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 and mutant.

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

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

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",'Replicate', 'media',
                  'mutant')
head(data)
```

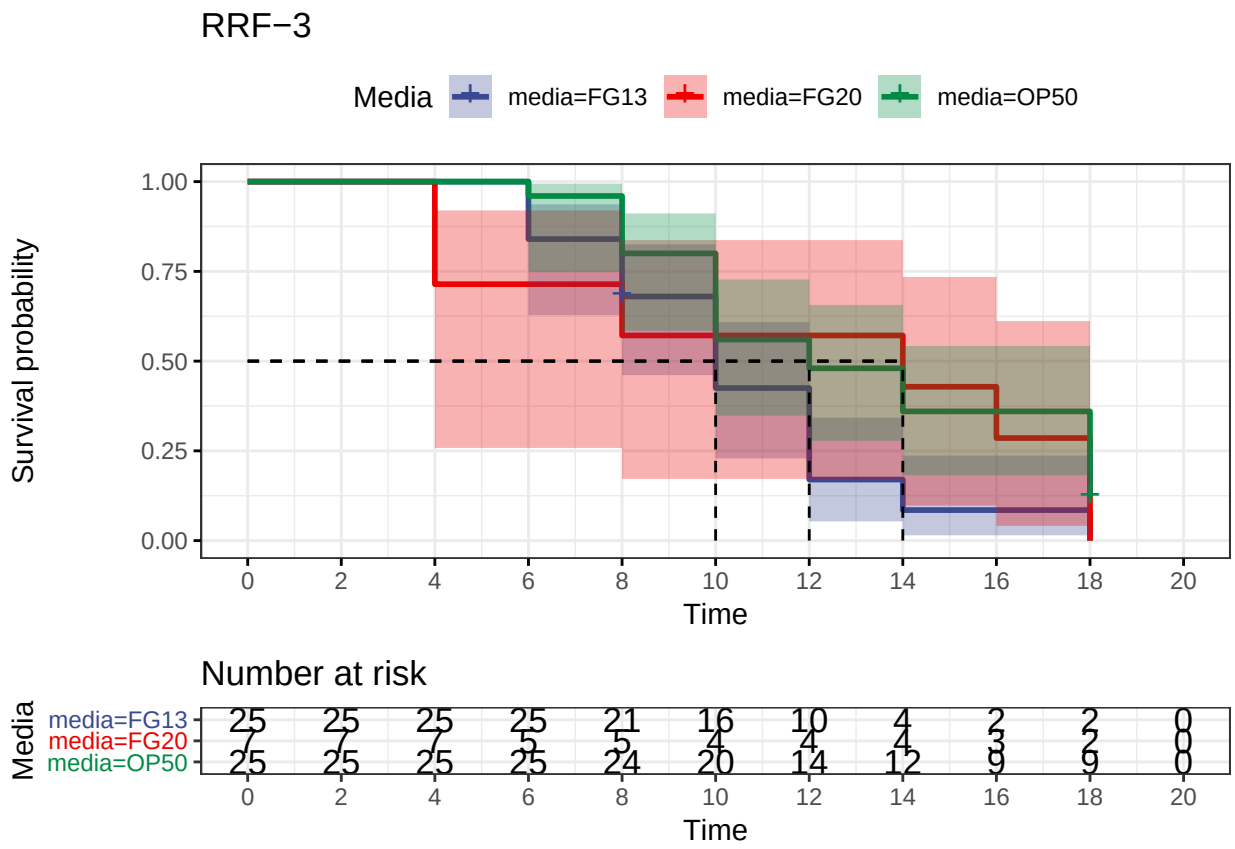
```
##   Time status Replicate media mutant
## 1    4      1        NA  FG13    PMK
## 2    6      1        NA  FG13    PMK
## 3    8      1        NA  FG13    PMK
## 4    8      1        NA  FG13    PMK
## 5   10      1        NA  FG13    PMK
## 6   10      1        NA  FG13    PMK
```

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

First we see for each mutant

```
d_RRF <- subset(data, mutant == 'RRF')
d_PMK <- subset(data,mutant == 'PMK')
d_DAF <- subset(data, mutant == 'DAF')
```

```
d_RRF$Survobj <- with(d_RRF,Surv(d_RRF$Time, event = d_RRF$status))
km_RRF <- survfit(Survobj ~ media, data = d_RRF,conf.type = "log-log", error = "greenwood")
s_km_RRF <- summary(km_RRF)
ggsurvplot(km_RRF,data = d_RRF, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'RRF-3')
```




```
df_fly_13_r <- data.frame(c(0,s_km_RRF$time[1:6]),c(1,s_km_RRF$surv[1:6]),                    c(0,s_
colnames(df_fly_13_r) <- c('Time','Surv','Std.error')
df_fly_20_r <- data.frame(c(0,s_km_RRF$time[7:11])
                          ,c(1,s_km_RRF$surv[7:11]),    c(0,s_km_RRF$std.err[7:11]))
colnames(df_fly_20_r) <- c('Time','Surv','Std.error')
df_OP50_r <- data.frame(c(0,s_km_RRF$time[12:17])
                        ,c(1,s_km_RRF$surv[12:17]),
                          c(0,s_km_RRF$std.err[12:17]))
colnames(df_OP50_r) <- c('Time','Surv','Std.error')

plot(df_fly_13_r$Time[2:7],df_fly_13_r$Surv[2:7], pch = 16, cex = 1.2, xlab = 'Time (hours)',
     ylab = 'Surviving fraction',
     main = expression('Heat knockdown survival of'~italic(C.) ~italic(elegans) ~'strain' ~ italic(rrf)=
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13_r$Time,df_fly_13_r$Surv)
arrows(df_fly_13_r$Time, df_fly_13_r$Surv-df_fly_13_r$Std.error, df_fly_13_r$Time,
       df_fly_13_r$Surv+df_fly_13_r$Std.error, length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_fly_13_r$Time, df_fly_13_r$Surv -
## df_fly_13_r$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

```
points(df_OP50_r$Time[2:9],df_OP50_r$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50_r$Time,df_OP50_r$Surv, lty = 2)
arrows(df_OP50_r$Time, df_OP50_r$Surv-df_OP50_r$Std.error,
```

