Survival of C.elegans with three different media

Jonas Gehrlein

12 nov 2018

code for fitting Kaplan-Meier and log-rank test and for displaying survival curves for each type of media

```
data <- read.table(here('data','Worm_4days.csv'), header = TRUE, sep = ";", dec = ",")
head(data)
```

```
##   ï..Time Status Group Replicate
## 1       4      1  FG13         1
## 2       6      1  FG13         1
## 3       6      1  FG13         1
## 4       8      1  FG13         1
## 5       8      1  FG13         1
## 6      10      1  FG13         1
```

here() should show that your position is in the folder 7.semester else select the active project to be 7.semester in the upper right corner.

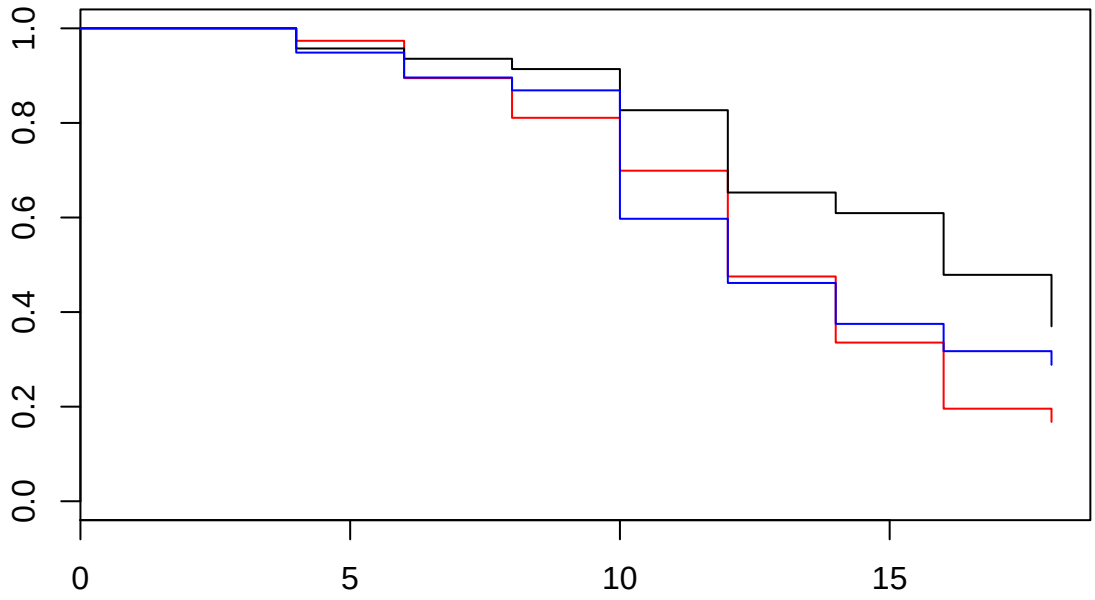
Or make a new .Rproj file in the folder 7.semester. There is a problem with the numbers of the worms so we change that with col.names()

```
colnames(data)<-c("Time","status","media",'Replicate')
head(data)
```

```
##   Time status media Replicate
## 1    4      1  FG13         1
## 2    6      1  FG13         1
## 3    6      1  FG13         1
## 4    8      1  FG13         1
## 5    8      1  FG13         1
## 6   10      1  FG13         1
```

Then we create a survival object with the survival package and make a kaplan-meier curve

```
data$Survobj <- with(data,Surv(data$Time, event = data$status))
km <- survfit(Survobj ~ media, data = data,conf.type = "log-log", error = "greenwood")
s_km <- summary(km)
plot(km, col = c("red","black",'blue'))
```




```
fit <- coxph(Survobj ~ media, data = data)
summary(fit)
```

```
## Call:
## coxph(formula = Survobj ~ media, data = data)
##
##   n= 124, number of events= 85
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG20 -0.6010    0.5482   0.2622 -2.293   0.0219 *
## mediaOP50 -0.1943    0.8234   0.2685 -0.724   0.4693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG20    0.5482      1.824    0.3280    0.9165
## mediaOP50    0.8234      1.214    0.4865    1.3936
##
## Concordance= 0.576  (se = 0.031 )
## Rsquare= 0.044   (max possible= 0.997 )
## Likelihood ratio test= 5.57  on 2 df,   p=0.06
## Wald test            = 5.44  on 2 df,   p=0.07
## Score (logrank) test = 5.57  on 2 df,   p=0.06
```

```
data$media <- factor(data$media, levels = c('OP50','FG13','FG20'))
fit <- coxph(Survobj ~ media, data = data)
summary(fit)
```

```
## Call:
## coxph(formula = Survobj ~ media, data = data)
##
##   n= 124, number of events= 85
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG13  0.1943    1.2144   0.2685  0.724    0.469
## mediaFG20 -0.4068    0.6658   0.2708 -1.502    0.133
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13    1.2144     0.8234    0.7176     2.055
## mediaFG20    0.6658     1.5019    0.3916     1.132
##
## Concordance= 0.576  (se = 0.031 )
## Rsquare= 0.044   (max possible= 0.997 )
## Likelihood ratio test= 5.57  on 2 df,   p=0.06
## Wald test            = 5.44  on 2 df,   p=0.07
## Score (logrank) test = 5.57  on 2 df,   p=0.06
```

Shows errorbars and lineplot based on one of Anders papers https://onlinelibrary.wiley.com/doi/full/10.1111/acel.12165 This is done by extracting the surviving proportion, standard errors and time from survfit(). and adding a startpoint where the survival is 100% at 0 hours.

```
df_fly_13 <- data.frame(c(0,s_km$time[1:8]),c(1,s_km$surv[1:8]),
                        c(0,s_km$std.err[1:8]))
colnames(df_fly_13) <- c('Time','Surv','Std.error')
df_fly_20 <- data.frame(c(0,s_km$time[9:16]),c(1,s_km$surv[9:16]),
                        c(0,s_km$std.err[9:16]))
colnames(df_fly_20) <- c('Time','Surv','Std.error')
df_OP50 <- data.frame(c(0,s_km$time[17:24]),c(1,s_km$surv[17:24]),
                        c(0,s_km$std.err[17:24]))
colnames(df_OP50) <- c('Time','Surv','Std.error')

plot(df_fly_13$Time[2:9],df_fly_13$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time (hours)', ylab = 'Survi
     main = expression('Heat knockdown survival of' ~italic(C.) ~italic(elegans)),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13$Time,df_fly_13$Surv)
arrows(df_fly_13$Time, df_fly_13$Surv-df_fly_13$Std.error, df_fly_13$Time,
       df_fly_13$Surv+df_fly_13$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_13$Time, df_fly_13$Surv - df_fly_13$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

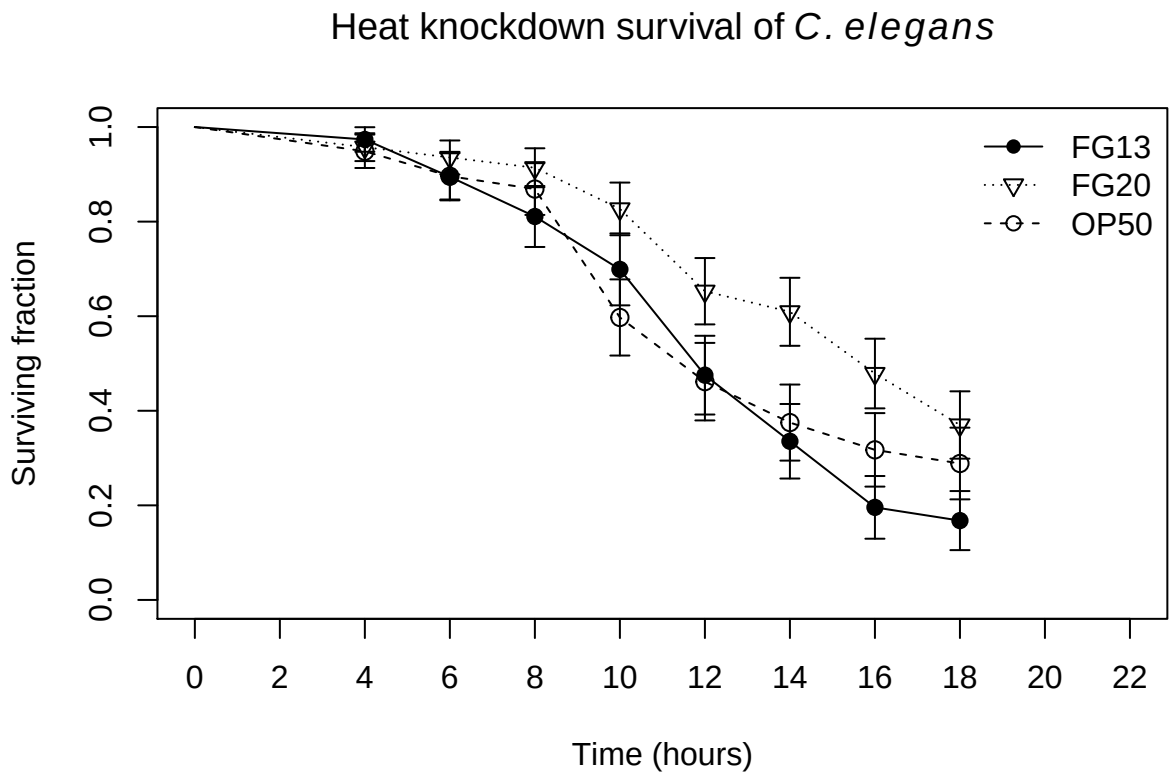
```
points(df_OP50$Time[2:9],df_OP50$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50$Time,df_OP50$Surv, lty = 2)
arrows(df_OP50$Time, df_OP50$Surv-df_OP50$Std.error,
      df_OP50$Time, df_OP50$Surv+df_OP50$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50$Time, df_OP50$Surv - df_OP50$Std.error,
## df_OP50$Time, : zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20$Time[2:9],df_fly_20$Surv[2:9], pch = 6)
lines(df_fly_20$Time,df_fly_20$Surv,lty = 3)
arrows(df_fly_20$Time, df_fly_20$Surv-df_fly_20$Std.error, df_fly_20$Time,
       df_fly_20$Surv+df_fly_20$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

