175 Project Final: Survival Analysis of Diabetic Retinopathy

Drew Pritchard Issam Zejli Janelle Samansky

12/5/2018

Background

Diabetic Retinopathy is a disease that affects some people with diabetes. It occurs when high blood sugar damages the blood vessels in the retina. These blood vessels can swell and leak. Or they can close, stopping blood from passing through. All of these changes can cause loss of vision.

Description

This dataset was a 50 percent random sample of patients with "high-risk" diabetic retinopathy. The data contains 197 patients, and each one had two types of a laser treatments on a random eye (left eye or right eye). The event of interest was the time(in months) from initiation of treatment to the time when visual acuity dropped below 5/200. Censoring was caused by death, dropout, or end of the study.

Scientific Questions of Interest

- Does the treatment of an eye effect the chance of survival? If so how are they different and to what degree?
- Is treatment at young age more effective than at an older age?
- Is there a difference in treatments between left eye and right eye?

Complication

Each observation is actually one of two parts. We can see that there are two of each patient id. For each patient only one eye was treated(trt == 1). Essentially we are given a control for each patient. Generally a dataset will only include one observation per subject. This allows us to look for a discrepancy amongst treated eye and untreated eye survival times. This is especially convenient because there are no other variables to interfere with our inference as each patient has one treated and one untreated. This does mean however that there is room for errors when performing our analysis and we have to be careful about the fact that our original dataset has a duplicate for each patient.

Table 1: Head of Retinopathy

id	laser	eye	age	type	trt	futime	status	risk
5	argon	left	28	adult	1	46.23	0	9
5	argon	left	28	adult	0	46.23	0	9
14	argon	right	12	juvenile	1	42.50	0	8
14	argon	right	12	juvenile	0	31.30	1	6
16	xenon	right	9	juvenile	1	42.27	0	11
16	xenon	right	9	juvenile	0	42.27	0	11

Our Dataset

Retinopathy is a survival data for trials of laser treatment on patients with diabetic retinopathy. A snapshot of the first few observations looks like:

An explanation of what the nine variables are is explained in the following table.

| Name	Description	Values
id	Numeric subject id. Two observations each.	(5 - 1749)
laser	Type of laser used.	xenon or argon
eye	Which eye was treated.	right or left
age	Age at diagnosis of diabetes.	1 - 58
type	Type of diabetes. Depending on the age of the subject.	juvenile or adult
trt	Indicator for whether an eye is treated(1) or controled(0).	0 or 1
futime	Time ot loss of vision or last follow-up	0.30 - 74.97
status	Indicator for whether subject lost their vision in that particular eye(1) or was censored from the trial(0)	0 or 1
risk	A risk score for the eye.	6 - 12

General Analysis

KM estimate

We use the Kaplan Meier curve to estimate the survival probability of Diabetic Retinopathy patients. As shown in the graph, the survival probability decreases with each passing month.

Table 2: Table of Quantiles

	Time (Months)	Lower Bound	Upper Bound
10	6.20	5.33	9.60
20	13.83	11.30	18.93
30	26.20	21.57	38.40
40	46.27	38.47	63.33
50	NA	59.80	NA

