Stats 504 Homework 2

October 10, 2021

1 Introduction

Diabetic Retinopathy is a retinal disorder in diabetes patients which can cause blindness. There exists two laser treatments, argon/xenon, that can help delay diabetic retinopathy. In this analysis, we plan to determine the efficacy and quantify the improvement of each laser treatment on visual acuity. Moreover, we want to explain the influence of age and clinical risk of diabetic retinopathy on visual acuity.

2 Method

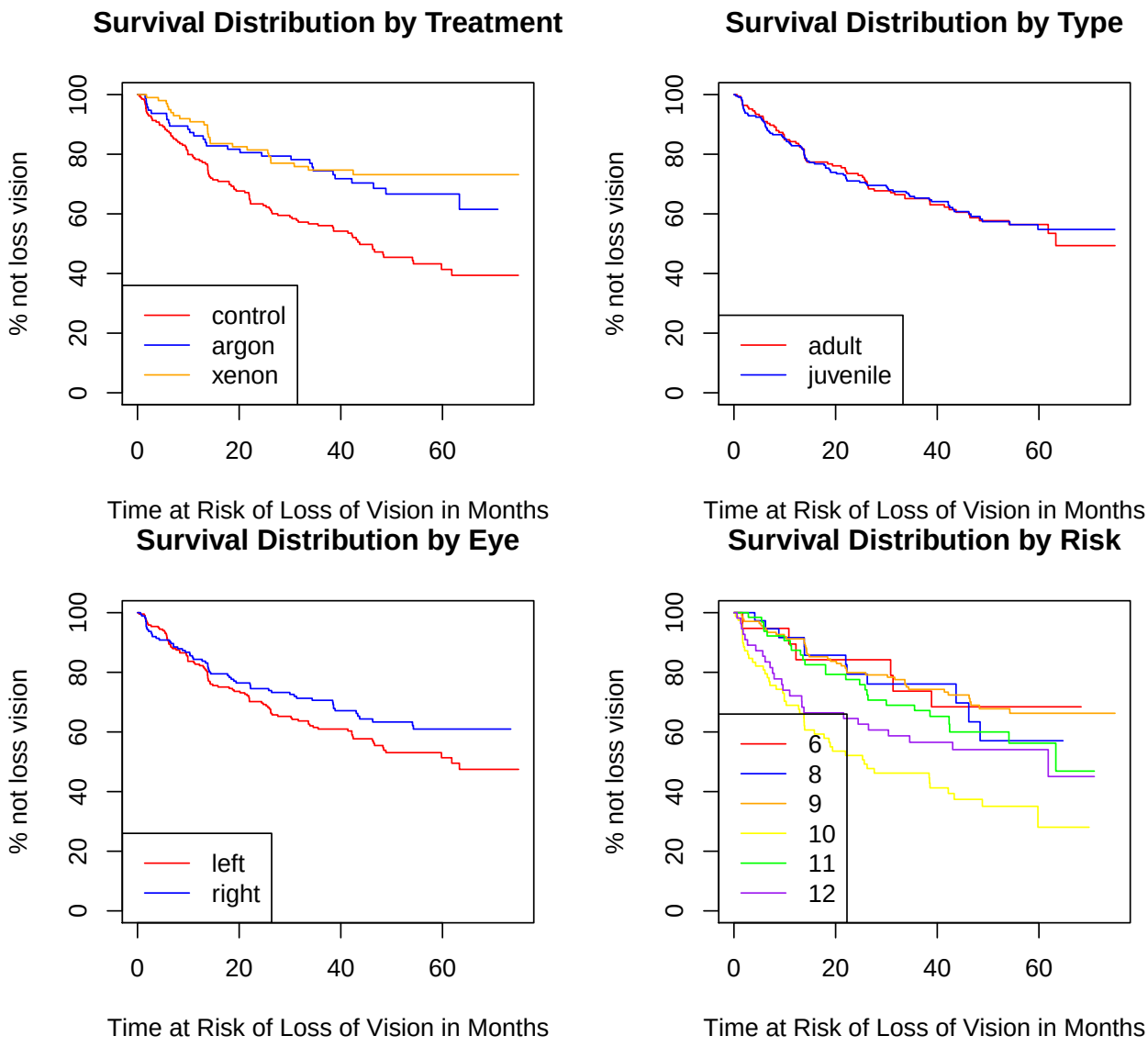
In this analysis, our target is to figure out the efficacy of two laser treatments and analyze the effect of age and clinical risk. We first use Kaplan-Meier Estimator to plot the survival curve curve and test if there is a difference between the survival curves. And then we use Cox Proportional Hazards Model to fit the data and quantify the relationship between the predictors and hazards functions. We also noticed that in this dataset, every two rows can be treated as a cluster, since they are just left and right eyes of one patient. There exists some associations between observations. In order to address this issue, we consider the frailty model, we just include a frailty term `frailty(id)` in our Cox PH model.