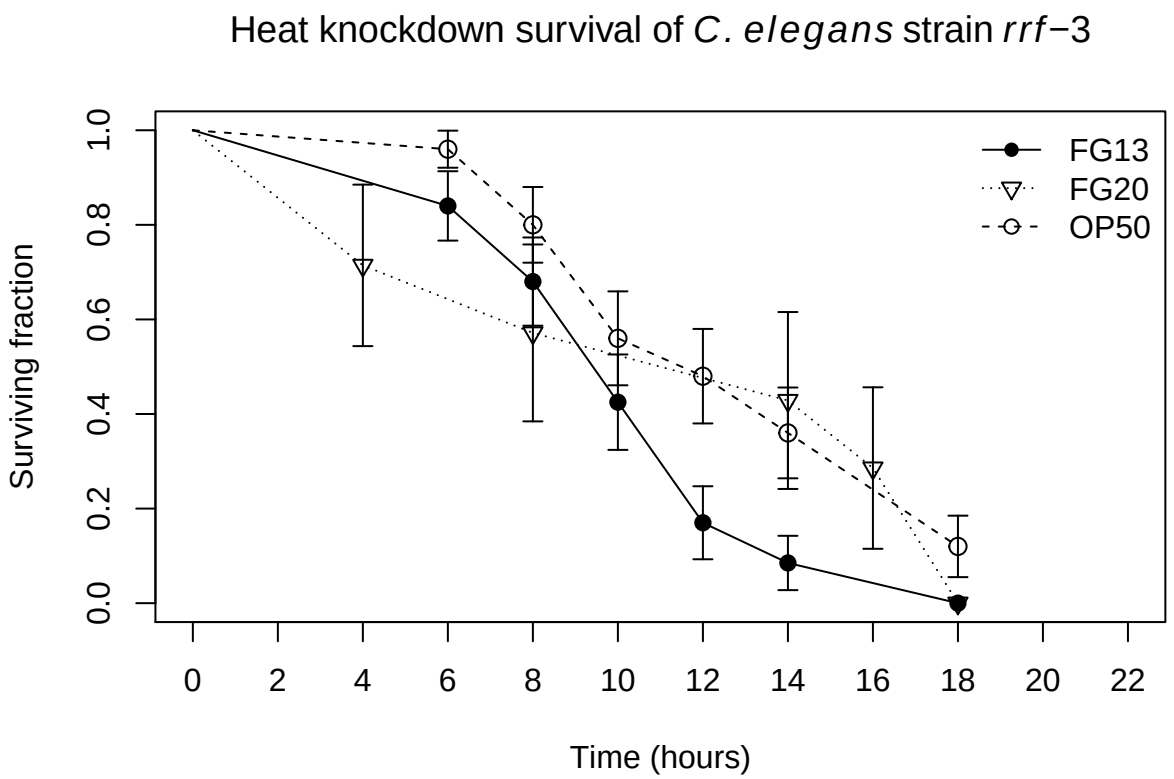
```
      df_OP50_r$Time, df_OP50_r$Surv+df_OP50_r$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50_r$Time, df_OP50_r$Surv - df_OP50_r$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

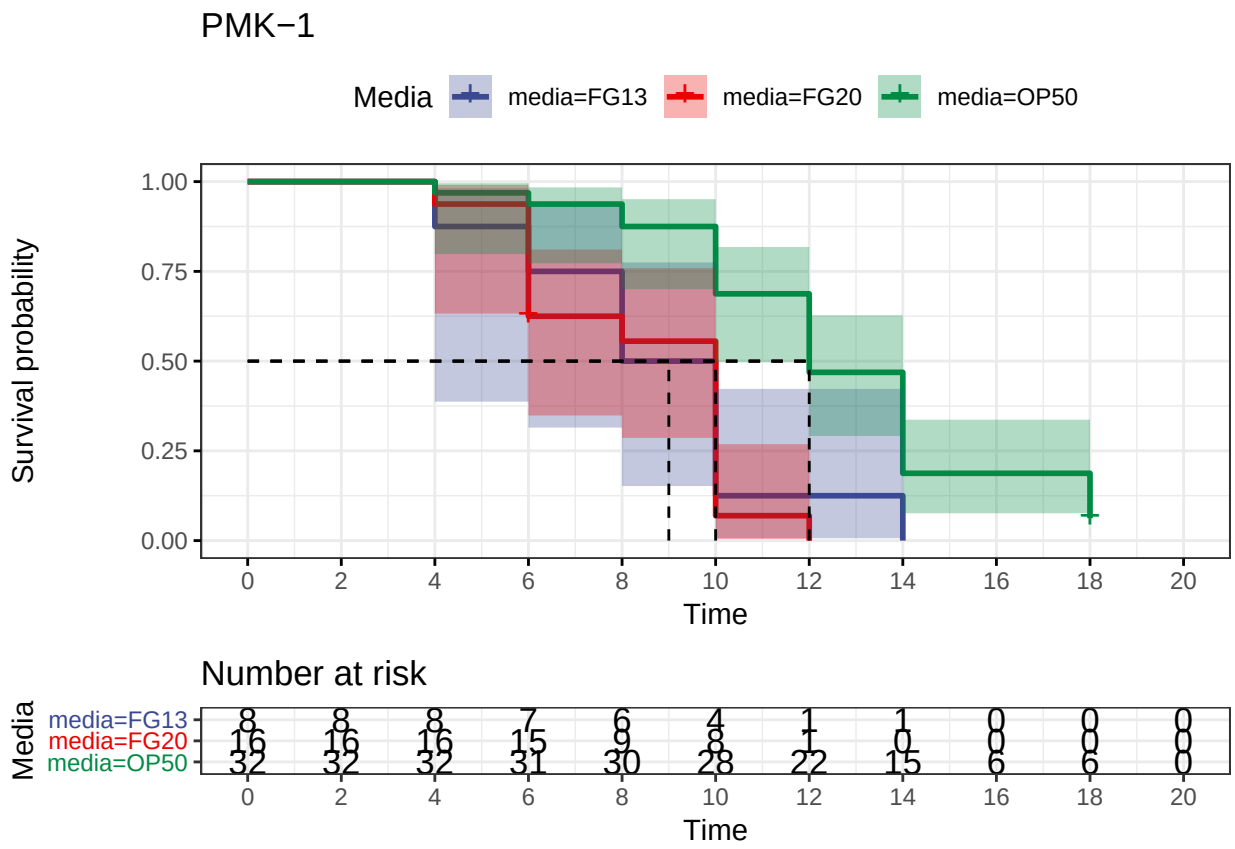
```
points(df_fly_20_r$Time[2:9],df_fly_20_r$Surv[2:9], pch = 6)
lines(df_fly_20_r$Time,df_fly_20_r$Surv,lty = 3)
arrows(df_fly_20_r$Time, df_fly_20_r$Surv-df_fly_20_r$Std.error, df_fly_20_r$Time,df_fly_20_r$Surv+df_fl
```

