were censored observations in our dataset.

Our dataset contains the covariates of laser (types of laser treatment used, either argon or xenon), eye (which randomized eye was operated on), age (age of the patient at the time of treatment), type (whether a person is a juvenile or and adult), and risk (a measure of how serious their diabetic retinopathy condition was before the treatment).

Our main point of interest is seeing if there is a difference in vision loss survival rates between the control and the treatment, and this will lead us to the question whether we should give laser eye surgeries to Diabetic Retinopathy patients. We want to fit a final model to test whether any of the covariates affected survival rates, and we especially want to see if type of laser affected the survival rates of the treated eyes and compare that to the control eyes. In our extension, we want to treat the treatment and control as not independent, and use a frailty model to observe the differences between each person using the covariates of our final model.

Model Setup

Because our dataset was from the R survival package, there was not much data wrangling to perform in order for our data to be organized to fit a proper model. This is a data summary table showing the format and distribution of the covariates in our dataset

```
##        id           laser        eye           age              type
##  Min.   :   5.0   xenon:200   Left :197   Min.   : 1.00   juvenile:228
##  1st Qu.: 480.0   argon:194   Right:197   1st Qu.:10.00   adult   :166
##  Median : 834.0                           Median :16.00
##  Mean   : 873.2                           Mean   :20.78
##  3rd Qu.:1296.0                           3rd Qu.:30.00
##  Max.   :1749.0                           Max.   :58.00
##       trt          futime          status            risk
##  Min.   :0.0   Min.   : 0.30   Min.   :0.0000   Min.   : 6.000
##  1st Qu.:0.0   1st Qu.:13.98   1st Qu.:0.0000   1st Qu.: 9.000
##  Median :0.5   Median :38.80   Median :0.0000   Median :10.000
##  Mean   :0.5   Mean   :35.58   Mean   :0.3934   Mean   : 9.698
##  3rd Qu.:1.0   3rd Qu.:54.25   3rd Qu.:1.0000   3rd Qu.:11.000
##  Max.   :1.0   Max.   :74.97   Max.   :1.0000   Max.   :12.000
```

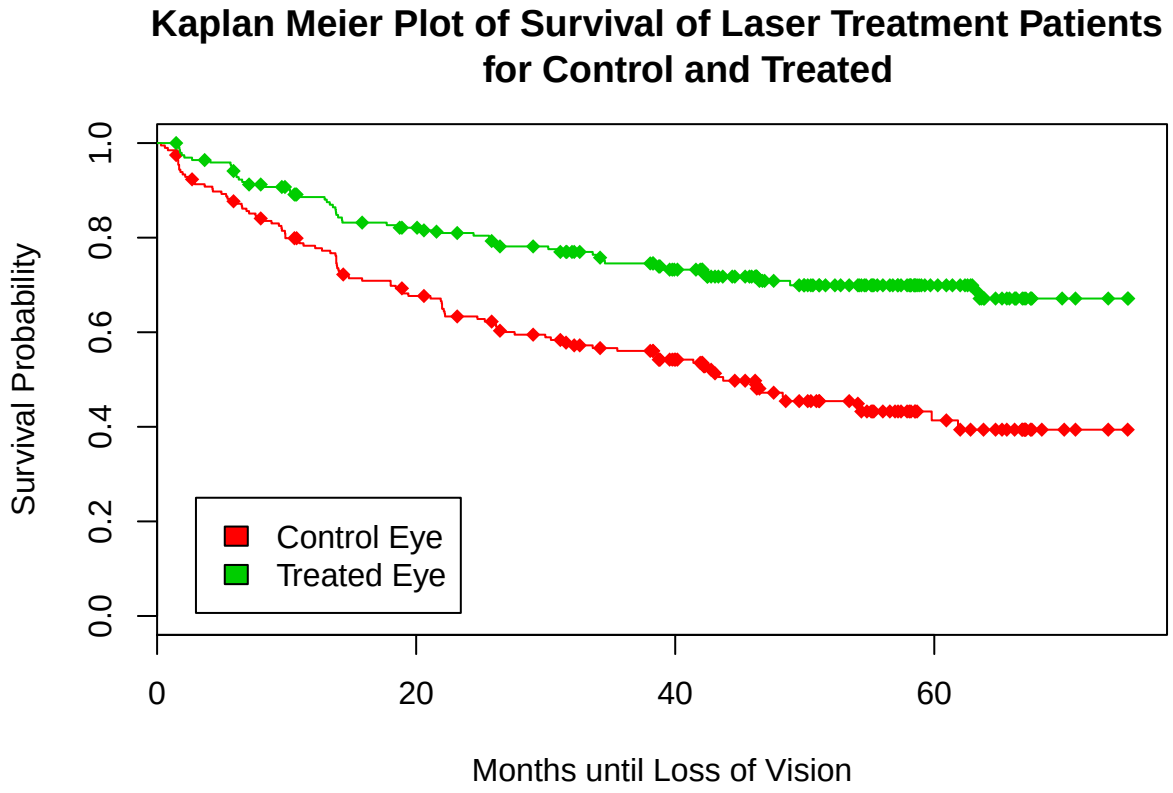
[1] A. L. Blair, D. R. Hadden, J. A. Weaver, D. B. Archer, P. B. Johnston and C. J. Maguire (1976). The 5-year prognosis for vision in diabetes, American Journal of Ophthalmology, 81:383-396.