3 Result

The whole dataset contains 197 patients and 394 observations. It has totally 10 columns and does not have any missing value. We defined a new variable `survobj` as our response variable in this analysis, which is a survival object and combines the information of `futime` and `status`. We also created a new variable `treatment` which encodes the `laser` to 'control' when `trt` equals 0. In this dataset, `age` and `type` contain redundant information, we just use `type` in Kaplan-Meier Estimator and `age` in Cox Proportional Hazards Model. Since there is no casual relationship between other predictors in this dataset, we also include `eye` and `risk` in our model.

3.1 Kaplan-Meier Estimation

The following figure shows survival curve against `treatment`, `type`, `eye` and `risk`. Through the figure we can see that in the plot against `treatment` and `risk`, the difference between curves is relatively larger than the rest two plots. Moreover, the difference increases as time increases. As time goes by, the survival curve of control group and high clinical risk group drcreases faster than the others.



We also conduct a log-rank test to test whether there is a true difference between survival curves. The hypothesis of the log-rank test as follows.
H_0 : There is no difference in the Kaplan-Meier survival curve between different groups.
H_1 : There is a difference in the Kaplan-Meier survival curve between different groups.
The result is shown in the following table. We can see that for the difference between `treatment` group and `risk` group they are statistical significant, and we do not have enough statistical evidence to conclude the difference between `eye` group and `type` group.

Table 1: The Log-rank Test

Variable	Chisq2	P-value
Treatment	22.666	2e-06
Type	0.017	0.894859
Eye	2.788	0.094968
Risk	32.962	0

3.2 Cox Proportional Hazards Model

In this section, we first conduct the classical Cox Proportional Hazards Model using `survobj` as response and `eye`, `age`, `risk`, `treatment` as predictors. Then we conduct hypotheses to test the proportional hazards assumptions for the Cox PH model. The result indicates that the proportional assumption of Cox PH model is satisfied quite well here.

Table 2: The Cox Proportion Hazard Model

	coef	z	Pr(>|z|)	exp(coef)	exp(-coef)	lower .95	upper .95
eyeright	-0.2225	-1.3281	0.1841	0.8005	1.2492	0.5765	1.1116
age	0.0059	1.0673	0.2858	1.0059	0.9941	0.9951	1.0168
risk	0.1426	2.5341	0.0113	1.1533	0.8671	1.0328	1.2877
treatmentargon	-0.6918	-3.2730	0.0011	0.5007	1.9974	0.3308	0.7576
treatmentxenon	-0.8762	-3.9054	0.0001	0.4163	2.4019	0.2682	0.6463

Table 3: Proportional Hazards Assumption Test

	chisq	df	p
eye	0.8010425	1	0.3707819
age	0.3635736	1	0.5465282
risk	1.6921822	1	0.1933136
treatment	0.5946964	2	0.7427853
GLOBAL	3.7861712	5	0.5805942

To address the association between the left and right eyes of one person, we conducted a Frailty Cox PH model, which adds a simple random effects term to allow intra-person correlation in one patient.