```
## Warning in arrows(df_fly_20_r$Time, df_fly_20_r$Surv -
## df_fly_20_r$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')
```




```
d_PMK$Survobj <- with(d_PMK,Surv(d_PMK$Time, event = d_PMK$status))
km_PMK <- survfit(Survobj ~ media, data = d_PMK,conf.type = "log-log", error = "greenwood")
s_km_PMK <- summary(km_PMK)
ggsurvplot(km_PMK,data = d_PMK, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'PMK-1')
```




```
df_fly_13_P <- data.frame(c(0,s_km_PMK$time[1:5]),
                          c(1,s_km_PMK$surv[1:5]),
                          c(0,s_km_PMK$std.err[1:5]))
colnames(df_fly_13_P) <- c('Time','Surv','Std.error')
df_fly_20_P <- data.frame(c(0,s_km_PMK$time[6:10]),
                          c(1,s_km_PMK$surv[6:10]),
                          c(0,s_km_PMK$std.err[6:10]))
colnames(df_fly_20_P) <- c('Time','Surv','Std.error')
df_OP50_P <- data.frame(c(0,s_km_PMK$time[11:17]),
                        c(1,s_km_PMK$surv[11:17]),
                          c(0,s_km_PMK$std.err[11:17]))
colnames(df_OP50_P) <- c('Time','Surv','Std.error')

plot(df_fly_13_P$Time[2:9],df_fly_13_P$Surv[2:9], pch = 16, cex = 1.2,
     xlab = 'Time (hours)', ylab = 'Surviving fraction',
     main = expression('Heat knockdown survival of'~italic(C.) ~ italic(elegans) ~'strain' ~ italic(rrf
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13_P$Time,df_fly_13_P$Surv)
arrows(df_fly_13_P$Time, df_fly_13_P$Surv-df_fly_13_P$Std.error, df_fly_13_P$Time,df_fly_13_P$Surv+df_fl
```

```
## Warning in arrows(df_fly_13_P$Time, df_fly_13_P$Surv -
## df_fly_13_P$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

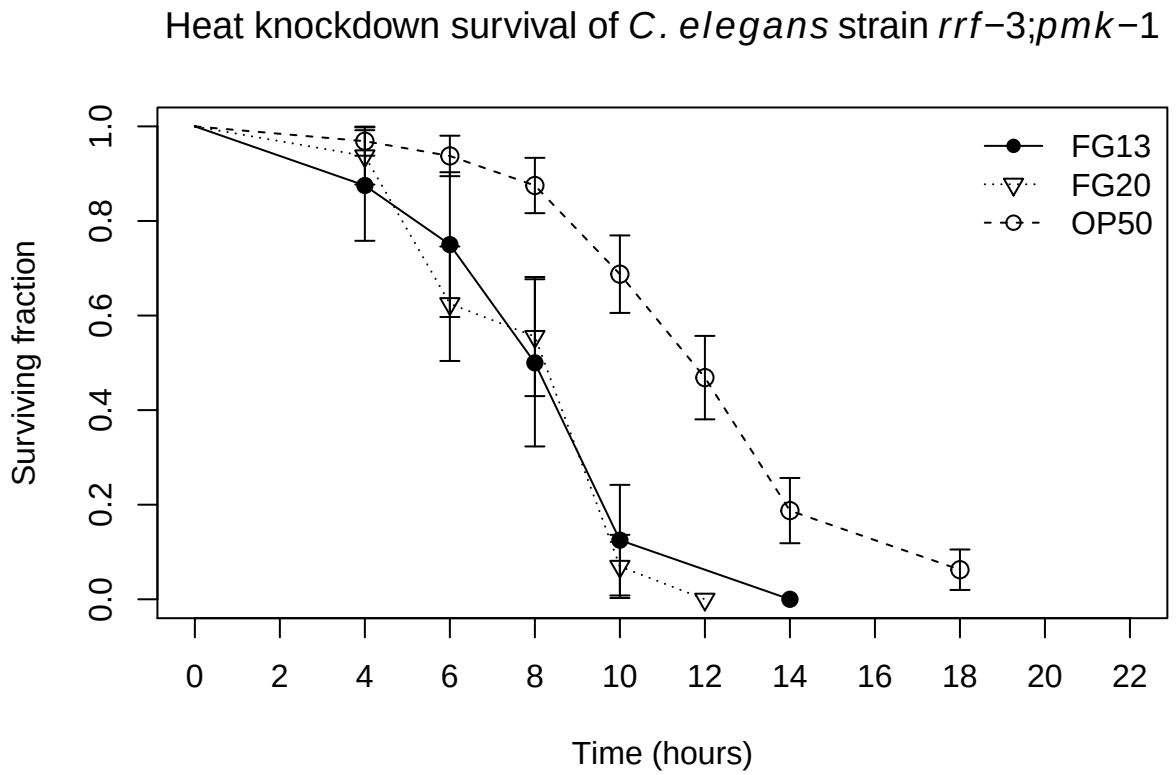
```
points(df_OP50_P$Time[2:9],df_OP50_P$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50_P$Time,df_OP50_P$Surv, lty = 2)
arrows(df_OP50_P$Time, df_OP50_P$Surv-df_OP50_P$Std.error,
      df_OP50_P$Time, df_OP50_P$Surv+df_OP50_P$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50_P$Time, df_OP50_P$Surv - df_OP50_P$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20_P$Time[2:9],df_fly_20_P$Surv[2:9], pch = 6)
lines(df_fly_20_P$Time,df_fly_20_P$Surv,lty = 3)
arrows(df_fly_20_P$Time, df_fly_20_P$Surv-df_fly_20_P$Std.error, df_fly_20_P$Time,df_fly_20_P$Surv+df_fl
```