Kaplan–Meier Curve of Retinopathy Data
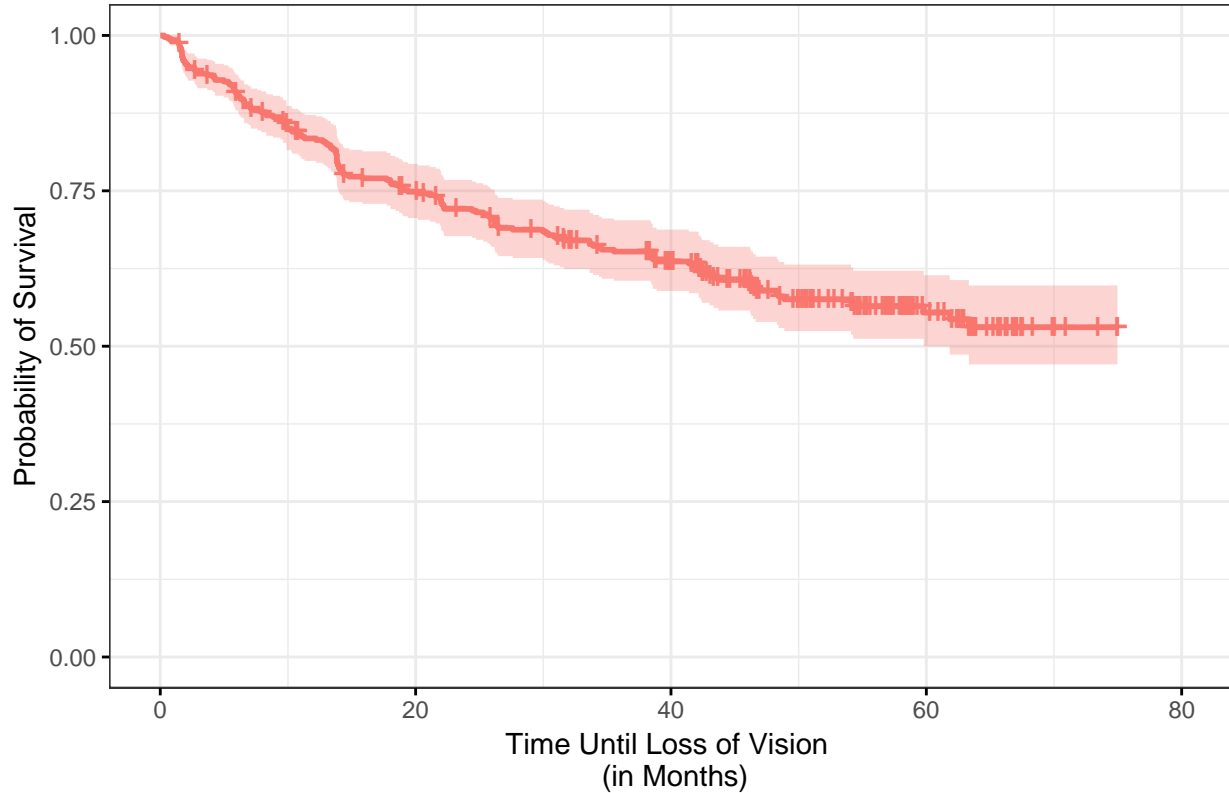
Quantiles of Survival Times

Analyzing the quantile survival(Table Above): the 10th quantile for the survival function is the time at which the survival probability is below the probability of 0.9 which is at 6.2 months. Similarly for the 20th quantile it is equal to 13.38 months.

In our data, no patient events occurred after 61 months and so we just assume that that the survival probability after that time is the same as the most recent event. In this case probability remains approximately 0.554. This can cause issues in inferences and analysis. Generally a statistician would want survial data where the probability of survival converges to zero. We will do our best to avoid any problems that may occur.