Table 4: The Frailty Cox Proportional Hazards Model

	coef	Chisq	p	exp(coef)	exp(-coef)	lower .95	upper .95
eyeright	-0.3127	2.0069	0.1566	0.7315	1.3671	0.4746	1.1274
age	0.0075	1.0279	0.3107	1.0075	0.9926	0.9931	1.0221
risk	0.1642	5.6454	0.0175	1.1785	0.8486	1.0292	1.3494
treatmentargon	-0.8504	14.1398	2e-04	0.4273	2.3405	0.2743	0.6656
treatmentxenon	-0.9553	16.5165	0	0.3847	2.5995	0.2427	0.6098
frailty(id)	NA	104.0195	0.0235				

After Comparing these two models, we can see that the coefficients and significant tests are quite similar. `Argon` treatment, `Xenon` treatment and `Risk` are statistically significant in both models, and rest predictors are not. Here we use Frailty Cox PH model to quantify the efficacy of two laser treatments. The exponential coefficient of `Argon` treatment is 0.4273 (95% [0.2743, 0.6656]) and the exponential coefficient of `Xenon` treatment is 0.3847 (95% [0.2427, 0.6098]). Comparing with the control group, `Argon` treatment can decrease the risk of visual loss to 42.73% (95% [27.43%, 66.56%]) and `Xenon` treatment can decrease the risk of visual loss to 38.47% (95% [24.27%, 60.98%]). The `Xenon` treatment can provide relatively better treatment effect. AS for the effect of `age` and `risk`, since `age` is not statistically significant we can not conclude any effect of `age` on the risk of visual loss. The exponential coefficient of `risk` treatment is 1.1785 (95% [1.0292, 1.3494]), which indicates that 1-unit increase in clinical risk, the risk of visual loss will also increase 17.85% (95% CI [2.92%, 34.94%]).

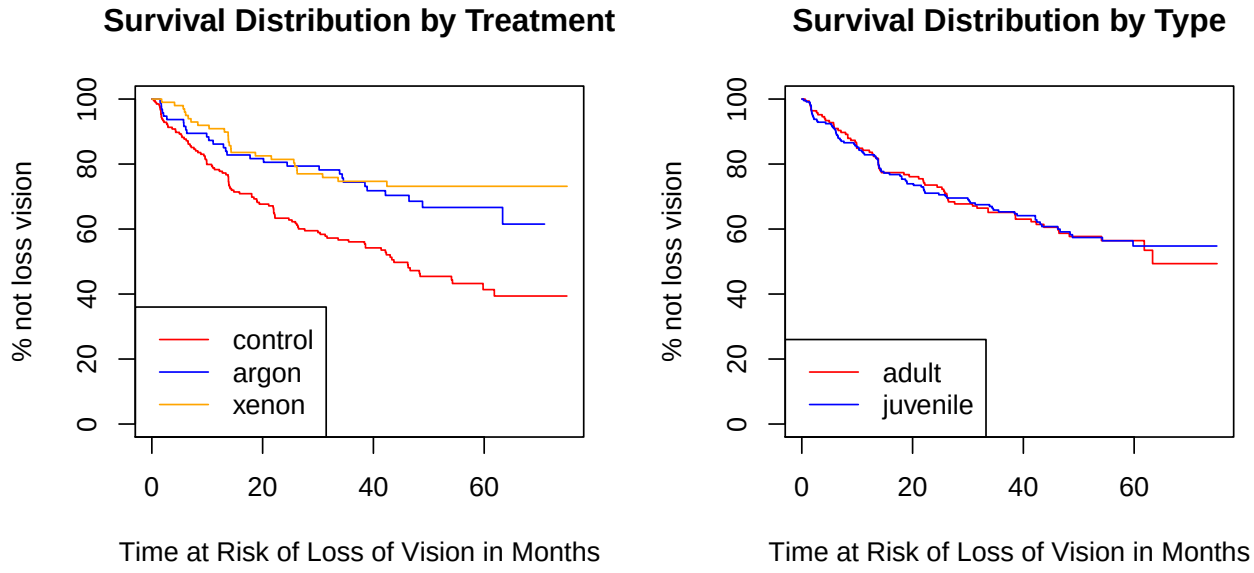
4 Conclusion

This analysis is aimed to quantify the efficacy of `Argon` and `Xenon` treatments, and explains the effect of `age` and `clinical risk` on the risk of visual loss. The Frailty Cox Proportional Hazards Model seems characterize the hazards function quite well in this question. In conclusion, both `Argon` and `Xenon` treatments certainly have significant positive treatment effect on visual acuity, `clinical risk` have negative influence on visual acuity, and we do not have enough evidence to show the effect of `age` on visual acuity.