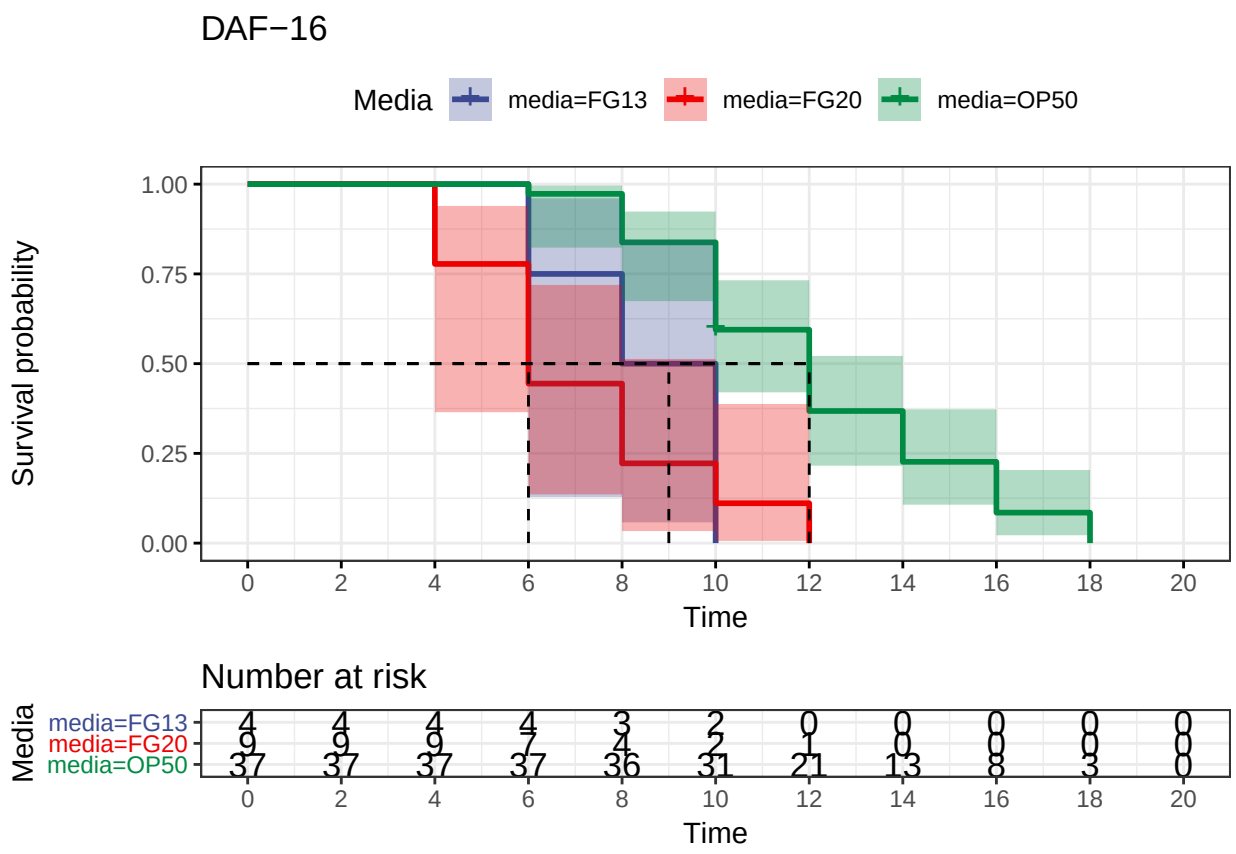
```
## Warning in arrows(df_fly_20_P$Time, df_fly_20_P$Surv -
## df_fly_20_P$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

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




```
d_DAF$Survobj <- with(d_DAF,Surv(d_DAF$Time, event = d_DAF$status))
km_DAF <- survfit(Survobj ~ media, data = d_DAF,conf.type = "log-log", error = "greenwood")
s_km_DAF <- summary(km_DAF)
```

```
ggsurvplot(km_DAF,data = d_DAF, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'DAF-16')
```




```
df_fly_13_d <- data.frame(c(0,s_km_DAF$time[1:3]),
                          c(1,s_km_DAF$surv[1:3]),
                          c(0,s_km_DAF$std.err[1:3]))
colnames(df_fly_13_d) <- c('Time','Surv','Std.error')
df_fly_20_d <- data.frame(c(0,s_km_DAF$time[4:8]),
                          c(1,s_km_DAF$surv[4:8]),
                          c(0,s_km_DAF$std.err[4:8]))
colnames(df_fly_20_d) <- c('Time','Surv','Std.error')
df_OP50_d <- data.frame(c(0,s_km_DAF$time[9:15]),
                        c(1,s_km_DAF$surv[9:15]),
                          c(0,s_km_DAF$std.err[9:15]))
colnames(df_OP50_d) <- c('Time','Surv','Std.error')

plot(df_fly_13_d$Time[2:9],df_fly_13_d$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Time (hours)',
     ylab = 'Surviving fraction',
     main = expression('Heat knockdown survival of'~italic(C.) ~italic(elegans) ~'strain' ~ italic(daf)
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_fly_13_d$Time,df_fly_13_d$Surv)
arrows(df_fly_13_d$Time, df_fly_13_d$Surv-df_fly_13_d$Std.error, df_fly_13_d$Time,df_fly_13_d$Surv+df_fl
```

```
## Warning in arrows(df_fly_13_d$Time, df_fly_13_d$Surv -
## df_fly_13_d$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