Kaplan-Meier Analysis

We continue our analysis of the KM estimate by looking at the effect of specific categorical variables in our data. In particular, we want to determine whether they are significant enough to include in our final model.
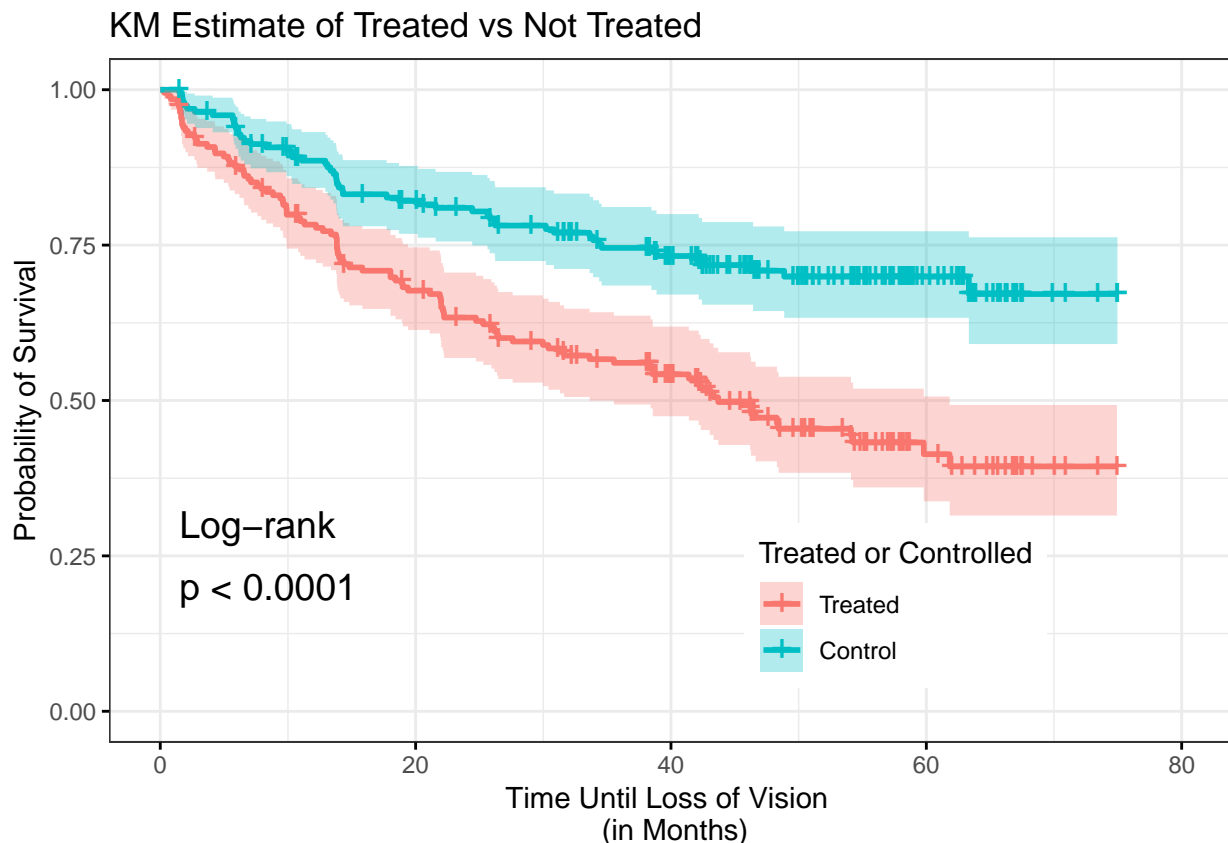
The test statistic we use is for the log-rank test and is a large chi-square test that uses as its criterion a