[2] W. J. Huster, R. Brookmeyer and S. G. Self (1989). Modelling paired survival data with covariates, Biometrics 45:145-156.

As we can see from the summary table above, we have 9 variables in our dataset, and we have 394 observations. Laser, eye, and type are all categorical variables. Additionally, we had no missing values in our dataset.

Basic Estimation

For the simplicity of our intial model, we are going to assume that the control and treatment are independent of one another. We want to first visualize the Kaplan Meier estimates between the treatment and control.



By looking at this plot, we notice that the treatment eye has higher survival rates than the control eye, which is to be expected by any treatment. To test whether this difference is significant, for this basic model, we fit a coxph model to test the covariate treatment to see if the treatment was significant.

```
## Call:
## coxph(formula = Surv(futime, status) ~ trt, data = retinopathy)
##
##       coef exp(coef) se(coef)    z       p
## trt -0.777     0.460    0.169 -4.6 4.2e-06
##
## Likelihood ratio test=22.37  on 1 df, p=2e-06
## n= 394, number of events= 155
```

When fitting a coxph model to test if there is a difference between hazard ratios between treatment and control, we obtain a p-value of 2e-06 with our likelihood ratio test. Because the p-value is lower than 0.05, we reject the null hypothesis. There is significant evidence to suggest that there is a difference between the hazard ratios of treatment and control. With the coxph model, we get that the hazard probability of the treated eye is 0.46 that of the hazard probability of the control eye. This means

that treated eyes have a lower hazard rate than control, meaning the treated eyes have a higher survival probability than control. This shows that what we observed in the Kaplan Meier Plot above is significant.

Model Fitting

In our next model, we still want to see the significance between the treatment and control and also its interaction between type of laser used, but we want to treat the other covariates as confounding variables. We are going to use the Anova function found in the car package in R to find the best model, and this will do so using likelihood ratio tests.

```
## Analysis of Deviance Table (Type III tests)
##           LR Chisq Df Pr(>Chisq)
## eye         1.8041  1    0.17921
## age         1.1008  1    0.29408
## type        0.2986  1    0.58477
## risk        6.5594  1    0.01043 *
## trt        17.7888  1  2.468e-05 ***
## laser:trt   0.4571  1    0.49896
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This model tests the covariates given all the other covariates are in our model to see if they are significant. We notice that only covariates that are significant are the risk factor and the treatment variable (treatment or control). Our interaction term for treatment and laser which we were interested in was not significant to our model. Another method to find the final model is by checking AIC for models with a different set of covariates.

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## Surv(futime, status) ~ eye + age + type + risk + laser:trt +
##     trt
##
## Final Model:
## Surv(futime, status) ~ risk + trt
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                 388   1703.123 1715.123
## 2      - type  1 0.2985857       389   1703.422 1713.422
## 3 - trt:laser  1 0.4530961       390   1703.875 1711.875
## 4       - age  1 1.1041625       391   1704.979 1710.979
## 5       - eye  1 1.6654657       392   1706.644 1710.644
```