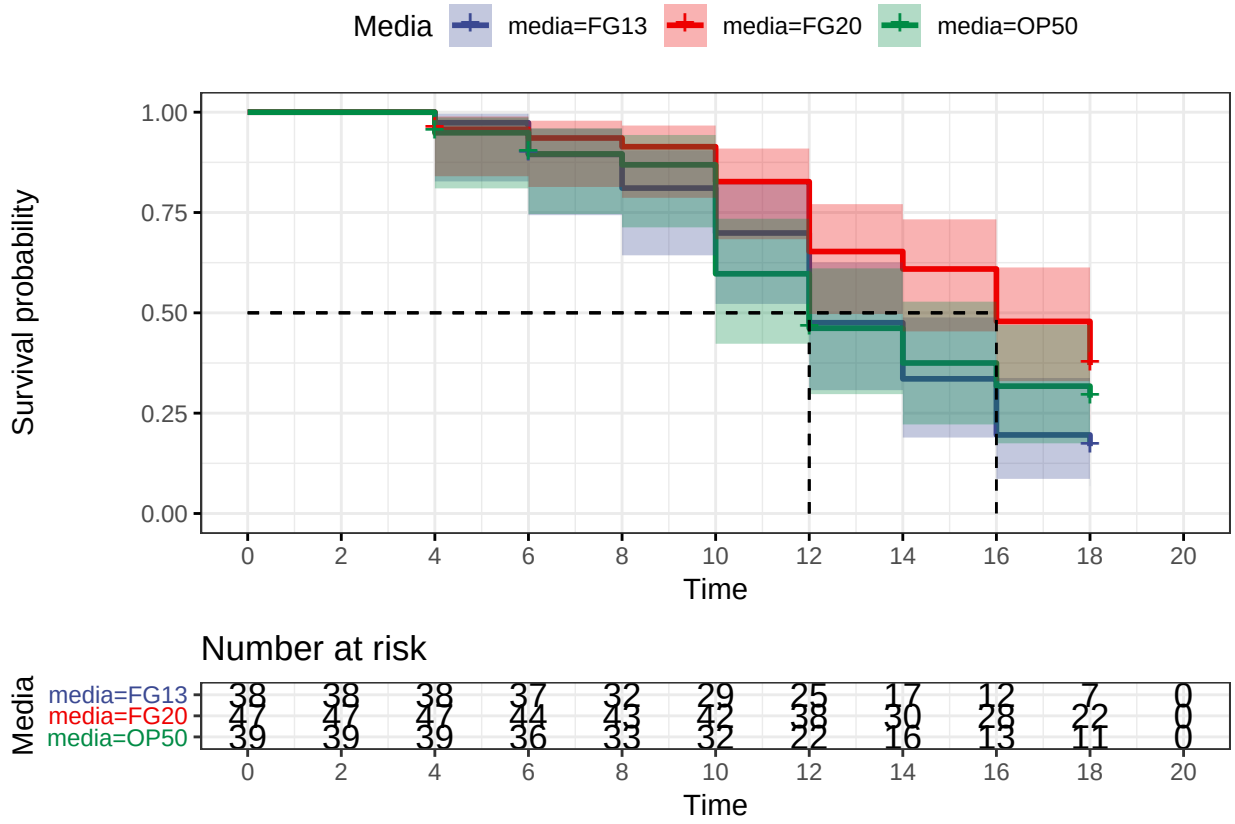
```
## Warning in arrows(df_fly_20$Time, df_fly_20$Surv - df_fly_20$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
legend('topright', pch = c(16,6,1),lty = c(1,3,2),
       legend = c('FG13','FG20','OP50'), bty = 'n')
```



Then we create a normal Kaplan-Meier curve

```
ggsurvplot(km,data = data, conf.int = TRUE, ggtheme = theme_bw(),
           risk.table = 0.25, palette ='aaas', surv.median.line = 'hv',legend.title = 'Media',break.x.by
```



Now we test for difference between the curves with both log-rank and gehan-wilcoxon and with an cox proportional hazard model

```
survdiff(Survobj ~media, data = data, rho = 0)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = data, rho = 0)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50  39       26     23.6     0.238     0.396
## media=FG13  38       30     23.0     2.115     3.498
## media=FG20  47       29     38.4     2.280     5.052
##
##  Chisq= 5.6  on 2 degrees of freedom, p= 0.06
```

```
survdiff(Survobj ~media, data = data, rho = 1)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = data, rho = 1)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50  39     20.0     17.0     0.531      1.13
## media=FG13  38     21.1     16.7     1.149      2.42
## media=FG20  47     18.6     26.0     2.101      5.80
##
##  Chisq= 5.9  on 2 degrees of freedom, p= 0.05
```

Then we compare the different groups individually

```
d_OP50 <- data[data$media == 'OP50',]
d_13 <- data[data$media == "FG13",]
d_20 <- data[data$media == "FG20",]
d_fly <- rbind(d_13, d_20)
d_2050 <- rbind(d_20, d_OP50)
d_1350 <- rbind(d_13, d_OP50)
```

First tests for each combination first fly media

```
survdiff(Surv(Time,status) ~ media, data = d_fly, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_fly,
##     rho = 0)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 38       30     21.8      3.11         6
## media=FG20 47       29     37.2      1.82         6
##
##  Chisq= 6  on 1 degrees of freedom, p= 0.01
```

```
survdiff(Surv(Time,status) ~ media, data = d_fly, rho = 1)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_fly,
##     rho = 1)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 38     22.0     16.0      2.23      5.56
## media=FG20 47     19.4     25.3      1.41      5.56
##
##  Chisq= 5.6  on 1 degrees of freedom, p= 0.02
```

```
fit_fly <- coxph(Surv(Time,status) ~media, data = d_fly)
summary(fit_fly)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ media, data = d_fly)
##
##   n= 85, number of events= 59
##
##             coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13 0.6352    1.8875   0.2632 2.413   0.0158 *
## mediaFG20     NA        NA   0.0000    NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13     1.887     0.5298     1.127     3.162
```

```
## mediaFG20        NA         NA        NA        NA
##
## Concordance= 0.587  (se = 0.035 )
## Rsquare= 0.065   (max possible= 0.996 )
## Likelihood ratio test= 5.75  on 1 df,   p=0.02
## Wald test            = 5.82  on 1 df,   p=0.02
## Score (logrank) test = 6.01  on 1 df,   p=0.01
```

Then 20 degrees and OP50

```
survdiff(Surv(Time,status) ~ media, data = d_2050, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_2050,
##     rho = 0)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 39       26     21.1     1.123      2.16
## media=FG20 47       29     33.9     0.701      2.16
##
##  Chisq= 2.2  on 1 degrees of freedom, p= 0.1
```

```
survdiff(Surv(Time,status) ~ media, data = d_2050, rho = 1)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_2050,
##     rho = 1)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 39     20.5     15.8      1.37      3.38
## media=FG20 47     19.4     24.1      0.90      3.38
##
##  Chisq= 3.4  on 1 degrees of freedom, p= 0.07
```

```
fit_2050 <- coxph(Surv(Time,status) ~media, data = d_2050)
summary(fit_2050)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ media, data = d_2050)
##
##   n= 86, number of events= 55
##
##              coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13      NA        NA   0.0000    NA       NA
## mediaFG20 -0.3933    0.6748   0.2712 -1.45    0.147
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13        NA         NA        NA        NA
## mediaFG20    0.6748      1.482    0.3966     1.148
##
## Concordance= 0.569  (se = 0.037 )
## Rsquare= 0.024   (max possible= 0.994 )
```

```
## Likelihood ratio test= 2.08  on 1 df,   p=0.1
## Wald test            = 2.1  on 1 df,   p=0.1
## Score (logrank) test = 2.13  on 1 df,   p=0.1
```

Then for 13 degrees and OP50

```
survdiff(Surv(Time,status) ~ media, data = d_1350, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_1350,
##     rho = 0)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 39       26     28.2     0.170     0.423
## media=FG13 38       30     27.8     0.172     0.423
##
##  Chisq= 0.4  on 1 degrees of freedom, p= 0.5
```

```
survdiff(Surv(Time,status) ~ media, data = d_1350, rho = 1)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_1350,
##     rho = 1)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 39     18.8     19.2   0.00879    0.0279
## media=FG13 38     19.5     19.1   0.00887    0.0279
##
##  Chisq= 0  on 1 degrees of freedom, p= 0.9
```

```
fit_1350 <- coxph(Surv(Time,status) ~media, data = d_1350)
summary(fit_1350)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ media, data = d_1350)
##
##   n= 77, number of events= 56
##
##             coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13 0.1731    1.1890   0.2685 0.645    0.519
## mediaFG20     NA        NA   0.0000    NA       NA
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13     1.189      0.841    0.7025     2.013
## mediaFG20        NA         NA        NA        NA
##
## Concordance= 0.506  (se = 0.04 )
## Rsquare= 0.005   (max possible= 0.996 )
## Likelihood ratio test= 0.42  on 1 df,   p=0.5
## Wald test            = 0.42  on 1 df,   p=0.5
## Score (logrank) test = 0.42  on 1 df,   p=0.5
```

Where only 20 degrees and 13 degrees are significantly different from each other. But there is also some difference between 20 degrees and OP50. Where there isn´t a big difference between 13 degrees and OP50.

Then dose response curves with first the two fly media.