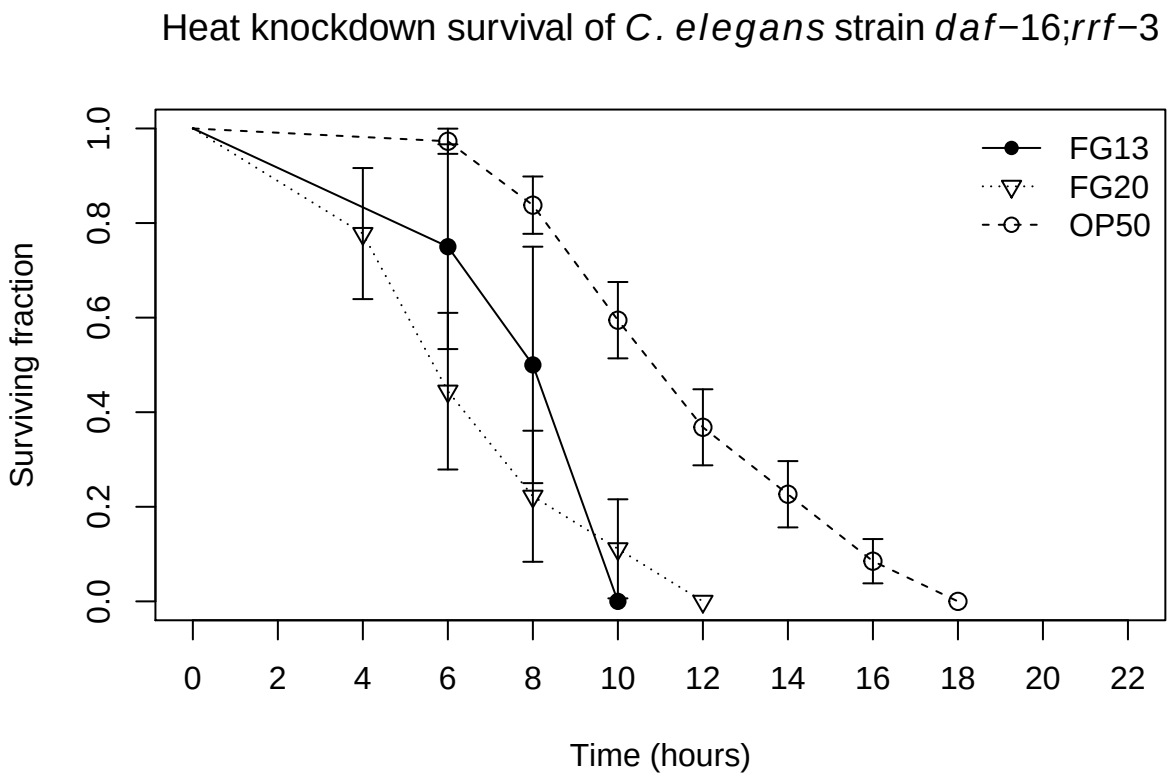
```
points(df_OP50_d$Time[2:9],df_OP50_d$Surv[2:9], pch = 1, cex = 1.2)
lines(df_OP50_d$Time,df_OP50_d$Surv, lty = 2)
arrows(df_OP50_d$Time, df_OP50_d$Surv-df_OP50_d$Std.error,
      df_OP50_d$Time, df_OP50_d$Surv+df_OP50_d$Std.error,
      length=0.05, angle=90, code=3, col = 'black')
```

```
## Warning in arrows(df_OP50_d$Time, df_OP50_d$Surv - df_OP50_d$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_fly_20_d$Time[2:9],df_fly_20_d$Surv[2:9], pch = 6)
lines(df_fly_20_d$Time,df_fly_20_d$Surv,lty = 3)
arrows(df_fly_20_d$Time, df_fly_20_d$Surv-df_fly_20_d$Std.error, df_fly_20_d$Time,df_fly_20_d$Surv+df_fl
```

```
## Warning in arrows(df_fly_20_d$Time, df_fly_20_d$Surv -
## df_fly_20_d$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

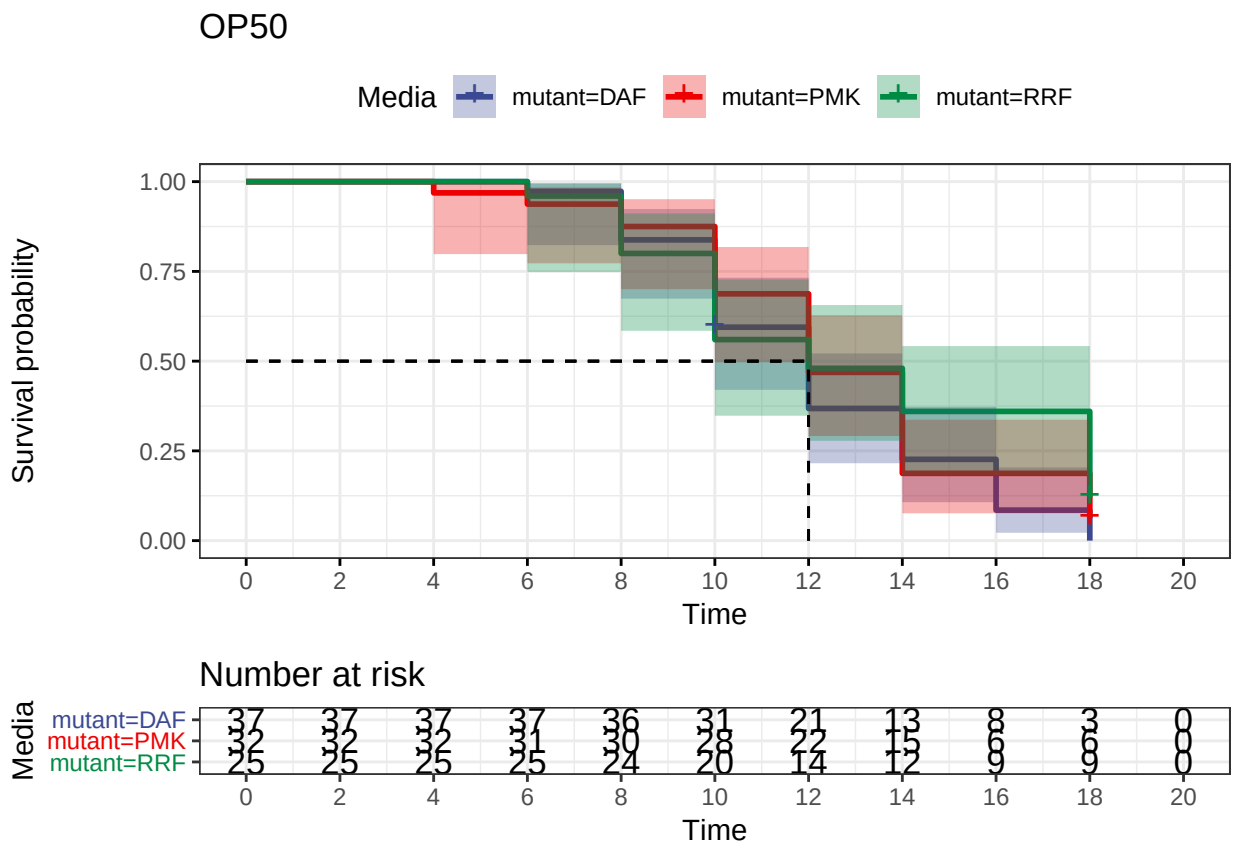
```
legend('topright', pch = c(16,6,1),lty = c(1,3,2),
       legend = c('FG13','FG20','OP50'), bty = 'n')
```



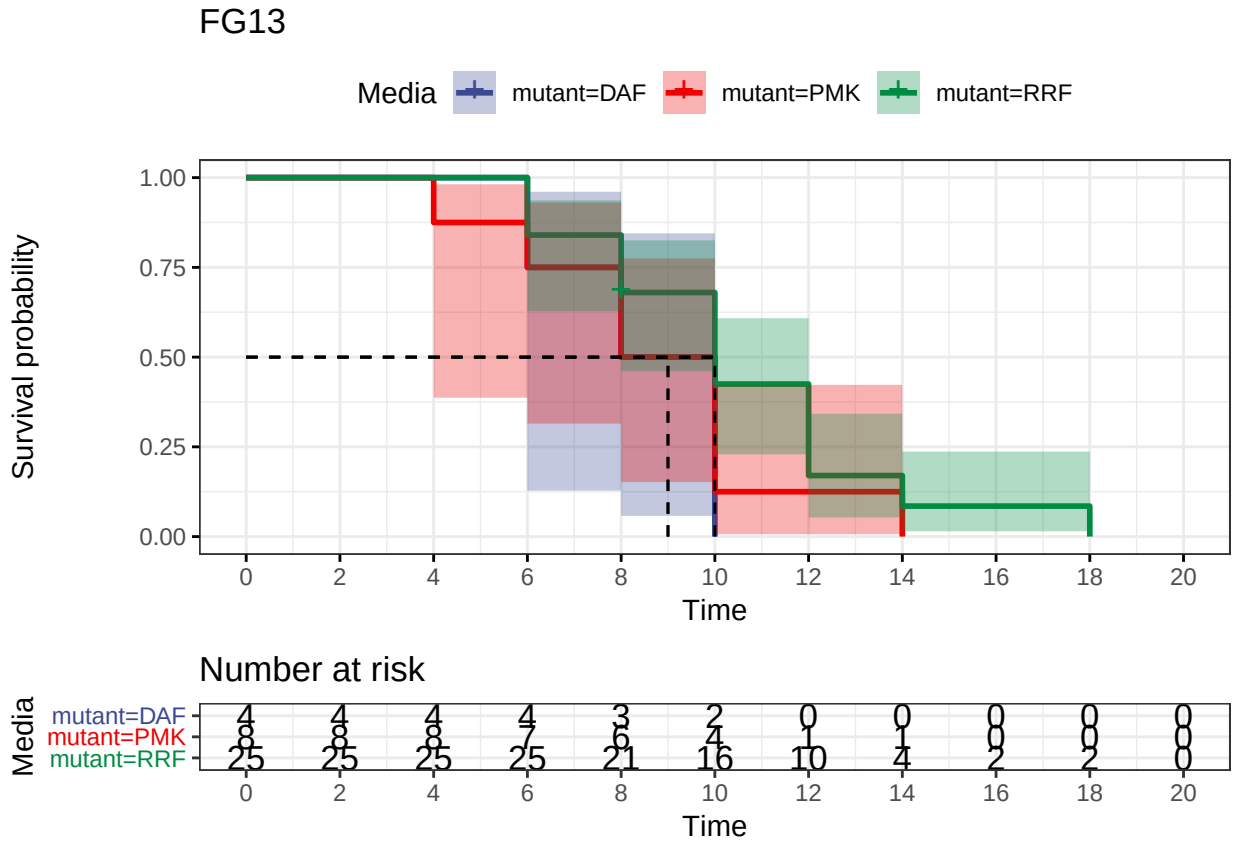
Then we compare by medium