The stepAIC uses multiple models to find AIC and compares them together to find the best model with the lowest AIC. Using stepAIC on both directions we get that the best model is the model with only risk factor and treatment. This is the same result that we found using the Anova function which calculates LRT tests. So our final model is the model with risk factor and treatment.

```
lrt <- 2*(retino.full$loglik[2] - retino.control$loglik[2])
lrt
```

```
## [1] 28.43032
```

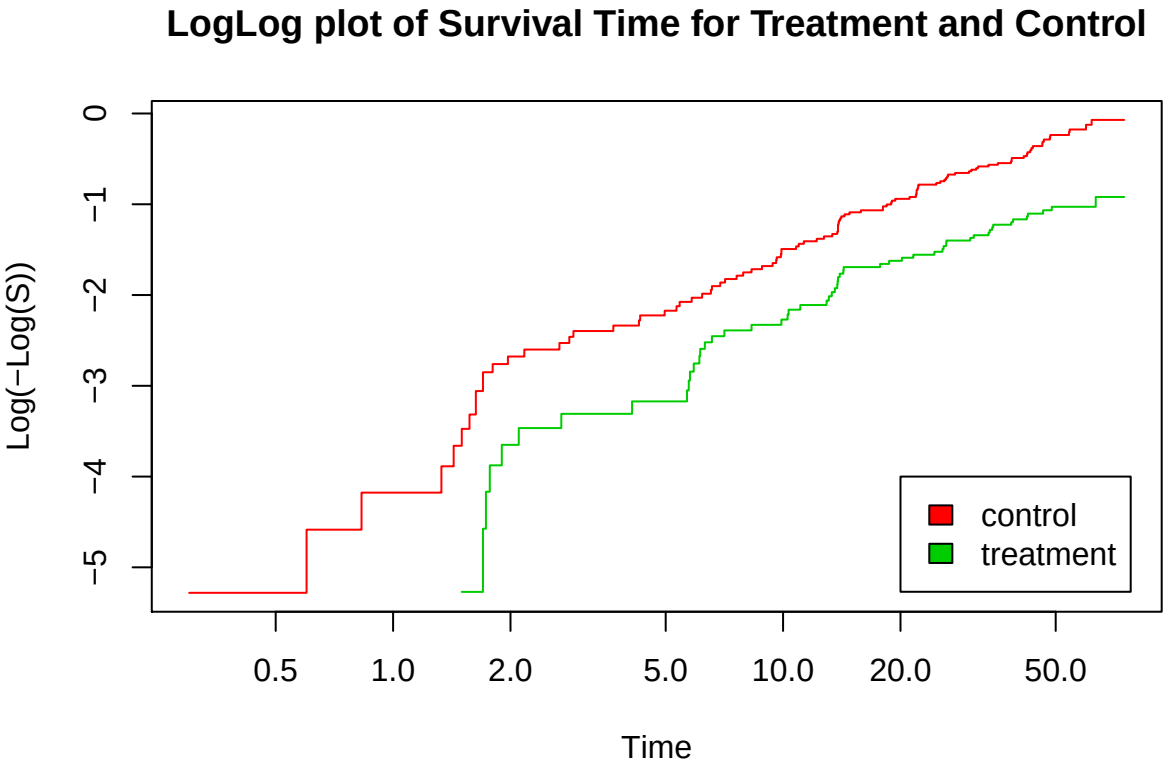
```
pchisq(lrt, df = 2, lower.tail = F)
```

```
## [1] 6.705539e-07
```

We obtain a p-value of 6.705e-07 which is less than 0.05, so we reject the null hypothesis. We conclude that when we take into account for the confounding effects of the other covariates in our dataset, the model with risk and trt is significant.

Model Checking

We want to check to see if our final model follows the proportional hazards assumption using loglog plots and residual tests.