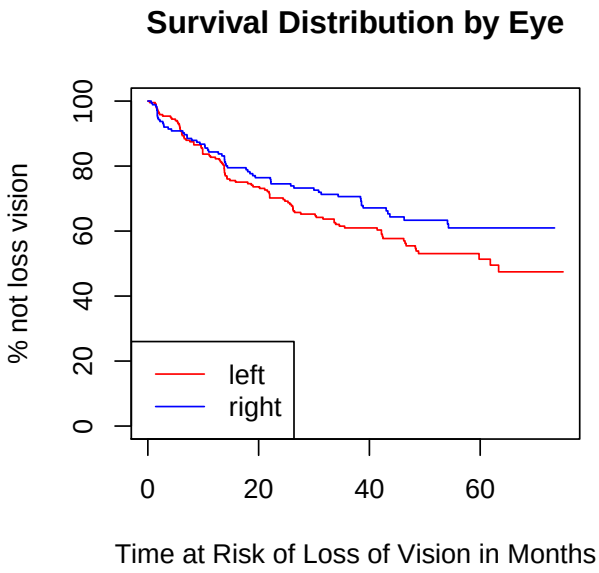
5 Appendix

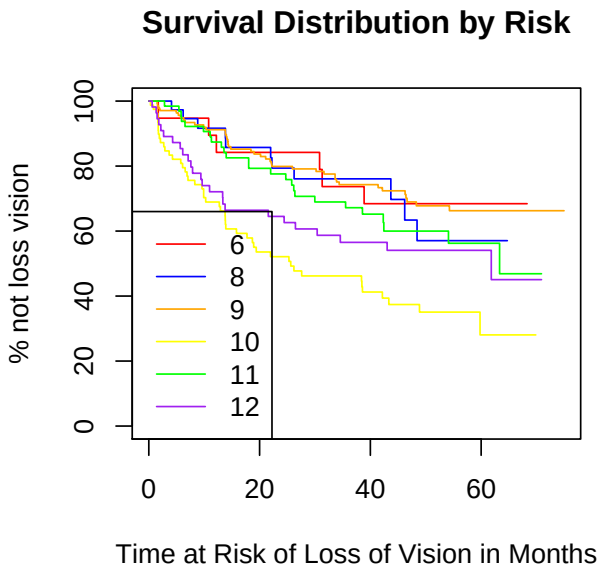
Source code of this report can be found *here*.

```
library(survival)
library(tidyverse)
library(Hmisc)

df = read.delim('diabeticVision.csv', sep=',')
df = df %>%
  mutate(
    tr_eye = ifelse(trt==1, eye, ifelse(eye=='left', 'right', 'left')),
    treatment = ifelse(trt==1, laser, 'control')
  )
df$treatment =  relevel(as.factor(df$treatment), ref='control')

survobj = with(df, Surv(futime, status))

plot_survfit = function(var){
  fitr <- survfit(as.formula(paste0('survobj~', var)), data=df)
  plot(fitr, xlab="Time at Risk of Loss of Vision in Months",
   ylab="% not loss vision", yscale=100,
   main=paste("Survival Distribution by", capitalize(var)),
   col = c('red', 'blue', 'orange', 'yellow', 'green', 'purple'))
  legend('bottomleft', legend=levels(as.factor(df[,var])),
         col = c('red', 'blue', 'orange', 'yellow', 'green', 'purple'), lty=1)
}
par(mfrow=c(1,2))
plot_survfit('treatment')
plot_survfit('type')
```




```
plot_survfit('eye')
plot_survfit('risk')
```






```
dif_res = rbind(
  c('Treatment', survdiff(survobj~treatment, data=df)$chisq,
    1 - pchisq(survdiff(survobj~treatment, data=df)$chisq, df=1)),
  c('Type', survdiff(survobj~type, data=df)$chisq,
    1 - pchisq(survdiff(survobj~type, data=df)$chisq, df=1)),
  c('Eye', survdiff(survobj~eye, data=df)$chisq,
    1 - pchisq(survdiff(survobj~eye, data=df)$chisq, df=1)),
  c('Risk', survdiff(survobj~risk, data=df)$chisq,
    1 - pchisq(survdiff(survobj~risk, data=df)$chisq, df=1)))
dif_res[,c(2)] = round(as.numeric(dif_res[,c(2)]), 3)
dif_res[,c(3)] = round(as.numeric(dif_res[,c(3)]), 6)
dif_res %>% knitr::kable(col.names = c('Variable', 'Chisq2', 'P-value'),
                         caption = 'The Log-rank Test')
```