```
data_OP50 <- subset(data, media == 'OP50')
data_FG13 <- subset(data, media == 'FG13')
data_FG20 <- subset(data, media == 'FG20')
```

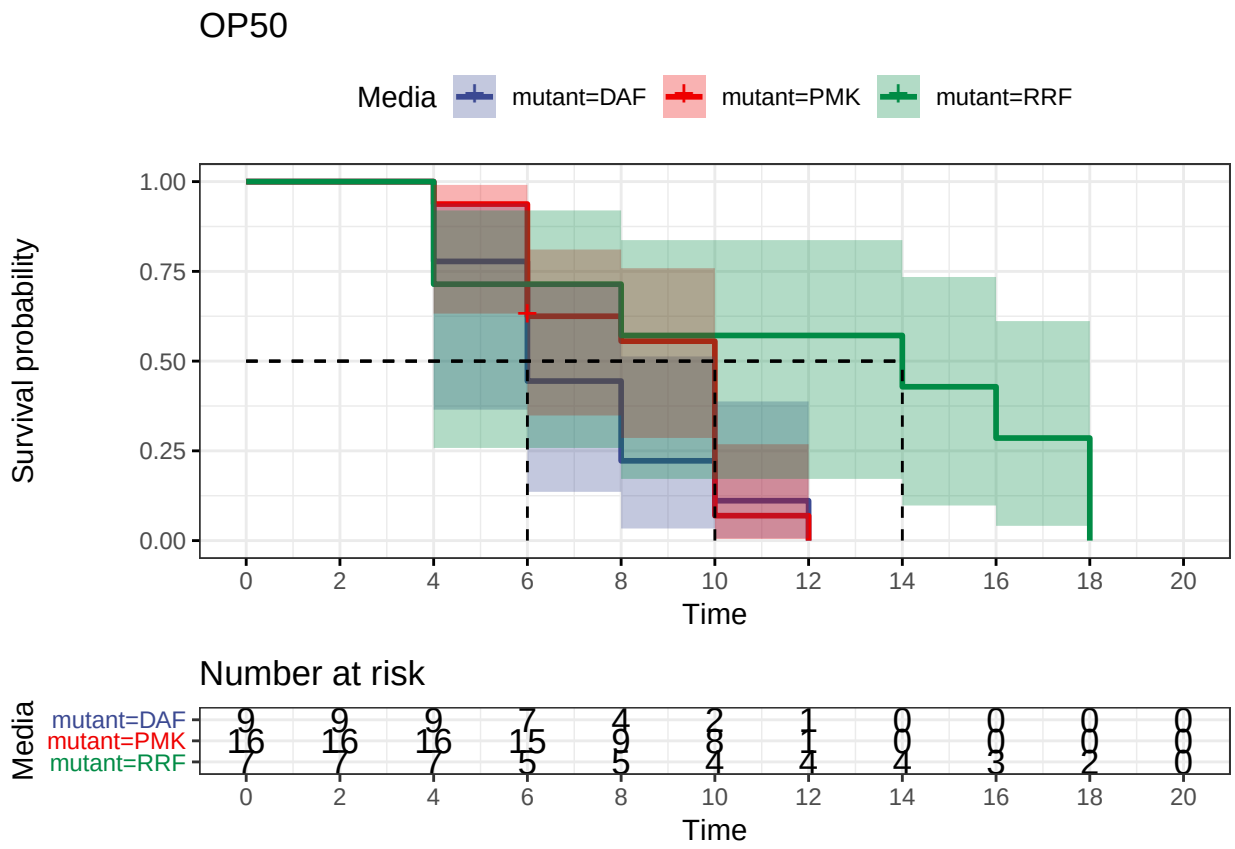
```
data_OP50$Survobj <- with(data_OP50,Surv(data_OP50$Time, event = data_OP50$status))
km_OP50 <- survfit(Survobj ~ mutant, data = data_OP50,conf.type = "log-log", error = "greenwood")
s_km_OP50 <- summary(km_OP50)
ggsurvplot(km_OP50,data = data_OP50, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'OP50')
```