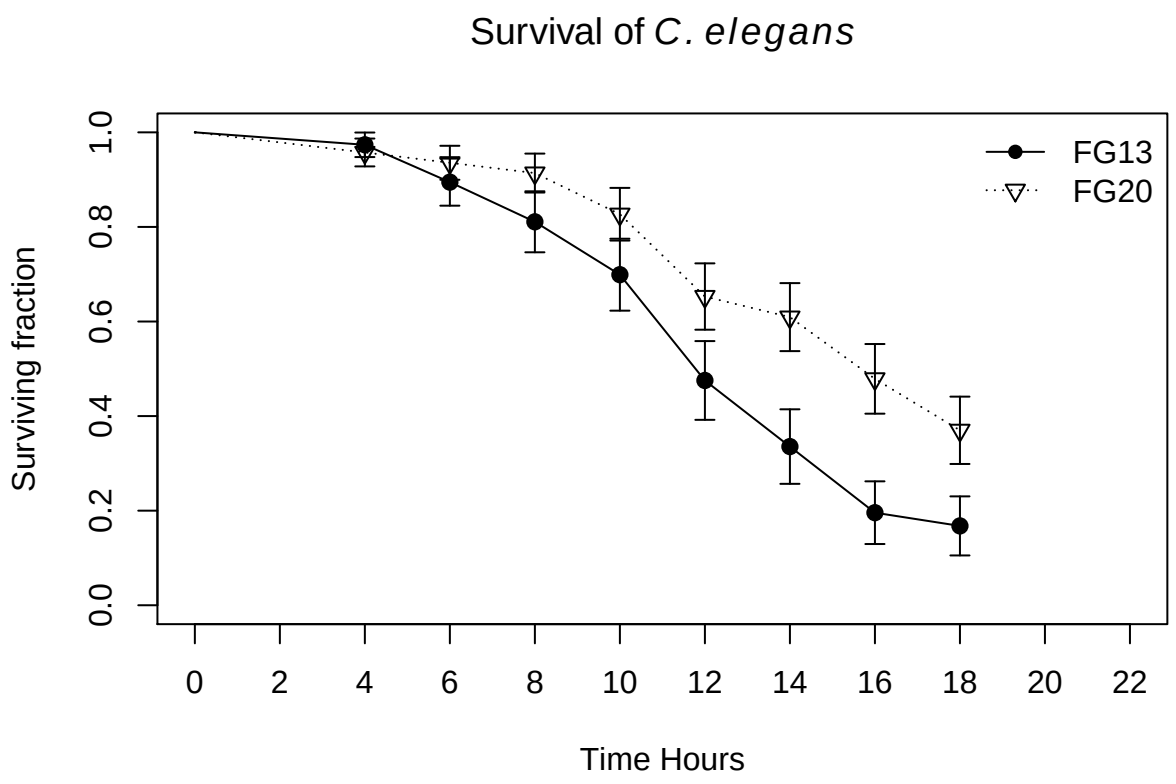
```
plot(df_fly_13$Time[2:9],df_fly_13$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time Hours',
     ylab = 'Surviving fraction',
     main = expression('Survival of'~italic(C.) ~italic(elegans)),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13$Time,df_fly_13$Surv)
arrows(df_fly_13$Time, df_fly_13$Surv-df_fly_13$Std.error, df_fly_13$Time,
       df_fly_13$Surv+df_fly_13$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_13$Time, df_fly_13$Surv - df_fly_13$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20$Time[2:9],df_fly_20$Surv[2:9], pch = 6)
lines(df_fly_20$Time,df_fly_20$Surv,lty = 3)
arrows(df_fly_20$Time, df_fly_20$Surv-df_fly_20$Std.error, df_fly_20$Time,
       df_fly_20$Surv+df_fly_20$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_20$Time, df_fly_20$Surv - df_fly_20$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
legend('topright', pch = c(16,6),lty = c(1,3),
       legend = c('FG13','FG20'), bty = 'n')
```



Than 20 degrees and OP50

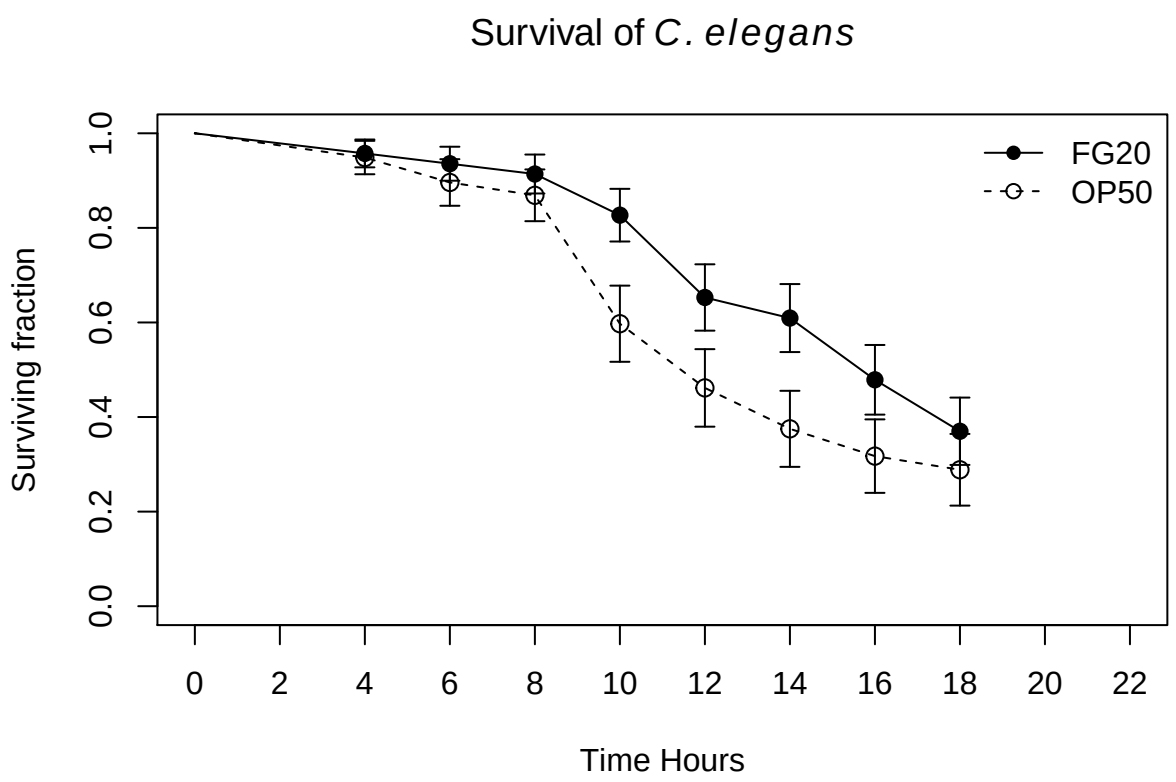
```
plot(df_fly_20$Time[2:9],df_fly_20$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time Hours',
     ylab = 'Surviving fraction',
     main = expression('Survival of'~italic(C.) ~italic(elegans)),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
points(df_OP50$Time[2:9],df_OP50$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50$Time,df_OP50$Surv, lty = 2)
arrows(df_OP50$Time, df_OP50$Surv-df_OP50$Std.error,
      df_OP50$Time, df_OP50$Surv+df_OP50$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50$Time, df_OP50$Surv - df_OP50$Std.error,
## df_OP50$Time, : zero-length arrow is of indeterminate angle and so skipped
```

```
lines(df_fly_20$Time,df_fly_20$Surv,lty = 1)
arrows(df_fly_20$Time, df_fly_20$Surv-df_fly_20$Std.error, df_fly_20$Time,
       df_fly_20$Surv+df_fly_20$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_20$Time, df_fly_20$Surv - df_fly_20$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
legend('topright', pch = c(16,1),lty = c(1,2),
       legend = c('FG20', 'OP50'), bty = 'n')
```