statistic that provides an overall comparison of the KM curves. The test compares the expected number of events with the actual number of events. The test won't be rejected as long as the expected number of events is the same as the actual number of events.
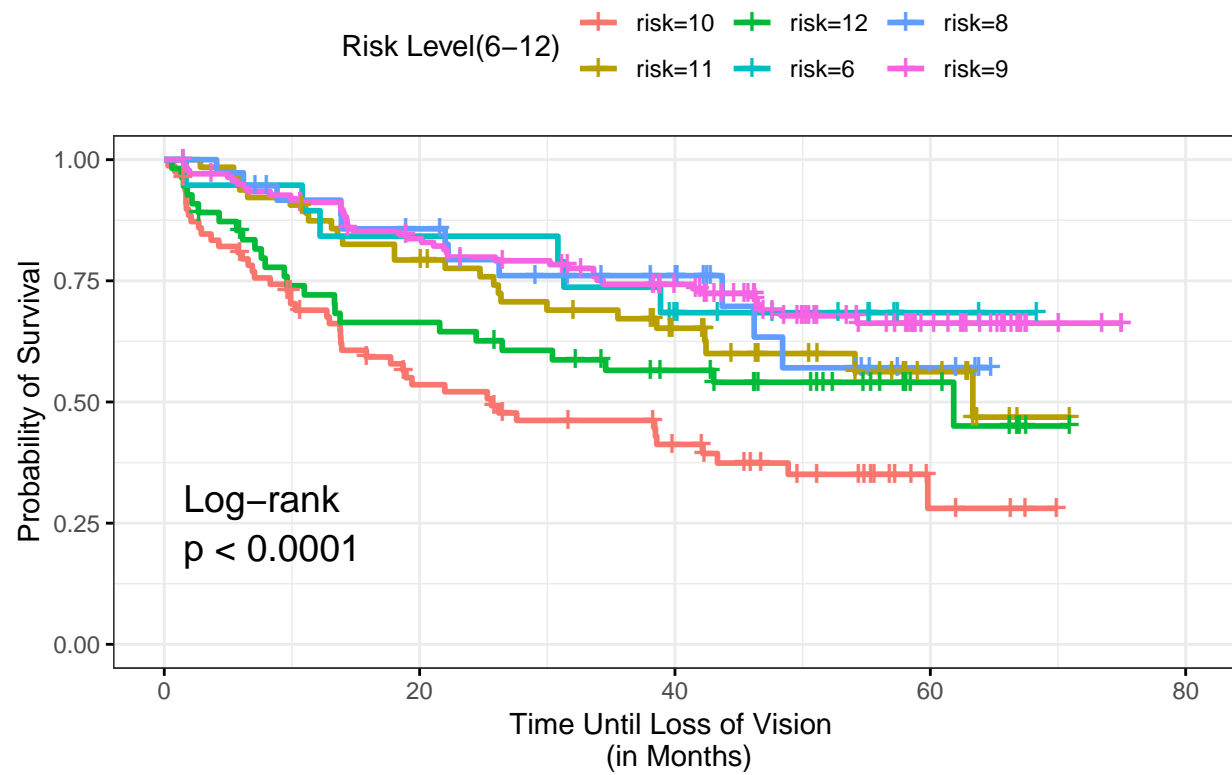
Our null hypothesis is as such: $H_0 : S_1(t) = S_2(t)$ (or in terms of hazard ratios: $H_0 : h_1(t) = h_2(t)$) In other words, we are trying ot find evidence for or against the assumption that survival function is the same for different values of a covariate.

- The p-value of the log-rank test for *laser* is 0.32 which is greater than our $\alpha = 0.05$, so we failt to reject the null hypothesis. Both the survival function where *laser* = Argon and *laser* = Xenon are identical. **No signifcant difference** is found.

- The p-value of the log-rank test for *age* is 0.89 which is also greater than α 0.05. Again, we fail to reject the null hypothesis → Survival functions of *age* are the same → **No significant difference.**

- The p-value of the log-rank test for *eye* is 0.09 which is greater than $\alpha = 0.05$. Again, we fail to reject the null hypothesis → Survival functions of *age* are the same → **No significant difference.**

- The p-value of the log-rank test for *trt*, *risk*, and *type*(separately) are less than $\alpha = 0.05$. We **reject** the null hypothesis → Survival functions including these covariates are not the same → **Significant Difference.**

We graph a visual representation such as the graph below using the ggsurvplot function in the survminer package. In the first graph we see the difference between treated and not-treated eyes. Similarly in the second graph we see the difference between all the different values of *risk*.

KM Estimate of Treated vs Not Treated

KM Estimate of All Risk Levels

Risk Level(6–12)

risk=10   risk=12   risk=8

risk=11   risk=6   risk=9

Log–rank
p < 0.0001

Probability of Survival

Time Until Loss of Vision
(in Months)

6

Cox PH Analysis - Univariate

Cox-proportional hazards modeling can be used for single variables in order to estimate their relative effect on hazard rate. For instance, if we fit a cox model on laser, the model shows that the hazard proportion(exp(coef) in the coxph() function) is 1.107 which indicates that the hazard rate for $laser =$ Argon is larger relative to Xenon with 10.7% increase in hazard rate(1.107 vs 1.0).

1. *laser*

- We conclude that with p-value of .3 (greater than .05) we say there is not a significant difference between the hazard rates for the patients that were treated with the two lasers.

2. *type*

- The p-value for our test is .9 which is greater than .05 so we conclude that there is no significant difference between the hazard rates of the patients with juvenile diabetes than those with adult diabetes.

3. *eye*

- The p-value for this test is .09 which is slightly above .05 therefore we say there is not much significant difference between the hazard rate of patients that receive treatment on their right eye than those that receive treatment on their left eye.

4. *risk*

- With a p-value less than .05 we conclude that there is a significant difference between the level of risk and the effect on patients with diabetes.

