Treating diabetic retinopathy

The purpose of this report is to determine the efficacy of treatment type on visual acuity and quantify the improvement between eyes by treatment type. Also, we would analyze the potential impact of age at diagnosis and clinical risk of diabetic retinopathy have on visual activity. The response variable would be the lag-corrected time to lose vision or last follow-up in months. The variables would be the type of laser used for the treatment, age of diagnosis, treatment types of eyes for either controlled or treated, the status of the eye, and clinical risk of loss of acuity. Since our response is the time until the loss of vision, we will use the method called survival analysis which uses the probability of not having an event (lost of vision, in this case) before time t. In conclusion, we will determine whether the efficiency of treatment type, age, and clinical risk on visual activity.

The dataset consists of 10 variables with 394 objects, with no missing data. In the dataset, each id received treatment for either the right eye or the left eye. The type variable is determined by the age variable. Since the treatment is only executed to the eyes with treatment = 1, we changed the type of the laser to none if the treatment = 0 for that object. Figure 1 represents the overall percentage of eyes that did not lose vision per time. Divided the data into each laser type, figure 2 represents the percentage of eyes that did not lose vision per time for each laser type. In the figure, the eyes with no treatment showed a higher ratio of eyes that did not lose vision than eyes with treatment. Xenon treatment showed less efficiency than argon treatment, but the results were not distinguished.

For the quantitative analysis, compared to the non-treatment cases, both argon and xenon treatment showed a statistically meaningful result. In table 1, the argon treatment showed 0.863

more percentage per month to lose vision than no treatment (control eye), and the xenon treatment showed 0.699 more percentage per month to lose vision than no treatment with both p-value smaller than 0.05. However, in table 2, the two treatments themselves are not distinguished with p-value 0.412.

We can also analyze the potential effect of age on the loss of vision. We divided the age into three groups: 0~18, 18~30, and 30~, and compared the survival distribution into graphs. Figure 3 represents the survival distribution over time by age. However, it seems age does not have a potential impact on the treatment since the results for each age distribution are not distinguishable. Table 3 represents the quantitative analysis for the relationship between age distribution and time to lose vision. We can find that both p-values are not significant in this case.

The last thing we would check is the potential effect of clinical risk of loss of acuity to the loss of vision. We divided the clinical risk into three groups: 5~8, 8~10, and 10~12 and regressed over the survival distribution. Figure 4 represents the survival distribution over time by clinical risk. It also seems that clinic risk does not have a potential impact on the treatment since the results for each clinic risk distribution are not distinguishable in the graph. Table 4 represents the quantitative analysis for the relationship between clinic risk distribution and time to lose vision. On the table, the p-values for each clinic risk are not significant.

There may have been interactions between the variables, so we regressed the percentage of loss the vision over time by laser type with age, and laser type with clinical risk. However, table 5 shows that both p-values are not significant in this case also.

In conclusion, we figured out that the treatment is significant to the lag-corrected time to loss of vision, and the eye with treatment showed less percentage of losing vision in the future. However, we could not find a significant result in between a type of lasers, age of diagnosis, and clinical risk of loss of acuity and log-corrected time to loss of vision. We used the survival method to analyze the relationship.

Table 1. Time to lose vision or last follow-up in months for each laser type, compare to the laser = none

	COEF	EXP(COEF)	SE(COEF)	Z	P
LASERARGON	0.863	2.370	0.197	4.388	1.14e-05
LASERXENON	0.699	2.012	0.195	3.580	0.000344

Table 2. Time to lose vision or last follow-up in months for each laser type, compare to the laser = ARGON

	COEF	EXP(COEF)	SE(COEF)	Z	P
LASENONE	-0.863	0.422	0.197	-4.389	1.14e-05
LASERXENON	-0.164	0.849	0.199	-0.821	0.412

Table 3. Time to lose vision or last follow-up in months for each age distribution, compare to the age distribution = (0,18]

	COEF	EXP(COEF)	SE(COEF)	Z	P
AGEDIST(18,30]	-0.0105	0.990	0.208	-0.051	0.960
AGEDIST(30,100]	0.0665	1.069	0.194	0.343	0.731

Table 4.Time to lose vision or last follow-up in months for each clinical risk distribution, compare to the risk distribution = (0,18]