Then 13 degrees and OP50

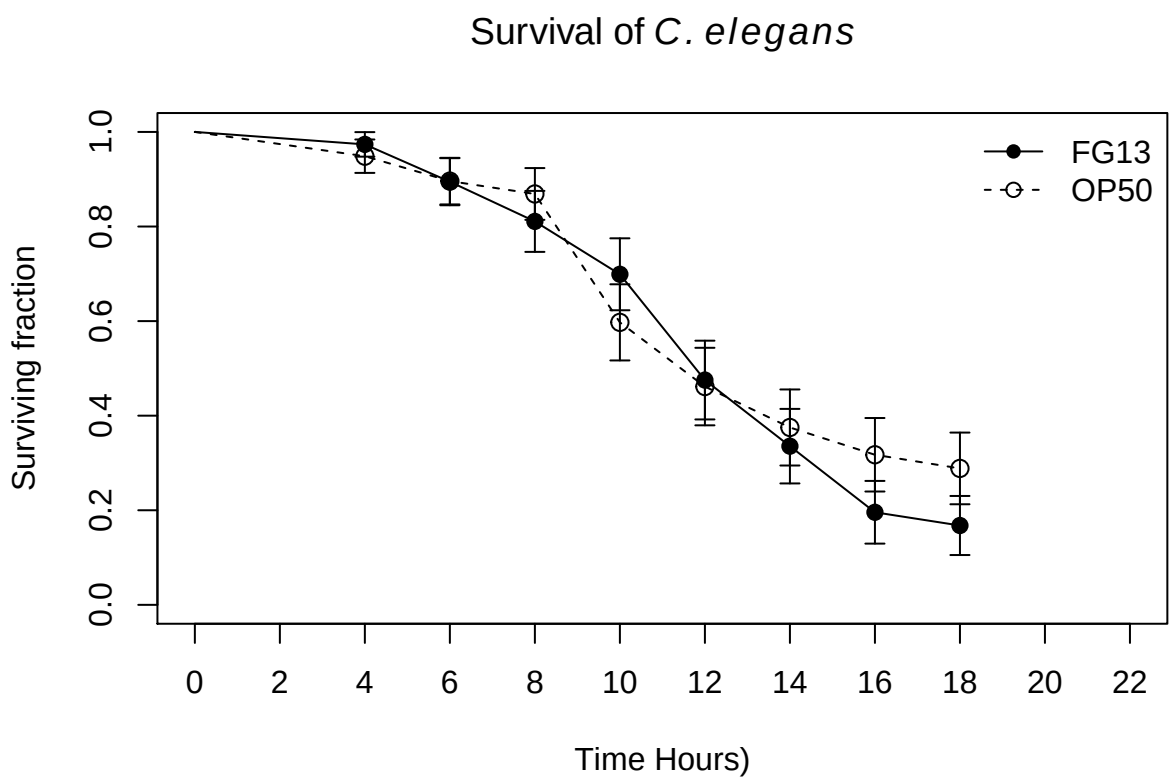
```
plot(df_fly_13$Time[2:9],df_fly_13$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time Hours)',
     ylab = 'Surviving fraction',
     main = expression('Survival of'~italic(C.) ~italic(elegans)),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13$Time,df_fly_13$Surv)
arrows(df_fly_13$Time, df_fly_13$Surv-df_fly_13$Std.error, df_fly_13$Time,
       df_fly_13$Surv+df_fly_13$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_13$Time, df_fly_13$Surv - df_fly_13$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_OP50$Time[2:9],df_OP50$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50$Time,df_OP50$Surv, lty = 2)
arrows(df_OP50$Time, df_OP50$Surv-df_OP50$Std.error,
      df_OP50$Time, df_OP50$Surv+df_OP50$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

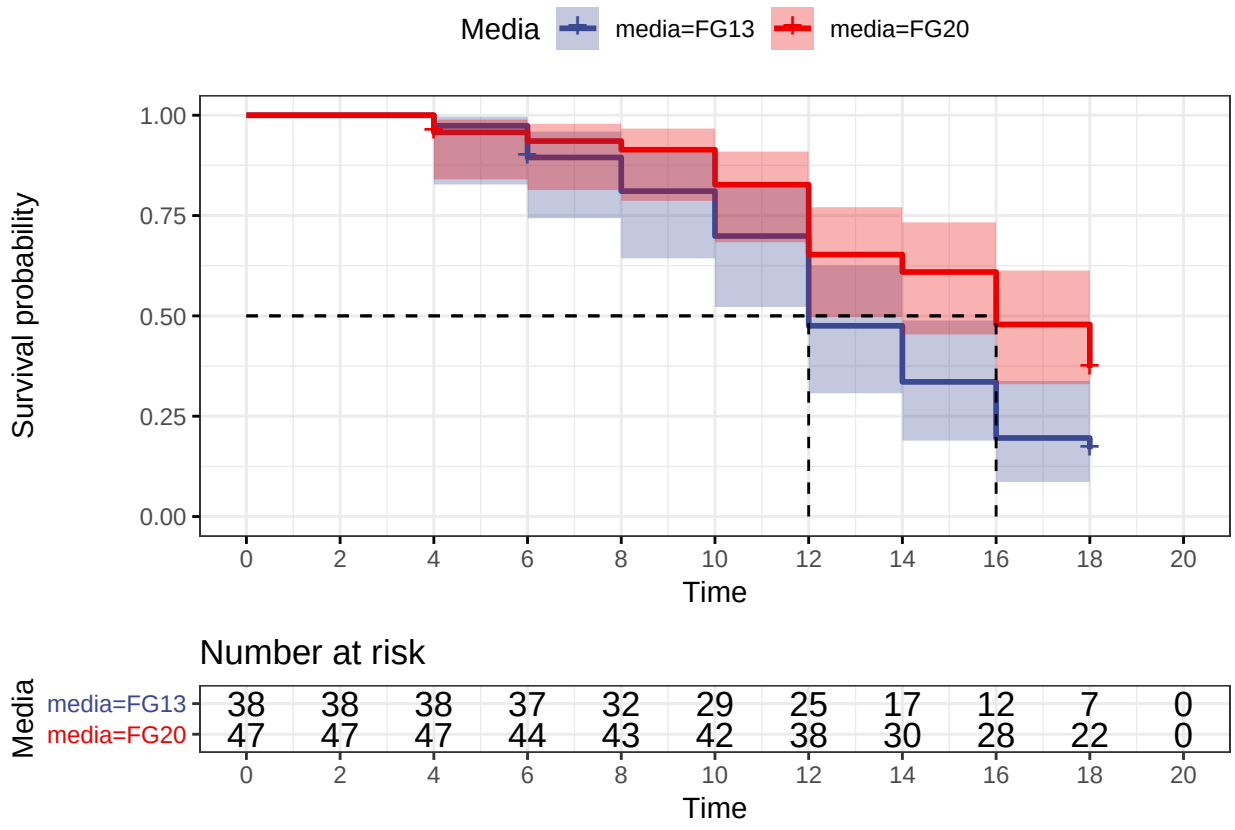
```
## Warning in arrows(df_OP50$Time, df_OP50$Surv - df_OP50$Std.error,
## df_OP50$Time, : zero-length arrow is of indeterminate angle and so skipped
```

```
legend('topright', pch = c(16,1),lty = c(1,2),
       legend = c('FG13','OP50'), bty = 'n')
```

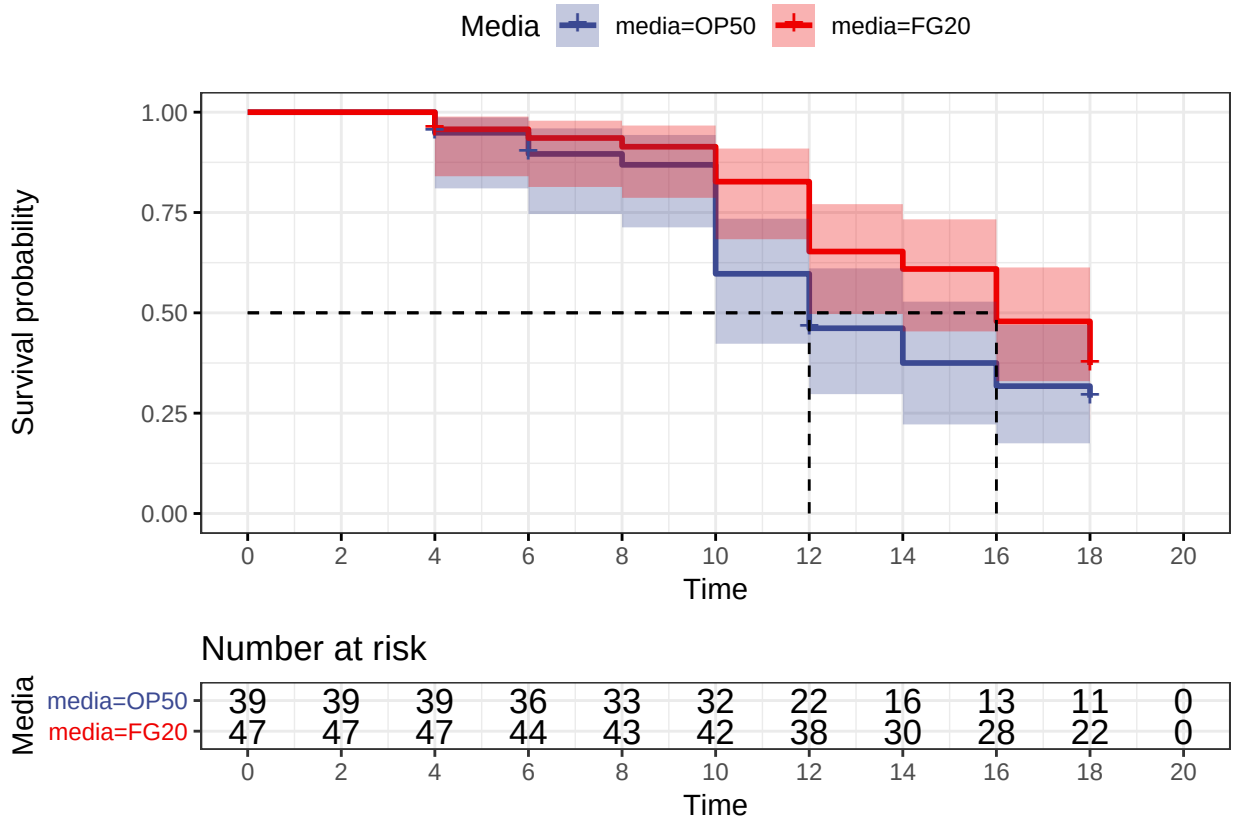


Then Kaplan meier curves

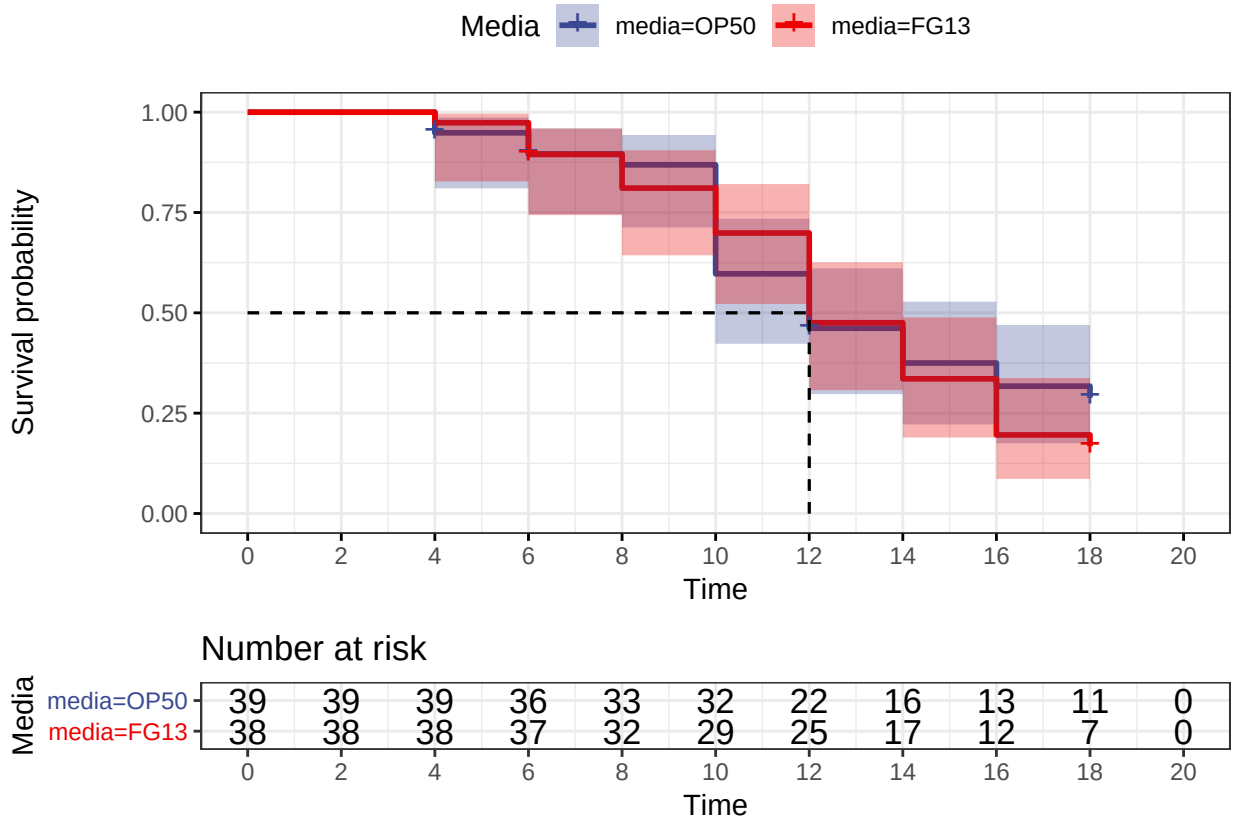
```r
km_fly <- survfit(Surv(Time,status) ~ media, data = d_fly,conf.type = "log-log", error = "greenwood")
km_2050 <- survfit(Surv(Time,status) ~ media, data = d_2050,conf.type = "log-log", error = "greenwood")
km_1350 <- survfit(Surv(Time,status) ~ media, data = d_1350,conf.type = "log-log", error = "greenwood")
ggsurvplot(km_fly,data = d_fly, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2)
```