5. *trt*

- From running a coxph test we see that the p value is 2e^-6 which is considerably less than 0.05 therefore we conclude that there is a significant difference between the patients that were in the treatment group than those that were in control.

Variable	exp(coef)	CI Interval of exp(coef)	p-value	LRT Test Statistic
Laser	1.175 (argon)	[0.857, 1.610]	0.3	1.0
Type	1.022 (adult)	[0.744, 1.404]	0.9	0.02
Eye	1.316 (left)	[0.952, 1.818]	0.09	2.89
Treatment	0.46	[0.3304, 0.6403]	2e-06	22.35
Age	1.003	[0.9922, 1.014]	0.6	0.27
Risk8	1.0851	[0.4012, 2.935]		
Risk9	0.9614	[0.4081, 2.265]		
Risk10	2.8826	[1.2313, 6.748]	2e-06	28.95
Risk11	1.3758	[0.5654, 3.361]		
Risk12	1.8351	[0.7526, 4.474]		

What We Infer

The covariate *trt* has the highest likelihood of 22.35 and p-value of 2e-06 which makes it significant and most likely should include in our model. The covariate *risk* also has a high likelihood of 28.95 and p-value of 2e-05 which makes it significant and must include in our model.

Covariate Selection

By performing the backward selection and comparing the log likelihood of the covariates we see that *risk* and *trt* are the best covariates for our survival model because they have the highest log-likelihood values and the smallest p-values. Below shows an example of the anova table results that we analyze for each model.(NA values refer to the NULL model with no covariates) Some of the models we used were:

fitb = coxph(ret.surv ~ laser + eye + age + type + risk + trt, data = ret)

fitc = coxph(ret.surv ~ eye + risk + trt, data = ret)

fitd = coxph(ret.surv ~ risk + trt, data = ret)

	loglik	Chisq	Df	Pr(>|Chi|)
NULL	-867.9858	NA	NA	NA
laser	-867.4834	1.0049019	1	0.3161273
eye	-866.2147	2.5372730	1	0.1111867
age	-865.9147	0.6000454	1	0.4385607
type	-865.7766	0.2761087	1	0.5992631
risk	-851.4023	28.7485982	5	0.0000260
trt	-838.4142	25.9762370	1	0.0000003

Note:
anova(fitb) results

	loglik	Chisq	Df	Pr(>|Chi|)
NULL	-867.9858	NA	NA	NA
eye	-866.5757	2.82022	1	0.0930837
risk	-852.0449	29.06155	5	0.0000226
trt	-839.4686	25.15256	1	0.0000005

Note:

anova(fitc) results

	loglik	Chisq	Df	Pr(>|Chi|)
NULL	-867.9858	NA	NA	NA
risk	-853.5087	28.95425	5	2.37e-05
trt	-840.9825	25.05236	1	6.00e-07

Note:

anova(fitd) results

We reach the same conclusion that only *risk* and *trt* are significant.

Cox PH Analysis - Multivariate

Testing all the covariates together, trt and *risk10* are the most significant. As shown below in the raw R output(All numbers refer to the p-value of our test).

Something we are weary about is the fact that only *risk10* is significant and the other 5 levels aren't. This is abnormal especially considering that *risk10* is right in the middle of the *risk* scale(6 - 12). We will keep this in mind throughout the analysis.

	p-value
laserargon	0.2376562
eyeleft	0.0934458
trt	0.0000007
age	0.5367502
typeadult	0.8644819
risk8	0.9206809
risk9	0.7317151
risk10	0.0189509
risk11	0.6618259
risk12	0.2828816

From our previous univariate analysis and covariate selection, we chose to test the covariates:
trt and *risk.*
Again, *trt* and *risk10* remains the most significant as shown below.

	p-value
trt	0.0000011
risk8	0.9236188
risk9	0.7972208
risk10	0.0161935
risk11	0.5627627
risk12	0.2420628

Variable	PH Assumption Check
Laser	Satisfied
Type	Satisfied
Eye	Satisfied
Treatment	Satisfied
Age	Satisfied
Risk8 Risk9 Risk10 Risk11 Risk12	Satisfied