```
data_FG13$Survobj <- with(data_FG13,Surv(data_FG13$Time, event = data_FG13$status))
km_FG13 <- survfit(Survobj ~ mutant, data = data_FG13,conf.type = "log-log", error = "greenwood")
s_km_FG13 <- summary(km_FG13)
ggsurvplot(km_FG13,data = data_FG13, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'FG13')
```




```
data_FG20$Survobj <- with(data_FG20,Surv(data_FG20$Time, event = data_FG20$status))
km_FG20 <- survfit(Survobj ~ mutant, data = data_FG20,conf.type = "log-log", error = "greenwood")
s_km_FG20 <- summary(km_FG20)
ggsurvplot(km_FG20,data = data_FG20, conf.int = TRUE,
           ggtheme = theme_bw(),risk.table = 0.25,
           palette ='aaas', surv.median.line = 'hv',
           legend.title = 'Media',break.x.by = 2,
           title = 'OP50')
```



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.

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

```
data$Survobj <- with(data,Surv(data$Time, event = data$status))
data$mutant <-factor(data$mutant, c('RRF','PMK','DAF'))
data$media <-factor(data$media, c('OP50','FG13','FG20'))
survdiff(Survobj ~media + mutant, data = data, rho = 0)
```

```
## Call:
## survdiff(formula = Survobj ~ media + mutant, data = data, rho = 0)
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50, mutant=RRF  25       22    33.34     3.855     7.900
## media=OP50, mutant=PMK  32       30    39.06     2.100     4.263
## media=OP50, mutant=DAF  37       36    38.84     0.208     0.404
## media=FG13, mutant=RRF  25       24    19.78     0.901     1.483
## media=FG13, mutant=PMK   8        8     4.22     3.384     4.615
## media=FG13, mutant=DAF   4        4     1.78     2.762     3.545
## media=FG20, mutant=RRF   7        7     8.29     0.200     0.341
## media=FG20, mutant=PMK  16       15     6.88     9.595    13.238
## media=FG20, mutant=DAF   9        9     2.82    13.539    17.336
##
##  Chisq= 51.7  on 8 degrees of freedom, p= 2e-08
```

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

```
## Call:
## survdiff(formula = Survobj ~ media + mutant, data = data, rho = 1)
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## media=OP50, mutant=RRF 25    11.37    17.04    1.8896    4.3076
## media=OP50, mutant=PMK 32    14.58    22.00    2.5059    6.0333
## media=OP50, mutant=DAF 37    18.26    23.74    1.2635    2.9998
## media=FG13, mutant=RRF 25    15.24    12.93    0.4151    0.8063
## media=FG13, mutant=PMK  8     6.12     3.18    2.7304    4.3317
## media=FG13, mutant=DAF  4     3.24     1.50    2.0432    2.9683
## media=FG20, mutant=RRF  7     3.62     3.99    0.0348    0.0719
## media=FG20, mutant=PMK 16    12.12     5.63    7.4975   11.8513
## media=FG20, mutant=DAF  9     7.76     2.32   12.7917   18.3933
##
##  Chisq= 48  on 8 degrees of freedom, p= 1e-07
```