	COEF	EXP(COEF)	SE(COEF)	Z	P
RISKDIST(8,10]	0.350	1.419	0.265	1.319	0.187
RISKDIST(10,12]	0.402	1.494	0.281	1.430	0.153

Table 5. Time to lose vision or last follow-up in months for laser, age and risk

	CHISQ	DF	P
LASER	1.783	2	0.41
AGE	0.968	1	0.33
RISK	1.230	1	0.27
LASER:AGE	0.986	2	0.61
LASER:RISK	1.548	2	0.46
AGE:RISK	0.719	1	0.40
LASER:AGE:RISK	0.972	2	0.62
GLOBAL	6.565	11	0.83
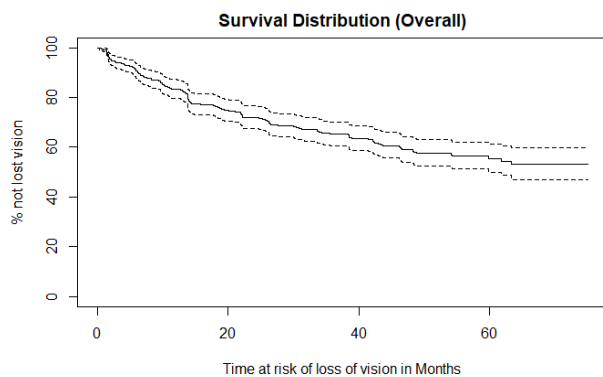
Figure 1. Survival Distribution (Overall)
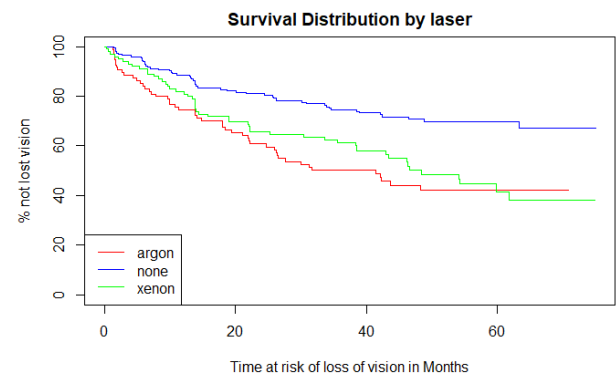
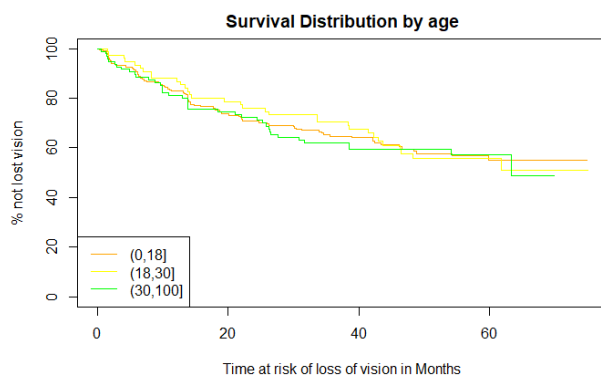*Figure 2. Survival Distribution by laser*
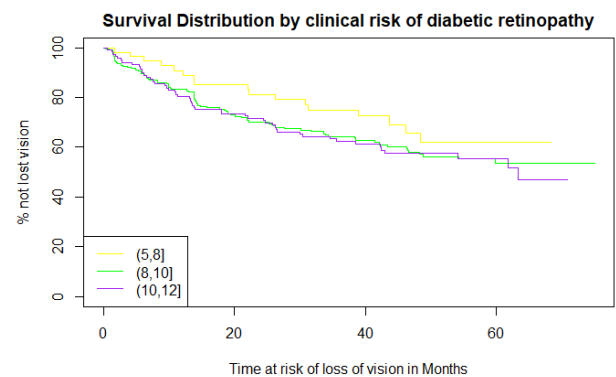