Model Checking

First we simply run a series of coxzph functions to test our Cox PH assumptions:

```
##              rho chisq     p
## laserargon -0.0764 0.898 0.343
```

```
##            rho chisq     p
## typeadult 0.0534 0.443 0.506
```

```
##          rho chisq     p
## eyeleft 0.0848  1.11 0.293
```

```
##            rho  chisq     p
## risk8   0.05107 0.4037 0.525
## risk9   0.00909 0.0128 0.910
## risk10 -0.03585 0.1975 0.657
## risk11  0.03880 0.2329 0.629
## risk12 -0.07470 0.8578 0.354
## GLOBAL       NA 8.1475 0.148
```

```
##        rho chisq     p
## trt -0.0564 0.495 0.482
```

None of our variables result in a p-value below .05 so we believe our assumptions are met. Next we use a complementary log-log plot to visualize whether the Cox PH assumption is appropriate for modeling the effect of our factorized covariates.

Log–Log of Treatment variable
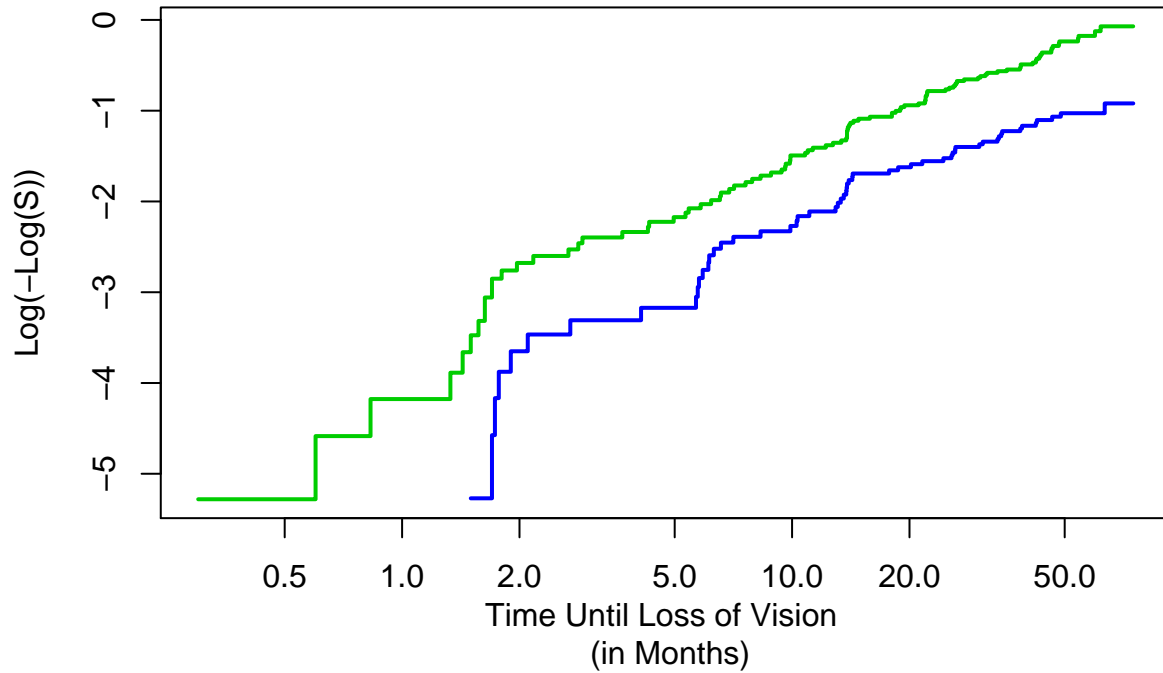
We see from the Log-log plot for treatment the two curves are parallel so we can say that the assumptions hold, which agrees with our conclusion after performing the coxzph test.