```r
ggsurvplot(km_2050,data = d_2050, conf.int = TRUE,
           ggtheme = theme_bw(), risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2)
```




```
ggsurvplot(km_1350,data = d_1350, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media', break.x.by = 2)
```



Now we use the replicate dataset

```
data_2 <- read.table(here('data', 'worm_second.csv'), header = TRUE,
                     sep = ';')
head(data_2)
```

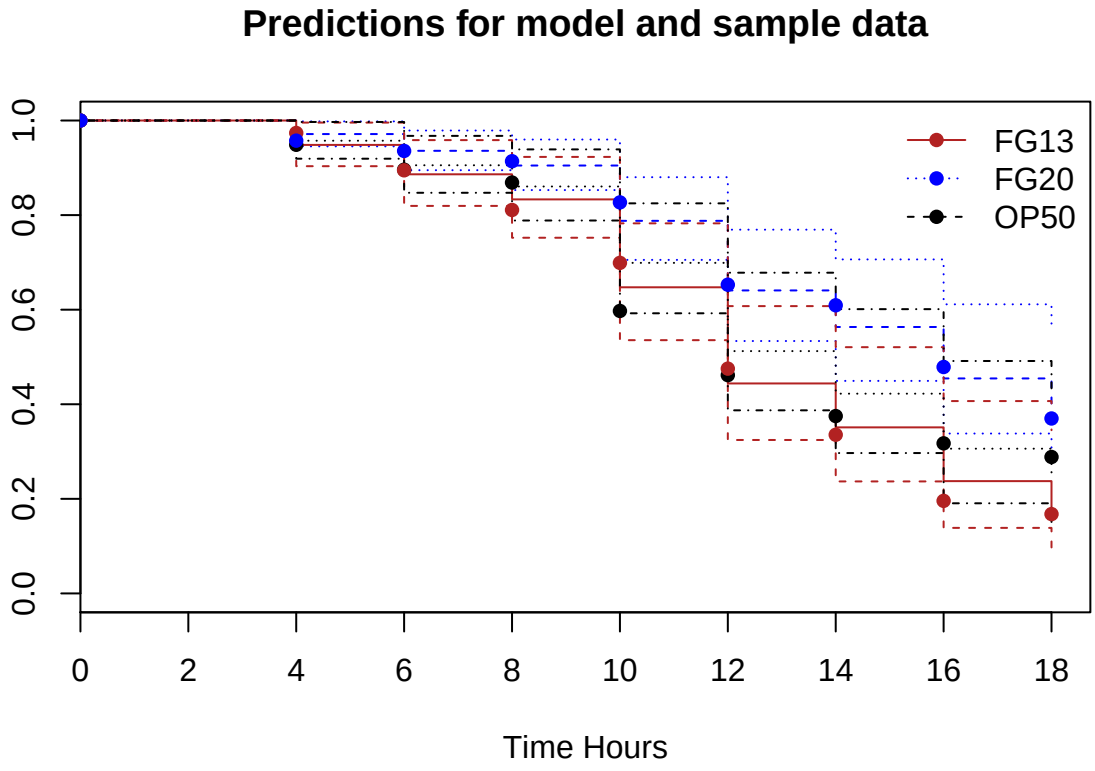
```
##   ï..Time Group Metabolite Status
## 1       4  FG20          0      1
## 2       4  FG20          0      1
## 3       8  FG20          0      1
## 4      12  FG20          0      1
## 5       4  FG20          1      1
## 6       4  FG20          1      1
```

```
colnames(data_2) <- c('Time','media','Metabolite','status')
data_2_rep <- subset(data_2, Metabolite == 0)
```

First we try to see if we can extend the cox model from the earlier dataset to the new one to see hoow well we can extrapolate from one dataset to the other

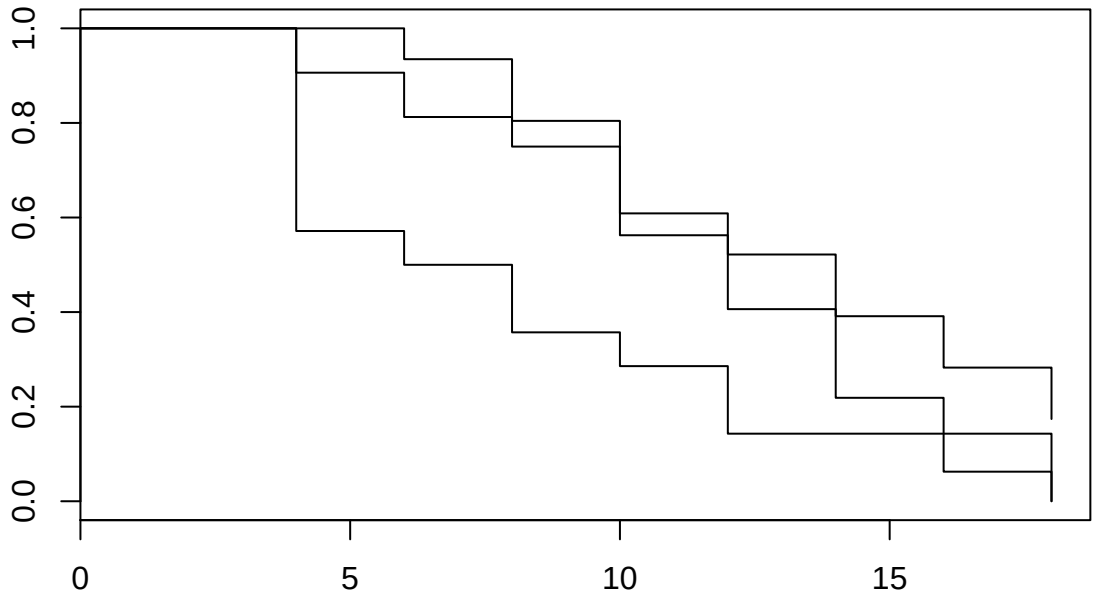
```
plot(survfit(fit, newdata = data.frame(media = 'FG20')), lty = 2, xaxp = c(0,18,9), col = 'blue',
     main = 'Predictions for model and sample data',
     xlab = 'Time Hours')
lines(survfit(fit, newdata = data.frame(media = 'FG13')), col = 'firebrick')
lines(survfit(fit, newdata = data.frame(media = 'OP50')), lty =3)
```

```
points(df_OP50$Time, df_OP50$Surv, col = 'black', pch = 16)
points(df_fly_13$Time,df_fly_13$Surv, col = 'firebrick', pch = 16)
points(df_fly_20$Time,df_fly_20$Surv, col = 'blue', pch = 16)
legend('topright', pch = 16,lty = c(1,3,2),
       col = c('firebrick','blue','black'),
       legend = c('FG13','FG20','OP50'), bty = 'n')
```



Here the data fits well if we do the same for the new data set.

```
data_2_rep$Survobj <- with(data_2_rep,
                           Surv(data_2_rep$Time, event = data_2_rep$status))
km_2 <- survfit(Survobj ~ media, data = data_2_rep,conf.type = "log-log",
                error = "greenwood")
s_km_2 <- summary(km_2)
plot(km_2)
```




```
fit_2 <- coxph(Survobj ~ media, data = data_2_rep)
summary(fit_2)
```

```
## Call:
## coxph(formula = Survobj ~ media, data = data_2_rep)
##
##   n= 92, number of events= 84
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG20  0.4415    1.5551   0.3237  1.364   0.1726
## mediaOP50 -0.5814    0.5591   0.2467 -2.357   0.0184 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG20    1.5551     0.6431    0.8246    2.9328
## mediaOP50    0.5591     1.7886    0.3447    0.9068
##
## Concordance= 0.618  (se = 0.036 )
## Rsquare= 0.118   (max possible= 0.999 )
## Likelihood ratio test= 11.54  on 2 df,   p=0.003
## Wald test            = 12.18  on 2 df,   p=0.002
## Score (logrank) test = 12.87  on 2 df,   p=0.002
```

```
data_2_rep$media <- factor(data_2_rep$media, c('OP50','FG13','FG20'))
fit_2 <- coxph(Survobj ~ media, data = data_2_rep)
summary(fit_2)
```

```
## Call:
## coxph(formula = Survobj ~ media, data = data_2_rep)
##
##   n= 92, number of events= 84
##
##             coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13 0.5814    1.7886   0.2467 2.357  0.01844 *
## mediaFG20 1.0230    2.7814   0.3152 3.245  0.00117 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13     1.789     0.5591     1.103     2.901
## mediaFG20     2.781     0.3595     1.499     5.159
##
## Concordance= 0.618  (se = 0.036 )
## Rsquare= 0.118   (max possible= 0.999 )
## Likelihood ratio test= 11.54  on 2 df,   p=0.003
## Wald test            = 12.18  on 2 df,   p=0.002
## Score (logrank) test = 12.87  on 2 df,   p=0.002
```

We see that Fly gut 20 has extremely bad survival and our former predictions do not hold for the new dataset.

If we try to plot the new data for old predictions