When looking at the loglog plot between Treatment and Control, we see that the two lines are parallel so our proportional hazards assumption is probably true. Because our risk factor is numerical, we will not use a loglog plot to determine if proprotional hazards assumption is true. Another way to check for the proportional hazards assumption is to look at the coxzph.

```
cox.zph(retino.final)
```

```
##            rho chisq     p
## risk   -0.1068 1.667 0.197
## trt    -0.0592 0.546 0.460
## GLOBAL      NA 2.221 0.329
```

When we run our coxzph test, the p-value for risk of 0.197 which is greater than 0.05 suggests that we fail to reject the null hypothesis that the proportional hazards assumption is true. The p-value for trt of 0.460 is greater than 0.05 suggests that we fail to reject the null hypothesis as well. Thus, our proportional hazard assumption is true for risk and trt. This shows us that our final model satisfies the proportional hazards assumption so we are justified to use our model.

Model Explanation/Conclusion

```
summary(retino.final)
```

```
## Call:
## coxph(formula = Surv(futime, status) ~ risk + trt, data = retinopathy)
##
##   n= 394, number of events= 155
##
##          coef exp(coef) se(coef)      z Pr(>|z|)
## risk  0.14600   1.15720  0.05588  2.613  0.00898 **
## trt  -0.77792   0.45936  0.16881 -4.608 4.06e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      exp(coef) exp(-coef) lower .95 upper .95
## risk    1.1572     0.8642     1.037    1.2911
## trt     0.4594     2.1769     0.330    0.6395
##
## Concordance= 0.626  (se = 0.024 )
## Rsquare= 0.072   (max possible= 0.988 )
## Likelihood ratio test= 29.33  on 2 df,   p=4e-07
## Wald test            = 27.97  on 2 df,   p=8e-07
## Score (logrank) test = 29.01  on 2 df,   p=5e-07
```

Using the likelihood ratio test, we obtain a p-value of 4e-07 which is less than 0.05, so we reject the null hypothesis. There is a signifiant difference between hazard rates for risk factor and treatment and control, and it shows our model is significant. In addition to what we saw for the treatment versus control from the earlier part, we also obtain that for every increase in 1 risk factor, the hazard ratio goes up by 1.158 times, which means that as risk factor increases, hazard ratio also increases, causing lower survival times. Looking at our confidence intervals, we are 95% confident that the true hazard ratio between treatments and control is between 0.330 and 0.6395. Because the upper bound is also less than 1, we can tell that it is significant. Additionally, we are 95% confident that the true hazard ratio for a unit increase in risk is between 1.037 and 1.2911. Because the lower bound is also greater than 1, we can tell that this covariate is significant.

Through our analysis, we found that there is a significant difference between the vision loss survival between control and treated eyes when we treat control and treated as independent. In our final model, we were mainly interested in seeing if the interaction between type of laser and treatment (we are using interaction because laser is not used on control eyes), but this interaction was not significant, so the type of laser being used did not affect the survival rates significantly. In conclusion, we showed that using laser treatment helps with reducing vision loss but the type of laser being used does not matter. We also found that the risk factor which is a measure of the severity of the diabetic retinopathy affected survival rates of the patients. Overall we conclude that the factors that contribute the most to vision loss is whether the person was treated with the treatment and whether their risk of diabetic retinopathy was high.

Extension

In this extension, we are not going to disregard the dependence between the treatment and control variable, as it is the same person with one of their eyes treated and the other not. Because of this dependence, we are going to observe a frailty model, adding to the final model that we found from model fitting in the part before. In our frailty model, we are going to use the frailty function on the id of the patients to account for