Log–Log of Risk variable
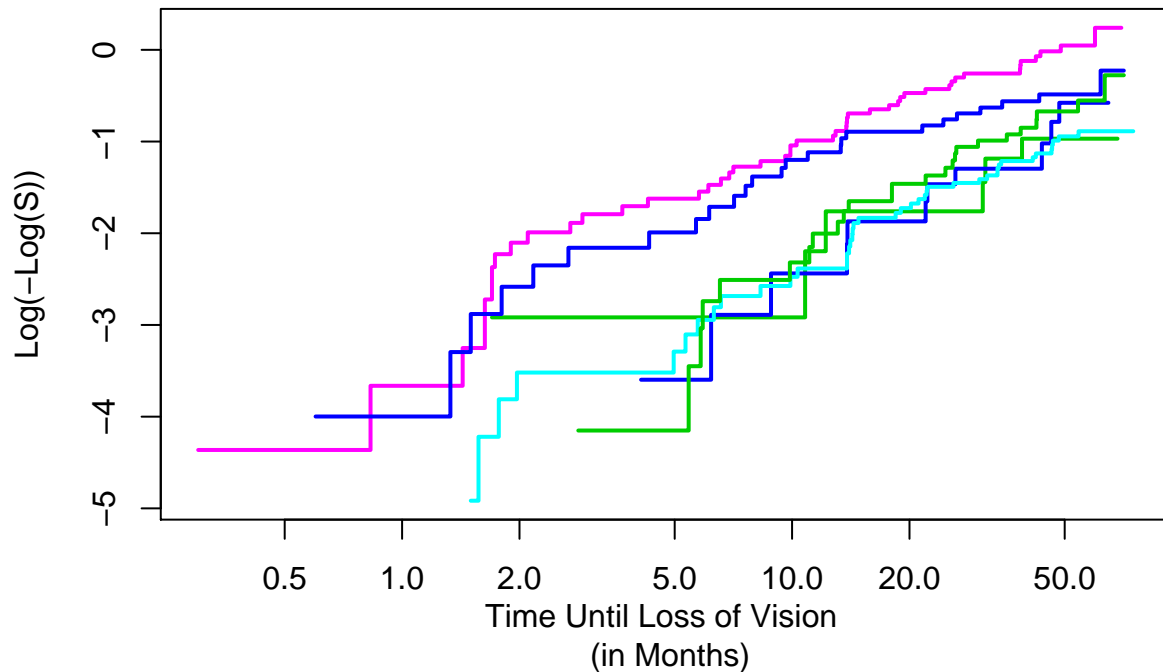
This log-log plot of our risk variable seems to contradict our previous coxzph analysis.. Although it could be because we have a total of six levels, we still decide it would be best to look at stratifying the risk variable.

Stratification

We cannot leave *risk* in our model as it is. While we do believe there is some significance in it, it does not hold up to our PH assumptions and is overall causing issues. To try and fix this, we stratify *risk* and compare it to our model without *risk* at all.

After stratifying on the risk variable using the strata() function we take a look at the difference with and without it.

```
## Call:
## coxph(formula = ret.surv ~ ret$trt + strata(ret$risk), data = ret)
##
##           coef exp(coef) se(coef)    z       p
## ret$trt -0.800     0.449    0.170 -4.7 2.6e-06
##
## Likelihood ratio test=23.33  on 1 df, p=1e-06
## n= 394, number of events= 155
```

```
## Call:
## coxph(formula = ret.surv ~ trt, data = ret)
##
##         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
```

We choose to perform an LRT test between the reduced and full model using the formula:
$-2 * log(L_R) - (-2 * log(L_F))$
From this we get:

```
## [1] 0
```

We find a chi-squared p-value of 0 and declare that the stratified term is significant.

Now we take a second to look over our current model

```
## Call:
## coxph(formula = ret.surv ~ ret$trt + strata(ret$risk), data = ret)
##
##           coef exp(coef) se(coef)    z       p
## ret$trt -0.800     0.449    0.170 -4.7 2.6e-06
##
## Likelihood ratio test=23.33  on 1 df, p=1e-06
## n= 394, number of events= 155
```

What we can infer here is that the hazard ratio(exp(coef)) for treated subjects is only 45% of those who are untreated. Meaning they keep their sight 55% longer.
Also, we can find the confidence interval by reading the summary.

We find it is (.322, .627). So we believe our range of reduced time to loss of eyesight is between 32.2% to 62.7%.