```
df_fly_13_2 <- data.frame(c(0,s_km_2$time[1:8]),c(1,s_km_2$surv[1:8]),
                          c(0,s_km_2$std.err[1:8]))
colnames(df_fly_13_2) <- c('Time','Surv','Std.error')
df_fly_20_2 <- data.frame(c(0,s_km_2$time[9:14]),c(1,s_km_2$surv[9:14]),
                          c(0,s_km_2$std.err[9:14]))
colnames(df_fly_20_2) <- c('Time','Surv','Std.error')
df_OP50_2 <- data.frame(c(0,s_km_2$time[15:24]),c(1,s_km_2$surv[15:24]),
                          c(0,s_km_2$std.err[15:24]))
colnames(df_OP50_2) <- c('Time','Surv','Std.error')

plot(df_fly_13_2$Time[2:9],df_fly_13_2$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time (hours)',
     ylab = 'Surviving fraction',
     main = expression('Survival of'~italic(C.) ~italic(elegans) ~'replicate 2' ),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13_2$Time,df_fly_13_2$Surv)
arrows(df_fly_13_2$Time, df_fly_13_2$Surv-df_fly_13_2$Std.error, df_fly_13_2$Time, df_fly_13_2$Surv+df_
```

```
## Warning in arrows(df_fly_13_2$Time, df_fly_13_2$Surv -
## df_fly_13_2$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

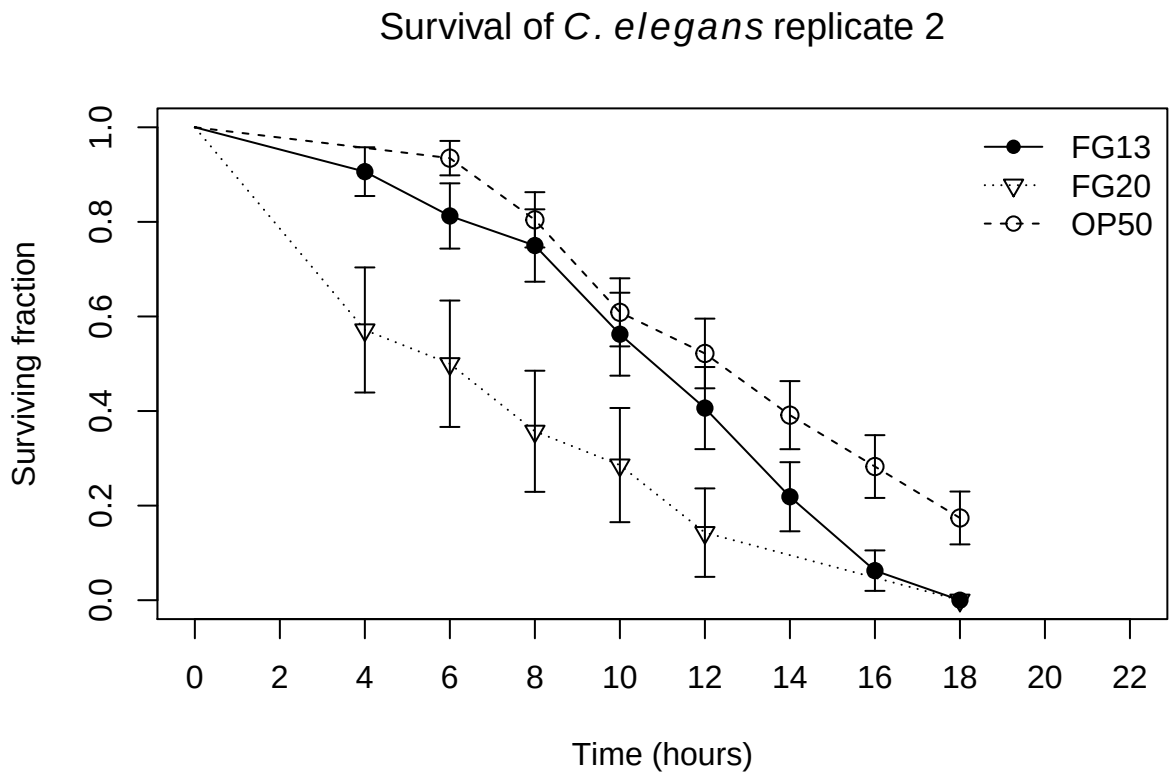
```
points(df_OP50_2$Time[2:9],df_OP50_2$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50_2$Time,df_OP50_2$Surv, lty = 2)
arrows(df_OP50_2$Time, df_OP50_2$Surv-df_OP50_2$Std.error,
      df_OP50_2$Time, df_OP50_2$Surv+df_OP50_2$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50_2$Time, df_OP50_2$Surv - df_OP50_2$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20_2$Time[2:9],df_fly_20_2$Surv[2:9], pch = 6)
lines(df_fly_20_2$Time,df_fly_20_2$Surv,lty = 3)
arrows(df_fly_20_2$Time, df_fly_20_2$Surv-df_fly_20_2$Std.error, df_fly_20_2$Time, df_fly_20_2$Surv+df_
```

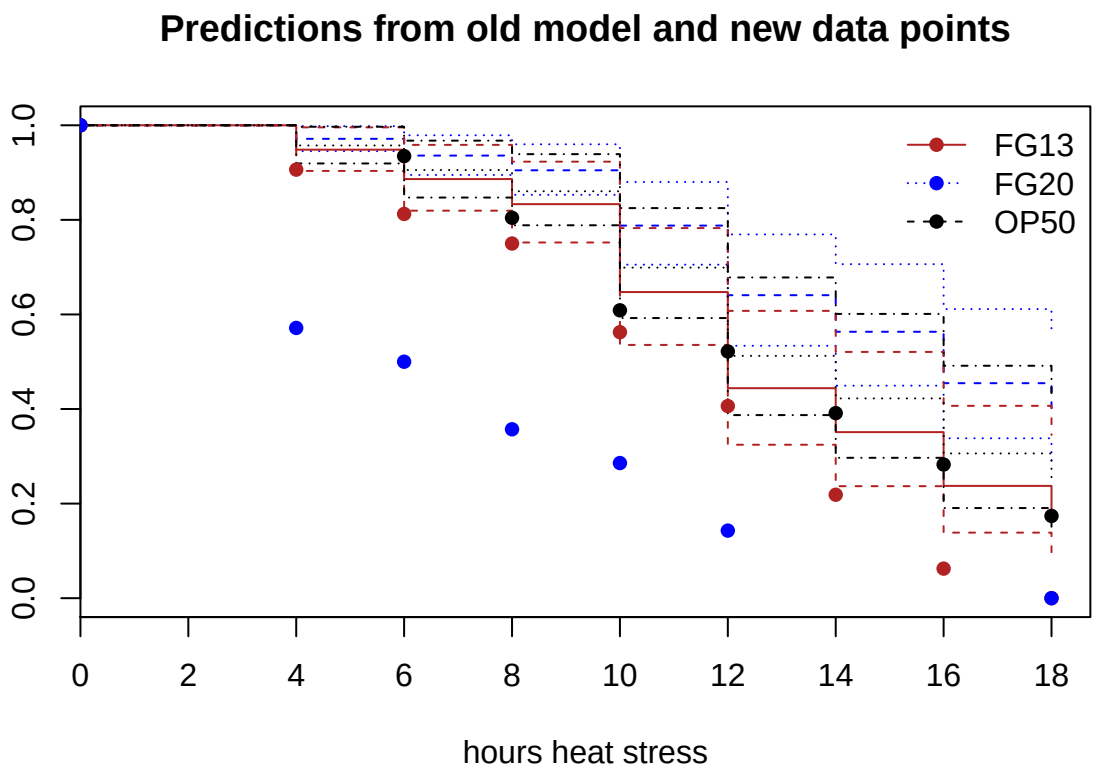
```
## Warning in arrows(df_fly_20_2$Time, df_fly_20_2$Surv -
## df_fly_20_2$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

```
legend('topright', pch = c(16,6,1),lty = c(1,3,2),
       legend = c('FG13','FG20','OP50'), bty = 'n')
```




```
plot(survfit(fit, newdata = data.frame(media = 'FG20')), lty = 2, xaxp = c(0,18,9), col = 'blue',
     main = 'Predictions from old model and new data points',
     xlab = 'hours heat stress')
```