Figure 3. Survival Distribution by age

Figure 4. Survival Distribution by clinical risk

Assignment 2 Appendix

```r
library(survival)
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.0.2     v forcats 0.5.1
```

```
## -- Conflicts -------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
knitr::opts_chunk$set(fig.width=3, fig.height=3)
dat = read.csv('C:/Users/krss9/Desktop/FS21/Stats 504/Assignment2/diabeticVision.csv')

#trt = 0 means that eye did not get a treatment. We will change those laser types to none.
for(i in 1:nrow(dat))
if(dat$trt[i] == 1){
  dat$laser[i] ='none'
}
#dat %>% filter(dat$laser == 'none') %>% head()
```

#Goals ### 1. Determine the efficacy of treatment type on visual acuity ### 2. Quantify the improvement between eyes by treatment type

**3. Understand the potential impact that age at diagnosis have on visual acuity**

4. understand the potential impact that clinical risk of diabetic retinopathy have on visual acuity

0. Overall survival

```r
survobj <- with(dat, Surv(futime, status))
fit0 <- survfit(survobj~1, data=dat)
```

```r
plot(fit0, xlab="Time at risk of loss of vision in Months",
ylab="% not lost vision", yscale=100,
main ="Survival Distribution (Overall)")
```
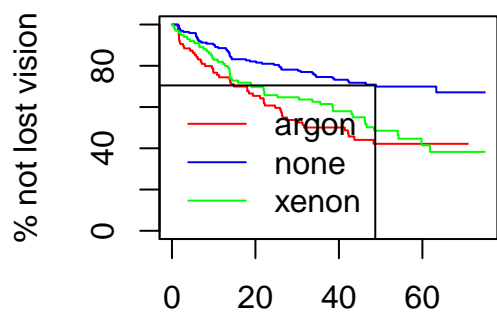
Survival Distribution (Overall
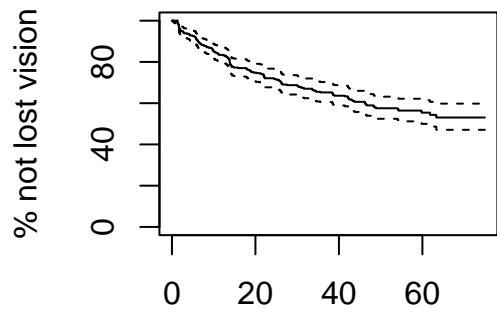
% not lost vision

Time at risk of loss of vision in Month

1. Survival distribution by laser

```
fitr <- survfit(survobj~laser, data=dat)
plot(fitr, xlab="Time at risk of loss of vision in Months",
ylab="% not lost vision", yscale=100,
main="Survival Distribution by laser",
col = c('red', 'blue','green'))
legend('bottomleft', legend=levels(as.factor(dat$laser)), col = c('red', 'blue','green'), lty=1)
```

Survival Distribution by laser

% not lost vision

argon
none
xenon

Time at risk of loss of vision in Month

2. Quantatitive analysis of the survival distribution by laser

```
survdiff(survobj~laser, data=dat)
```

```
## Call:
## survdiff(formula = survobj ~ laser, data = dat)
##
##                N Observed Expected (O-E)^2/E (O-E)^2/V
## laser=argon   97       50     32.6      9.24     11.74
## laser=none   197       54     83.2     10.27     22.25
## laser=xenon  100       51     39.1      3.60      4.82
##
##   Chisq= 23.2  on 2 degrees of freedom, p= 9e-06
```
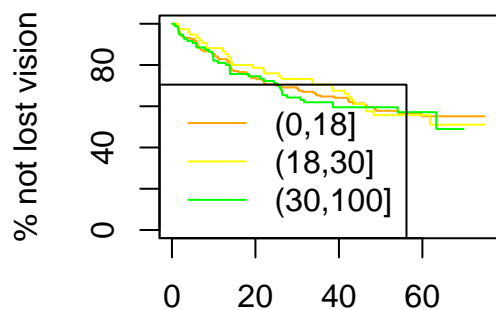
3. Survival Distrubution by age - graph

```
summary(dat$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   10.00   16.00   20.78   30.00   58.00
```

```
dat$agedist <- cut(dat$age, breaks=c(0,18,30,100))
fitc <- survfit(survobj~agedist, data=dat)
plot(fitc,xlab="Time at risk of loss of vision in Months",
ylab="% not lost vision", yscale=100,
main="Survival Distribution by age",
col = c('orange', 'yellow', 'green'))
legend('bottomleft', legend=levels(as.factor(dat$agedist)), col = c('orange', 'yellow', 'green'), lty=1)
```