Paired observations are not independent(Extension)

We have paired observations. For every subject there are two observations: one treated eye and one control eye. Because there is almost certainly correlation amongst each subject's eyes(each person has a different

propensity of having their sight fail in either eye) we must look to see if using a **cluster** or **frailty** model provides more information than without them.

Essentially what we're doing is looking at the subject's two observations separately to reduce the effect of a subject's eye correlation has on our analysis.

First we look at our current model:

```
## Call:
## coxph(formula = ret.surv ~ trt, data = ret)
##
##        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
```

Things to note down: LRT = 23.37 on 1df, exp(coef) = 0.449, p = 2.6e-06.

Now let us take a look at the same model with an added cluster term:

```
## Call:
## coxph(formula = ret.surv ~ trt + cluster(id), data = ret)
##
##        coef exp(coef) se(coef) robust se     z       p
## trt -0.777     0.460    0.169     0.147 -5.27 1.4e-07
##
## Likelihood ratio test=22.37  on 1 df, p=2e-06
## n= 394, number of events= 155
```

We see that it's very similar to our previous model. Almost no change except for the p-value and they are both so small the change is trivial.

Lastly we look at the frailty model:

```
## Call:
## coxph(formula = ret.surv ~ trt + frailty(id, distribution = "gamma") +
##     strata(risk), data = ret)
##
##                           coef se(coef)   se2  Chisq   DF       p
## trt                     -0.888    0.176 0.173 25.359  1.0 4.8e-07
## frailty(id, distribution                      72.002 57.1   0.088
##
## Iterations: 7 outer, 42 Newton-Raphson
##      Variance of random effect= 0.508   I-likelihood = -591.5
## Degrees of freedom for terms=  1.0 57.1
## Likelihood ratio test=143  on 58 df, p=4e-09
## n= 394, number of events= 155
```

Very different values. LRT = 201 on 85.6 df, etc.

We run another LRT between the frailty model and our current stratafied model.

```
## [1] 0
```

We find the frailty term is significant!

When we look at the results after using the cluster function on id against all covariates, we see that again only treatment and risk prove to be significant covariates in our analysis. When we cluster *id* against *laser*, we get a p value a little bigger than .3 showing that even with modifying the data so as to account for the fact we have multiple observation for the same subject the type of *laser* has no significant effect on patients

alone. When we use cluster function against treatment, we again see that this covariate has significant effect on survival rate even with our adjustment to the data.

Our final model we arrive at is:

coxph(ret.surv ~ trt + strata(risk) + frailty(id))

If we run summary() on this model we can look at the confidence intervals of the hazard ratios.

```
## Call:
## coxph(formula = ret.surv ~ trt + frailty(id, distribution = "gamma") +
##     strata(risk), data = ret)
##
##   n= 394, number of events= 155
##
##                        coef   se(coef) se2   Chisq DF    p
## trt                  -0.8881 0.1764   0.1727 25.36 1.00 4.8e-07
## frailty(id, distribution                    72.00 57.07 8.8e-02
##
##     exp(coef) exp(-coef) lower .95 upper .95
## trt   0.4114    2.431     0.2912    0.5813
##
## Iterations: 7 outer, 42 Newton-Raphson
##      Variance of random effect= 0.5081123   I-likelihood = -591.5
## Degrees of freedom for terms=  1.0 57.1
## Concordance= 0.802  (se = 0.052 )
## Likelihood ratio test= 143.3  on 58.02 df,    p=4e-09
```

	exp(coef)	exp(-coef)	lower .95	upper .95
trt	0.4114204	2.430604	0.2911816	0.5813098

Conclusion

The main takeaway from our analysis is that there is a distinct difference amongst the treated vs nontreated groups. Patients' treated eyes survived 58.9% longer until failure compared to their non-treated counterparts. Using the table from above we can even construct a 95% confidence interval for our hazard ratios([0.2919, 0.5813]). So ultimately we believe the treatment to allow paitents keep their eye sight between 42% and 61% **longer** than those who forego the treatment.

Aside from that we don't find any other variable that contributed towards patients losing their eyesight or preventing sight failure(such as left or right eye, etc.).

References

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