```
lines(survfit(fit, newdata = data.frame(media = 'FG13')), col = 'firebrick')
lines(survfit(fit, newdata = data.frame(media = 'OP50')), lty =3)
points(df_OP50_2$Time, df_OP50_2$Surv, col = 'black', pch = 16)
points(df_fly_13_2$Time,df_fly_13_2$Surv, col = 'firebrick', pch = 16)
points(df_fly_20_2$Time,df_fly_20_2$Surv, col = 'blue', pch = 16)
legend('topright', pch = 16,lty = c(1,3,2),
       col = c('firebrick','blue','black'),
       legend = c('FG13','FG20','OP50'), bty = 'n')
```



Both FG13 and OP50 is slightly lower than before but the survival in F20 has cratered and this may have happened because of something which liquefied the worms on the plates.

Then we compare data with standard errors with one group in each plot

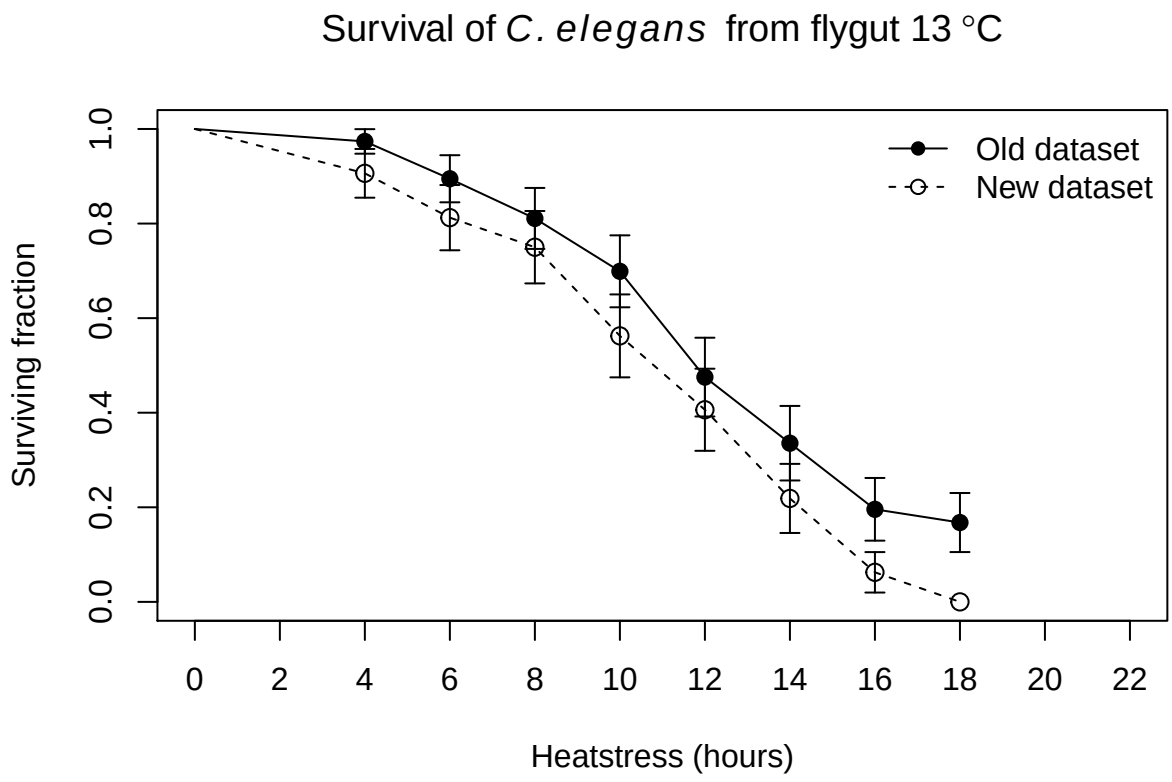
```
plot(df_fly_13$Time[2:9],df_fly_13$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Heatstress (hours)',
     ylab = 'Surviving fraction',
     main = expression('Survival of' ~ italic(C.) ~italic(elegans) ~' from flygut 13' ~degree*C),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13$Time,df_fly_13$Surv)
arrows(df_fly_13$Time, df_fly_13$Surv-df_fly_13$Std.error, df_fly_13$Time,
       df_fly_13$Surv+df_fly_13$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_13$Time, df_fly_13$Surv - df_fly_13$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_13_2$Time[2:9],df_fly_13_2$Surv[2:9], pch = 1, cex = 1.2)
lines(df_fly_13_2$Time,df_fly_13_2$Surv, lty = 2)
arrows(df_fly_13_2$Time, df_fly_13_2$Surv-df_fly_13_2$Std.error, df_fly_13_2$Time, df_fly_13_2$Surv+df_
```

```
## Warning in arrows(df_fly_13_2$Time, df_fly_13_2$Surv -
## df_fly_13_2$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

```
legend('topright', pch = c(16,1),lty = c(1,2),
       legend = c('Old dataset', 'New dataset'), bty = 'n')
```




```
data$Group <- rep(1,length(data$Time))
data <- cbind(data[1:3],data[5:6])
data_2_rep$Group <- rep(2,length(data_2_rep$Time))
data_2_rep <- cbind(data_2_rep[,1:2],data_2_rep[,4:6])
d_OP50 <- data[data$media == 'OP50',]
d_13 <- data[data$media == 'FG13',]
d_20 <- data[data$media == 'FG20',]
d_OP50_2 <- data_2_rep[data_2_rep$media == 'OP50',]
d_13_2 <- data_2_rep[data_2_rep$media == "FG13",]
d_20_2 <- data_2_rep[data_2_rep$media == "FG20",]

fly_13 <- rbind(d_13,d_13_2)
survdiff(Surv(Time,status) ~Group, data = fly_13, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ Group, data = fly_13,
##     rho = 0)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## Group=1 38       30     36.2      1.06      3.47
## Group=2 32       32     25.8      1.49      3.47
##
##  Chisq= 3.5  on 1 degrees of freedom, p= 0.06
```

```
fit_fly_13 <- coxph(Surv(Time,status) ~Group, data = fly_13)
summary(fit_fly_13)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ Group, data = fly_13)
##
##   n= 70, number of events= 62
##
##         coef exp(coef) se(coef)    z Pr(>|z|)
## Group 0.4841    1.6228   0.2576 1.88   0.0602 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##       exp(coef) exp(-coef) lower .95 upper .95
## Group     1.623     0.6162    0.9795     2.688
##
## Concordance= 0.556  (se = 0.039 )
## Rsquare= 0.049   (max possible= 0.998 )
## Likelihood ratio test= 3.52  on 1 df,   p=0.06
## Wald test            = 3.53  on 1 df,   p=0.06
## Score (logrank) test = 3.6  on 1 df,   p=0.06
```

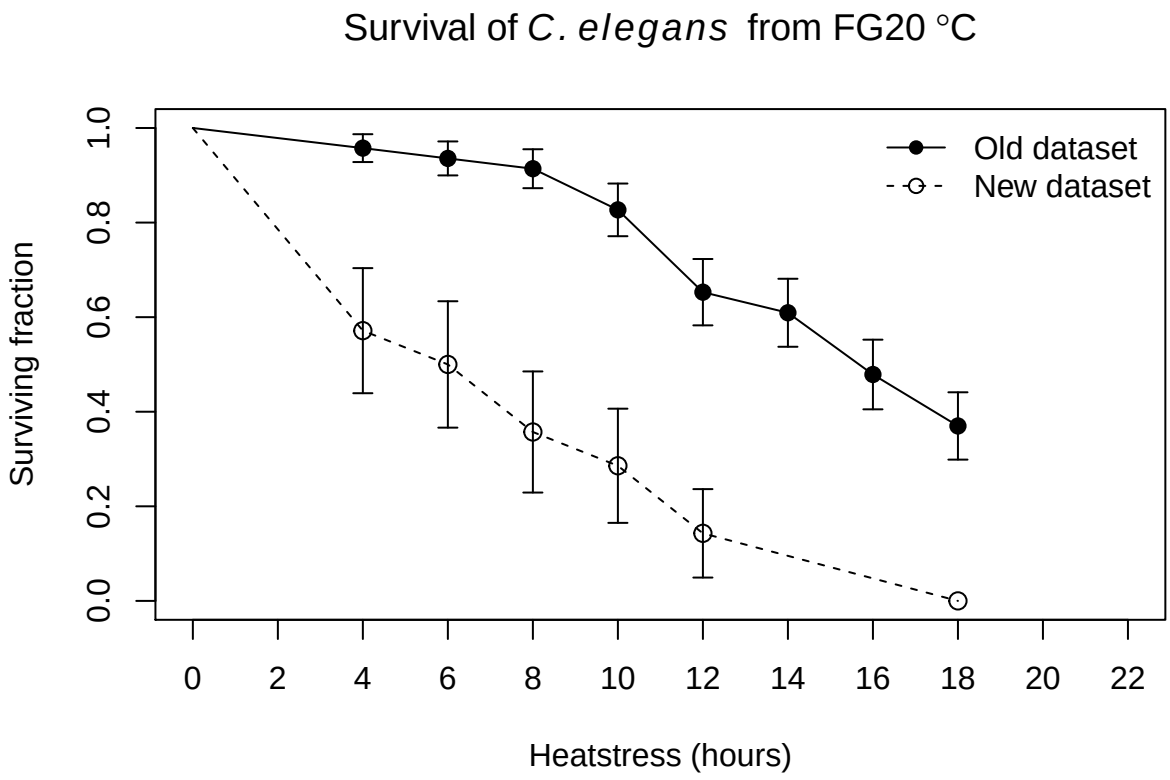
```
plot(df_fly_20$Time[2:9],df_fly_20$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Heatstress (hours)',
     ylab = 'Surviving fraction',
     main = expression('Survival of' ~ italic(C.) ~italic(elegans) ~' from FG20' ~degree*C),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_20$Time,df_fly_20$Surv)
arrows(df_fly_20$Time, df_fly_20$Surv-df_fly_20$Std.error, df_fly_20$Time,
       df_fly_20$Surv+df_fly_20$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_20$Time, df_fly_20$Surv - df_fly_20$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20_2$Time[2:9],df_fly_20_2$Surv[2:9], pch = 1, cex = 1.2)
lines(df_fly_20_2$Time,df_fly_20_2$Surv, lty = 2)
arrows(df_fly_20_2$Time, df_fly_20_2$Surv-df_fly_20_2$Std.error, df_fly_20_2$Time, df_fly_20_2$Surv+df_
```

```
## Warning in arrows(df_fly_20_2$Time, df_fly_20_2$Surv -
## df_fly_20_2$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

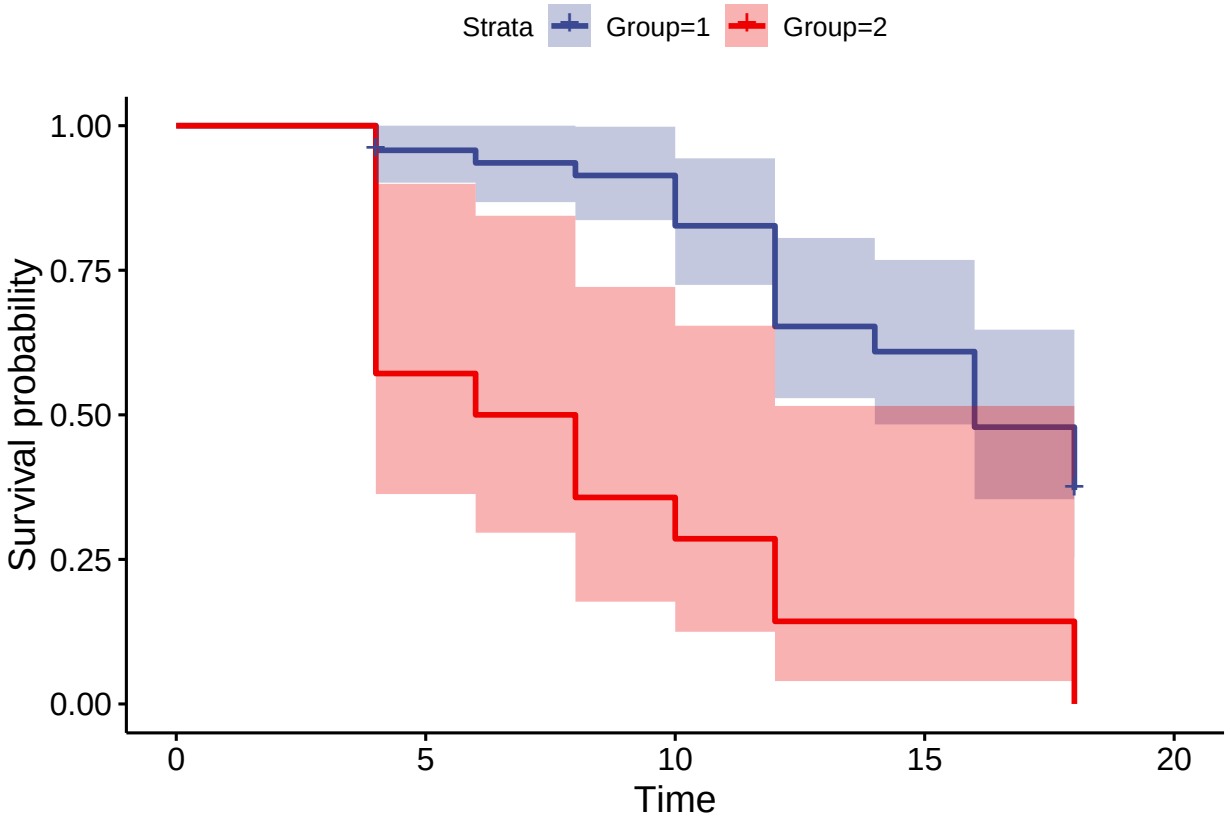
```
legend('topright', pch = c(16,1),lty = c(1,2),
       legend = c('Old dataset', 'New dataset'), bty = 'n')
```