Table 5: The Log-rank Test

Variable	Chisq2	P-value
Treatment	22.666	2e-06
Type	0.017	0.894859
Eye	2.788	0.094968
Risk	32.962	0

```
cox_model = summary(coxph(survobj~eye+age+risk+treatment, data=df))
cbind(round(cox_model$coefficients[,c(1,4,5)], 4),
      round(cox_model$conf.int,4)) %>%
  knitr::kable(caption = 'The Cox Proportion Hazard Model')
```

Table 6: The Cox Proportion Hazard Model

	coef	z	Pr(>|z|)	exp(coef)	exp(-coef)	lower .95	upper .95
eyeright	-0.2225	-1.3281	0.1841	0.8005	1.2492	0.5765	1.1116
age	0.0059	1.0673	0.2858	1.0059	0.9941	0.9951	1.0168
risk	0.1426	2.5341	0.0113	1.1533	0.8671	1.0328	1.2877
treatmentargon	-0.6918	-3.2730	0.0011	0.5007	1.9974	0.3308	0.7576

	coef	z	Pr(>|z|)	exp(coef)	exp(-coef)	lower .95	upper .95
treatmentxenon	-0.8762	-3.9054	0.0001	0.4163	2.4019	0.2682	0.6463

```
gof = cox.zph(coxph(survobj~eye+age+risk+treatment, data=df))
gof$table %>% knitr::kable(caption = 'Proportional Hazards Assumption Test')
```

Table 7: Proportional Hazards Assumption Test

	chisq	df	p
eye	0.8010425	1	0.3707819
age	0.3635736	1	0.5465282
risk	1.6921822	1	0.1933136
treatment	0.5946964	2	0.7427853
GLOBAL	3.7861712	5	0.5805942

```
# library(coxme)
# sfrail <- coxme(survobj~laser+eye+type+risk + (1 | id),  data = df)

cox_model_frailty = summary(coxph(survobj~eye+age+risk+treatment+frailty(id), data=df))
cbind(round(cox_model_frailty$coefficients[,c(1,4,6)], 4),
      rbind(round(cox_model_frailty$conf.int,4), c(' ', ' ', ' ',' ')))%>%
  knitr::kable(caption = 'The Frailty Cox Proportional Hazards Model')
```

Table 8: The Frailty Cox Proportional Hazards Model

	coef	Chisq	p	exp(coef)	exp(-coef)	lower .95	upper .95
eyeright	-0.3127	2.0069	0.1566	0.7315	1.3671	0.4746	1.1274
age	0.0075	1.0279	0.3107	1.0075	0.9926	0.9931	1.0221
risk	0.1642	5.6454	0.0175	1.1785	0.8486	1.0292	1.3494
treatmentargon	-0.8504	14.1398	2e-04	0.4273	2.3405	0.2743	0.6656
treatmentxenon	-0.9553	16.5165	0	0.3847	2.5995	0.2427	0.6098
frailty(id)	NA	104.0195	0.0235				