the individual level differences. [3]

```
## Call:
## coxph(formula = Surv(futime, status) ~ risk + trt + frailty(id,
##     distribution = "gamma"), data = retinopathy)
##
##   n= 394, number of events= 155
##
##                            coef    se(coef) se2     Chisq  DF    p
## risk                        0.1669 0.06909  0.05929   5.84  1.00 1.6e-02
## trt                        -0.9050 0.17466  0.17125  26.85  1.00 2.2e-07
## frailty(id, distribution                             107.44 80.12 2.3e-02
##
##      exp(coef) exp(-coef) lower .95 upper .95
## risk    1.1817     0.8463    1.0320    1.3530
## trt     0.4045     2.4720    0.2873    0.5697
##
## Iterations: 6 outer, 30 Newton-Raphson
##      Variance of random effect= 0.7934231   I-likelihood = -847.9
## Degrees of freedom for terms=  0.7  1.0 80.1
## Concordance= 0.842  (se = 0.024 )
## Likelihood ratio test= 199  on 81.81 df,   p=9e-12
```

```
## Call:
## coxph(formula = Surv(futime, status) ~ laser:trt + age + risk +
##     trt + frailty(id, distribution = "gamma"), data = retinopathy)
##
##   n= 394, number of events= 155
##
##                            coef      se(coef) se2      Chisq  DF    p
## age                         0.006255 0.007313 0.005611   0.73  1.00 3.9e-01
## risk                        0.168861 0.068963 0.059280   6.00  1.00 1.4e-02
## trt                        -0.983397 0.234855 0.227697  17.53  1.00 2.8e-05
## frailty(id, distribution                                104.95 78.28 2.4e-02
## laserargon:trt              0.143658 0.299635 0.281256   0.23  1.00 6.3e-01
##
##                exp(coef) exp(-coef) lower .95 upper .95
## age                1.006     0.9938    0.9920    1.0208
## risk               1.184     0.8446    1.0343    1.3553
## trt                0.374     2.6735    0.2361    0.5927
## laserargon:trt     1.154     0.8662    0.6417    2.0770
##
## Iterations: 6 outer, 31 Newton-Raphson
##      Variance of random effect= 0.7748154   I-likelihood = -847.4
## Degrees of freedom for terms=  0.6  0.7  0.9 78.3  0.9
## Concordance= 0.839  (se = 0.024 )
## Likelihood ratio test= 197.3  on 81.43 df,   p=1e-11
```

The p-value for the frailty component is 0.023 which is less than 0.05, indicating that the frailty component is significant to our model. Because our frailty component is significant, we conclude that the variance of the random component is non-zero for this model. We get that the variance of our random effect is 0.793 and this is significant according to our test. We notice that the hazard ratio we received in our final model without frailty was 1.1572 and the model with frailty has a hazard ratio of $\exp(0.1669) = 1.1816$. So with our frailty model, we observe that our hazard ratio is slighly larger than that without the frailty componenet.

[3] D. Kleinbaum, M. Klein (2012) Survival Analysis: A Self-Learning Text 3rd ed.

Additionally, we notice that the hazard ratio we received in our final model without frailty was 0.4594 and the model with frailty has a hazard ratio of exp(-0.9050) = 0.4045. With our frailty model, we observe that our hazard ratio is slighly lower than that without the frailty component. We notice that with both covariates, our hazard ratios for the covariates in our frailty model deviate more from 1 then the hazard ratios from the model without frailty. This shows that our frailty model, after accounting for the individual level differences, emphasizes the differences in all risk levels and between treatment and control. We are 95% confident that the true hazard ratio for risk is between 1.032 and 1.353. We are 95% confident that the the true hazard ratio for treatment is between 0.2873 and 0.5697. For the risk factor, we found that the confidence interval for the frailty model is wider than the one without the frailty component. For the treatment, we found that the confidencew interval for the model without the frailty componenet was wider than the one with the frailty component.

We also fit a model that included the covartiates, age, risk, trt, laser:trt, and the frailty component on id. We find that the 95% confidence interval for age is between 0.992 and 1.021 and this interval includes 1 so age is not signficant. We find that the 95% confidence interval for interaction between laser and treatment is between 0.6417 and 2.077 and this includes 1 so interaction between laser and treatment is not significant either.

We can also use a cluster method instead of frailty because cluster does not assume a gamma distribution.

```
cox.cluster <- coxph(Surv(futime,status) ~ risk + trt + cluster(id), data=retinopathy)
summary(cox.cluster)
```

```
## Call:
## coxph(formula = Surv(futime, status) ~ risk + trt + cluster(id),
##     data = retinopathy)
##
##   n= 394, number of events= 155
##
##          coef exp(coef) se(coef) robust se      z Pr(>|z|)
## risk  0.14600   1.15720  0.05588   0.05913  2.469   0.0135 *
## trt  -0.77792   0.45936  0.16881   0.14975 -5.195 2.05e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      exp(coef) exp(-coef) lower .95 upper .95
## risk    1.1572     0.8642    1.0306    1.2994
## trt     0.4594     2.1769    0.3425    0.6161
##
## Concordance= 0.626  (se = 0.024 )
## Rsquare= 0.072   (max possible= 0.988 )
## Likelihood ratio test= 29.33  on 2 df,   p=4e-07
## Wald test            = 29.8  on 2 df,   p=3e-07
## Score (logrank) test = 29.01  on 2 df,   p=5e-07,   Robust = 28.92  p=5e-07
##
##   (Note: the likelihood ratio and score tests assume independence of
##      observations within a cluster, the Wald and robust score tests do not).
```