Survival Distribution by age


Survival distribution by clinical risk - graph

```
summary(dat$risk)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   6.000   9.000  10.000   9.698  11.000  12.000
```

```
dat$riskdist <- cut(dat$risk, breaks=c(5,8,10,12))
fitr <- survfit(survobj~riskdist, data=dat)
plot(fitr,xlab="Time at risk of loss of vision in Months",
ylab="% not lost vision", yscale=100,
main="Survival Distribution by clinical risk of diabetic retinopathy",
col = c('yellow', 'green', 'purple'))
legend('bottomleft', legend=levels(as.factor(dat$riskdist)), col = c('yellow', 'green', 'purple'), lty=
```

stribution by clinical risk of diabe
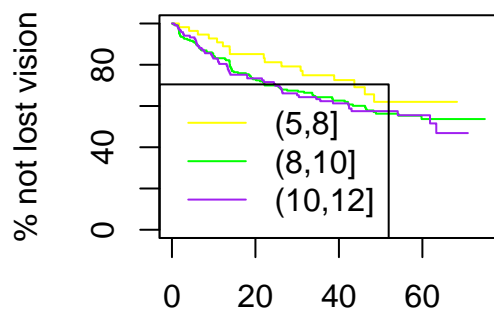
More Quantatitive analysis

```
dat$laser = relevel(as.factor(dat$laser), ref='none')
coxph(survobj~laser, data=dat)
```

```
## Call:
## coxph(formula = survobj ~ laser, data = dat)
##
##               coef exp(coef) se(coef)     z        p
## laserargon 0.8628    2.3697   0.1966 4.388 1.14e-05
## laserxenon 0.6992    2.0122   0.1953 3.580 0.000344
##
## Likelihood ratio test=23.05  on 2 df, p=9.903e-06
## n= 394, number of events= 155
```

```
dat$laser = relevel(as.factor(dat$laser), ref='argon')
coxph(survobj~laser, data=dat)
```

```
## Call:
## coxph(formula = survobj ~ laser, data = dat)
##
##              coef exp(coef) se(coef)      z        p
## lasernone -0.8628    0.4220   0.1966 -4.388 1.14e-05
## laserxenon -0.1635    0.8492   0.1992 -0.821    0.412
##
## Likelihood ratio test=23.05  on 2 df, p=9.903e-06
## n= 394, number of events= 155
```

```
coxph(survobj~agedist, data=dat)
```

```
## Call:
## coxph(formula = survobj ~ agedist, data = dat)
##
##                    coef exp(coef) se(coef)      z     p
## agedist(18,30]  -0.01051   0.98954  0.20789 -0.051 0.960
## agedist(30,100]  0.06657   1.06884  0.19388  0.343 0.731
##
## Likelihood ratio test=0.14  on 2 df, p=0.9316
## n= 394, number of events= 155
```

```
coxph(survobj~riskdist, data=dat)
```

```
## Call:
## coxph(formula = survobj ~ riskdist, data = dat)
##
##                   coef exp(coef) se(coef)     z     p
## riskdist(8,10]  0.3496    1.4185   0.2650 1.319 0.187
## riskdist(10,12] 0.4015    1.4941   0.2808 1.430 0.153
##
## Likelihood ratio test=2.34  on 2 df, p=0.3102
## n= 394, number of events= 155
```

Check the interactino between variables

```
test.dat1 <- cox.zph(coxph(survobj~laser*age*risk, data=dat))
test.dat1
```

```
##               chisq df    p
## laser         1.783  2 0.41
## age           0.968  1 0.33
## risk          1.230  1 0.27
## laser:age     0.986  2 0.61
## laser:risk    1.548  2 0.46
## age:risk      0.719  1 0.40
## laser:age:risk 0.972  2 0.62
## GLOBAL        6.565 11 0.83
```

```
test.dat2 <- cox.zph(coxph(survobj~laser*risk, data=dat))
test.dat2
```

```
##            chisq df    p
## laser       1.84  2 0.40
## risk        1.53  1 0.22
## laser:risk  1.55  2 0.46
## GLOBAL      3.55  5 0.62
```