```
fly_20 <- rbind(d_20,d_20_2)
survdiff(Surv(Time,status) ~Group, data = fly_20, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ Group, data = fly_20,
##     rho = 0)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## Group=1 47       29    37.84      2.07      20.7
## Group=2 14       14     5.16     15.17      20.7
##
##  Chisq= 20.7  on 1 degrees of freedom, p= 5e-06
```

```
km_fly_20 <-survfit(Surv(Time,status) ~Group, data = fly_20)
ggsurvplot(km_fly_20, conf.int = TRUE, palette = 'aaas')
```




```
fit_fly_20 <- coxph(Surv(Time,status) ~Group, data = fly_20)
summary(fit_fly_20)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ Group, data = fly_20)
##
##   n= 61, number of events= 43
##
##        coef exp(coef) se(coef)     z Pr(>|z|)
## Group 1.480     4.391    0.336 4.404 1.06e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##       exp(coef) exp(-coef) lower .95 upper .95
## Group     4.391     0.2277     2.273     8.483
##
## Concordance= 0.649  (se = 0.035 )
## Rsquare= 0.229   (max possible= 0.994 )
## Likelihood ratio test= 15.87  on 1 df,   p=7e-05
## Wald test            = 19.4  on 1 df,   p=1e-05
## Score (logrank) test = 22.83  on 1 df,   p=2e-06
```

```
plot(df_OP50$Time[2:9],df_OP50$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Heatstress (hours)',
     ylab = 'Surviving fraction',
     main = expression('Survival of' ~ italic(C.) ~italic(elegans) ~ 'from OP50'),
```

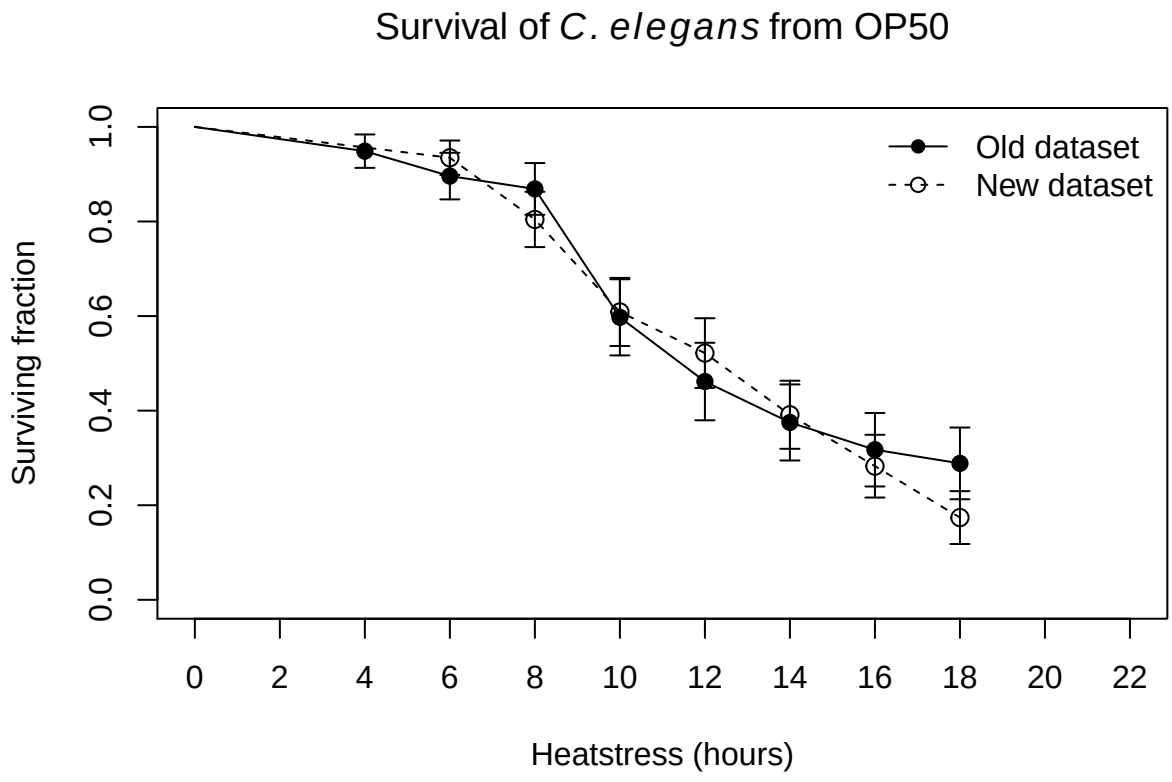
```
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_OP50$Time,df_OP50$Surv)
arrows(df_OP50$Time, df_OP50$Surv-df_OP50$Std.error, df_OP50$Time,
       df_OP50$Surv+df_OP50$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50$Time, df_OP50$Surv - df_OP50$Std.error,
## df_OP50$Time, : zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_OP50_2$Time[2:9],df_OP50_2$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50_2$Time,df_OP50_2$Surv, lty = 2)
arrows(df_OP50_2$Time, df_OP50_2$Surv-df_OP50_2$Std.error, df_OP50_2$Time, df_OP50_2$Surv+df_OP50_2$Std
```

```
## Warning in arrows(df_OP50_2$Time, df_OP50_2$Surv - df_OP50_2$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
legend('topright', pch = c(16,1),lty = c(1,2),
       legend = c('Old dataset', 'New dataset'), bty = 'n')
```




```
OP50 <- rbind(d_OP50,d_OP50_2)
survdiff(Surv(Time,status) ~Group, data = OP50, rho = 0)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ Group, data = OP50, rho = 0)
```

```
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## Group=1 39       26     28.2     0.173     0.381
## Group=2 46       38     35.8     0.136     0.381
##
##  Chisq= 0.4  on 1 degrees of freedom, p= 0.5
```

```
fit_OP50 <- coxph(Surv(Time,status) ~Group, data = OP50)
summary(fit_OP50)
```

```
## Call:
## coxph(formula = Surv(Time, status) ~ Group, data = OP50)
##
##   n= 85, number of events= 64
##
##         coef exp(coef) se(coef)     z Pr(>|z|)
## Group 0.1540    1.1665   0.2548 0.604    0.546
##
##       exp(coef) exp(-coef) lower .95 upper .95
## Group     1.166     0.8573    0.7079     1.922
##
## Concordance= 0.507  (se = 0.037 )
## Rsquare= 0.004   (max possible= 0.997 )
## Likelihood ratio test= 0.37  on 1 df,   p=0.5
## Wald test            = 0.37  on 1 df,   p=0.5
## Score (logrank) test = 0.37  on 1 df,   p=0.5
```

```
d_1350_2 <- rbind(d_OP50_2,d_13_2)
survdiff(Surv(Time,status) ~media, data = d_1350_2)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_1350_2)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 46       38     46.1      1.41      5.49
## media=FG13 32       32     23.9      2.72      5.49
##
##  Chisq= 5.5  on 1 degrees of freedom, p= 0.02
```

```
d_2050_2 <- rbind(d_OP50_2,d_20_2)
survdiff(Surv(Time,status) ~media, data = d_2050_2)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_2050_2)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50 46       38    45.05      1.10      10.3
## media=FG20 14       14     6.95      7.16      10.3
##
##  Chisq= 10.3  on 1 degrees of freedom, p= 0.001
```

```r
d_1320_2 <- rbind(d_13_2,d_20_2)
survdiff(Surv(Time,status) ~media, data = d_1320_2)
```

```
## Call:
## survdiff(formula = Surv(Time, status) ~ media, data = d_1320_2)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 32       32     35.2     0.283      1.92
## media=FG20 14       14     10.8     0.919      1.92
##
##  Chisq= 1.9  on 1 degrees of freedom, p= 0.2
```