When we use a cluster method as opposed to the frailty method, we get that the 95% confidence interval is between 1.0306 and 1.2994. We find that the confidence interval for our cluster method is narrower than that of our frailty method. We find that the 95% confidence interval is between 0.3425 and 0.6161. We find that the confidence interval for our cluster method is narrower than that of our frailty model. In general, our confidence intervals for our cluster method are narrower than that of the frailty method.

We should also check whether our new model with the frailty component satisfies the proportional hazards assumption. We will use a coxzph test to test this.

```
cox.zph(cox.frail)
```

```
##            rho chisq     p
## risk   -0.0881  1.72 0.190
## trt    -0.0867  1.26 0.261
## GLOBAL      NA  3.05 0.218
```

When we run our coxzph test, the p-value for risk of 0.190 which is greater than 0.05 suggests that we fail to reject the null hypothesis that the proportional hazards assumption is true. The p-value for trt of 0.261 is greater than 0.05 suggests that we fail to reject the null hypothesis as well. Thus, our proportional hazard assumption is true for risk and trt. This shows us that our frailty model satisfies the proportional hazards assumption so we are justified to use this as our model.

References

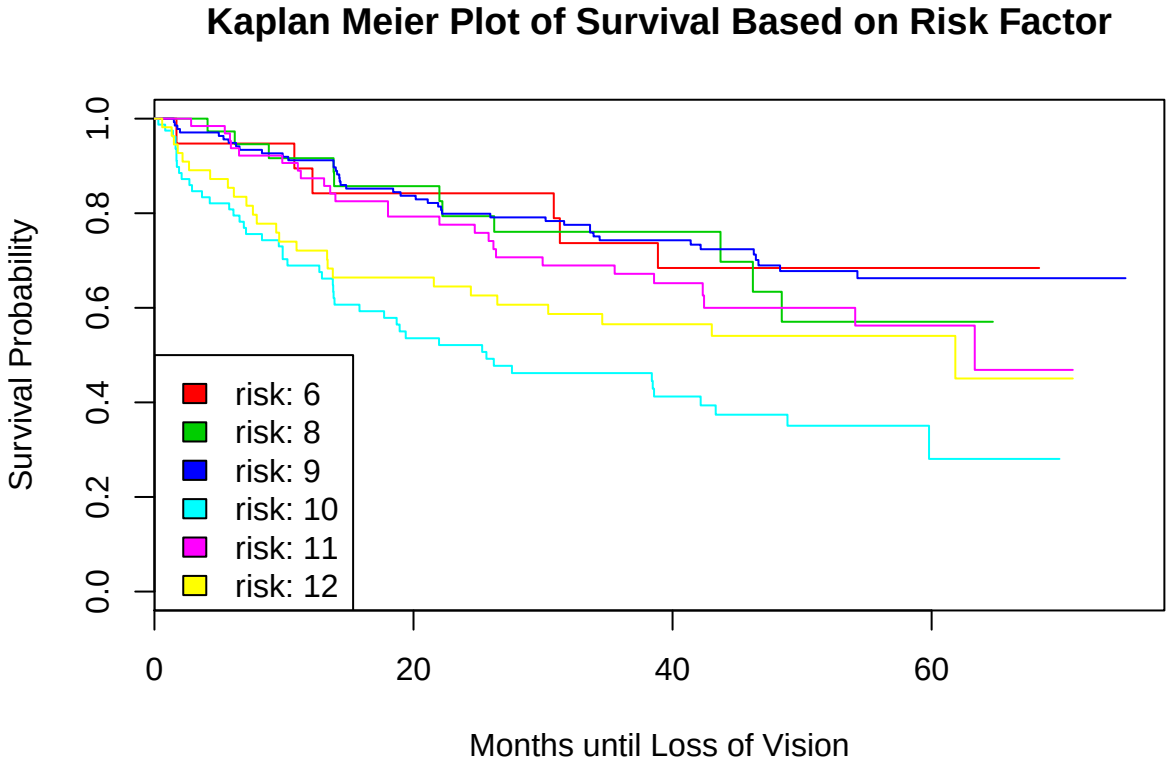
A. L. Blair, D. R. Hadden, J. A. Weaver, D. B. Archer, P. B. Johnston and C. J. Maguire (1976). The 5-year prognosis for vision in diabetes, American Journal of Ophthalmology, 81:383-396.

D. Kleinbaum, M. Klein (2012) Survival Analysis: A Self-Learning Text 3rd ed.

W. J. Huster, R. Brookmeyer and S. G. Self (1989). Modelling paired survival data with covariates, Biometrics 45:145-156.

Appendix

Because we already visualized treatment and control, we decide to make a Kaplan Meier plot to visualize the risk factor as well.



We can see that in general, higher risk leads to lower survival times. There seem to be some abnormalities which may have been the result of those eyes being the control rather than the treatment. It can also be contributed by the lack of observations for some of the risk factors such as for risk 6 and 8.

R Code

```
#SETUP:
library(survival)
library(car)
library(MASS)

#MODEL SETUP
summary(retinopathy)

#BASIC ESTIMATION
ret.trt.control.survfit <- survfit(Surv(futime, status) ~ trt, data = retinopathy)
plot(ret.trt.control.survfit, main = "Kaplan Meier Plot of Survival of Laser Treatment Patients
     for Control and Treated",
     xlab = "Months until Loss of Vision", ylab = "Survival Probability", col = c(2,3),
     mark.time = TRUE, mark = 18)
legend(3,0.25, c("Control Eye","Treated Eye"), fill = c(2,3))

treat.cox <- coxph(Surv(futime, status) ~ trt, data = retinopathy)
treat.cox

#MODEL FITTING
retino.full <- coxph(Surv(futime, status) ~  eye + age + type + risk + laser:trt + trt,
                     data = retinopathy)#full model of all covariates we want to test
Anova(retino.full, type = "III", test.statistic = "LR") # model fitting using likelihood
#ratio tests
step <- stepAIC(retino.full, direction = "both") #stepwise anova, another model fitting technique
step$anova

retino.control <- coxph(Surv(futime, status) ~ eye + age + type + laser, data = retinopathy)#model
#that we compare to full model to see if risk and trt are significant
retino.final <- coxph(Surv(futime,status) ~ risk + trt, data = retinopathy) #final model that
#we obtain
lrt <- 2*(retino.full$loglik[2] - retino.control$loglik[2])
lrt
pchisq(lrt, df = 2, lower.tail = F)

#MODEL CHECKING
retino_loglog <- survfit(Surv(futime, status) ~ trt, data = retinopathy)
plot(retino_loglog, fun = "cloglog", col = c(2,3),
     xlab = "Time", ylab = "Log(-Log(S))",
     main = "LogLog plot of Survival Time for Treatment and Control") #loglog plot for trt
legend(20,-4, c("control","treatment"), fill = c(2,3))

cox.zph(retino.final)

#MODEL EXPLANATION/SUMMARY
summary(retino.final)
```

```r
#EXTENSION
#fit a coxph to a frailty model
cox.frail <- coxph(Surv(futime,status) ~ risk + trt + frailty(id, distribution="gamma"),
                   data=retinopathy)
cox.frail2 <- coxph(Surv(futime,status) ~ laser:trt + age + risk + trt +
                      frailty(id, distribution="gamma"), data=retinopathy)
summary(cox.frail)
summary(cox.frail2)

cox.cluster <- coxph(Surv(futime,status) ~ risk + trt + cluster(id), data=retinopathy)
summary(cox.cluster)

cox.zph(cox.frail) #model checking for frailty model


#APPENDIX
riskfit <- survfit(Surv(futime,status)~risk, data = retinopathy)
plot(riskfit, col = 2:7, main = "Kaplan Meier Plot of Survival Based on Risk Factor",
     xlab = "Months until Loss of Vision", ylab = "Survival Probability")
legend(0,0.5, c("risk: 6","risk: 8", "risk: 9", "risk: 10", "risk: 11", "risk: 12"), fill = 2:7)
```