```
fit_both<- coxph(Survobj ~media + mutant, data = data)
summary(fit_both)
```

```
## Call:
## coxph(formula = Survobj ~ media + mutant, data = data)
##
##   n= 163, number of events= 155
##
##              coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13 1.1472    3.1494   0.2316 4.953 7.32e-07 ***
## mediaFG20 1.1531    3.1681   0.2234 5.161 2.45e-07 ***
## mutantPMK 0.6167    1.8529   0.2121 2.907 0.003644 **
## mutantDAF 0.8707    2.3885   0.2297 3.790 0.000151 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13     3.149     0.3175     2.000     4.959
## mediaFG20     3.168     0.3157     2.045     4.909
## mutantPMK     1.853     0.5397     1.223     2.808
## mutantDAF     2.388     0.4187     1.523     3.747
##
## Concordance= 0.67  (se = 0.028 )
## Rsquare= 0.211   (max possible= 1 )
## Likelihood ratio test= 38.69  on 4 df,   p=8e-08
## Wald test            = 38.5  on 4 df,   p=9e-08
## Score (logrank) test = 39.91  on 4 df,   p=5e-08
```

```
AIC(fit_both)
```

```
## [1] 1273.218
```

```r
fit_int <- coxph(Survobj ~media + mutant + mutant:media, data = data)
AIC(fit_int)
```

```
## [1] 1274.41
```

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

```
## Call:
## coxph(formula = Survobj ~ media, data = data)
##
##   n= 163, number of events= 155
##
##              coef exp(coef) se(coef)     z Pr(>|z|)
## mediaFG13 0.7195    2.0535   0.2019 3.564 0.000365 ***
## mediaFG20 0.9166    2.5008   0.2114 4.335 1.46e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaFG13     2.053     0.4870     1.382     3.050
## mediaFG20     2.501     0.3999     1.652     3.785
##
## Concordance= 0.649  (se = 0.025 )
## Rsquare= 0.131   (max possible= 1 )
## Likelihood ratio test= 22.98  on 2 df,   p=1e-05
## Wald test            = 24.52  on 2 df,   p=5e-06
## Score (logrank) test = 25.84  on 2 df,   p=2e-06
```

```r
AIC(fit_med)
```

```
## [1] 1284.932
```

```r
fit_mut <- coxph(Survobj ~mutant, data = data)
summary(fit_mut)
```

```
## Call:
## coxph(formula = Survobj ~ mutant, data = data)
##
##   n= 163, number of events= 155
##
##              coef exp(coef) se(coef)     z Pr(>|z|)
## mutantPMK 0.2473    1.2806   0.1955 1.265    0.206
## mutantDAF 0.3310    1.3923   0.2014 1.643    0.100
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mutantPMK     1.281     0.7809    0.8730     1.879
## mutantDAF     1.392     0.7182    0.9382     2.066
##
## Concordance= 0.526  (se = 0.029 )
## Rsquare= 0.018   (max possible= 1 )
```

```
## Likelihood ratio test= 3.03  on 2 df,   p=0.2
## Wald test            = 2.97  on 2 df,   p=0.2
## Score (logrank) test = 2.99  on 2 df,   p=0.2
```

```
AIC(fit_mut)
```

```
## [1] 1304.882
```

```
data$mutant <-factor(data$mutant, c('DAF','PMK','RRF'))
data$media <-factor(data$media, c('FG13','OP50','FG20'))
fit_both<- coxph(Survobj ~media + mutant, data = data)
summary(fit_both)
```

```
## Call:
## coxph(formula = Survobj ~ media + mutant, data = data)
##
##   n= 163, number of events= 155
##
##                coef exp(coef)  se(coef)      z Pr(>|z|)
## mediaOP50 -1.147199  0.317525  0.231637 -4.953 7.32e-07 ***
## mediaFG20  0.005922  1.005940  0.250871  0.024 0.981166
## mutantPMK -0.253909  0.775762  0.202013 -1.257 0.208791
## mutantRRF -0.870655  0.418677  0.229717 -3.790 0.000151 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaOP50    0.3175     3.1494    0.2017    0.5000
## mediaFG20    1.0059     0.9941    0.6152    1.6448
## mutantPMK    0.7758     1.2891    0.5221    1.1526
## mutantRRF    0.4187     2.3885    0.2669    0.6568
##
## Concordance= 0.67  (se = 0.028 )
## Rsquare= 0.211   (max possible= 1 )
## Likelihood ratio test= 38.69  on 4 df,   p=8e-08
## Wald test            = 38.5  on 4 df,   p=9e-08
## Score (logrank) test = 39.91  on 4 df,   p=5e-08
```

```
AIC(fit_both)
```

```
## [1] 1273.218
```

```
fit_int <- coxph(Survobj ~media + mutant + mutant:media, data = data)
AIC(fit_int)
```

```
## [1] 1274.41
```

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

```
## Call:
## coxph(formula = Survobj ~ media, data = data)
##
##   n= 163, number of events= 155
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## mediaOP50 -0.7195    0.4870   0.2019 -3.564 0.000365 ***
## mediaFG20  0.1971    1.2178   0.2456  0.802 0.422369
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mediaOP50     0.487     2.0535    0.3278    0.7234
## mediaFG20     1.218     0.8211    0.7525    1.9709
##
## Concordance= 0.649  (se = 0.025 )
## Rsquare= 0.131   (max possible= 1 )
## Likelihood ratio test= 22.98  on 2 df,   p=1e-05
## Wald test            = 24.52  on 2 df,   p=5e-06
## Score (logrank) test = 25.84  on 2 df,   p=2e-06
```

```
AIC(fit_med)
```

```
## [1] 1284.932
```

```
fit_mut <- coxph(Survobj ~mutant, data = data)
summary(fit_mut)
```

```
## Call:
## coxph(formula = Survobj ~ mutant, data = data)
##
##   n= 163, number of events= 155
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## mutantPMK -0.08361   0.91979  0.19910 -0.420    0.675
## mutantRRF -0.33095   0.71824  0.20139 -1.643    0.100
##
##           exp(coef) exp(-coef) lower .95 upper .95
## mutantPMK    0.9198      1.087    0.6226     1.359
## mutantRRF    0.7182      1.392    0.4840     1.066
##
## Concordance= 0.526  (se = 0.029 )
## Rsquare= 0.018   (max possible= 1 )
## Likelihood ratio test= 3.03  on 2 df,   p=0.2
## Wald test            = 2.97  on 2 df,   p=0.2
## Score (logrank) test = 2.99  on 2 df,   p=0.2
```

```
AIC(fit_mut)
```

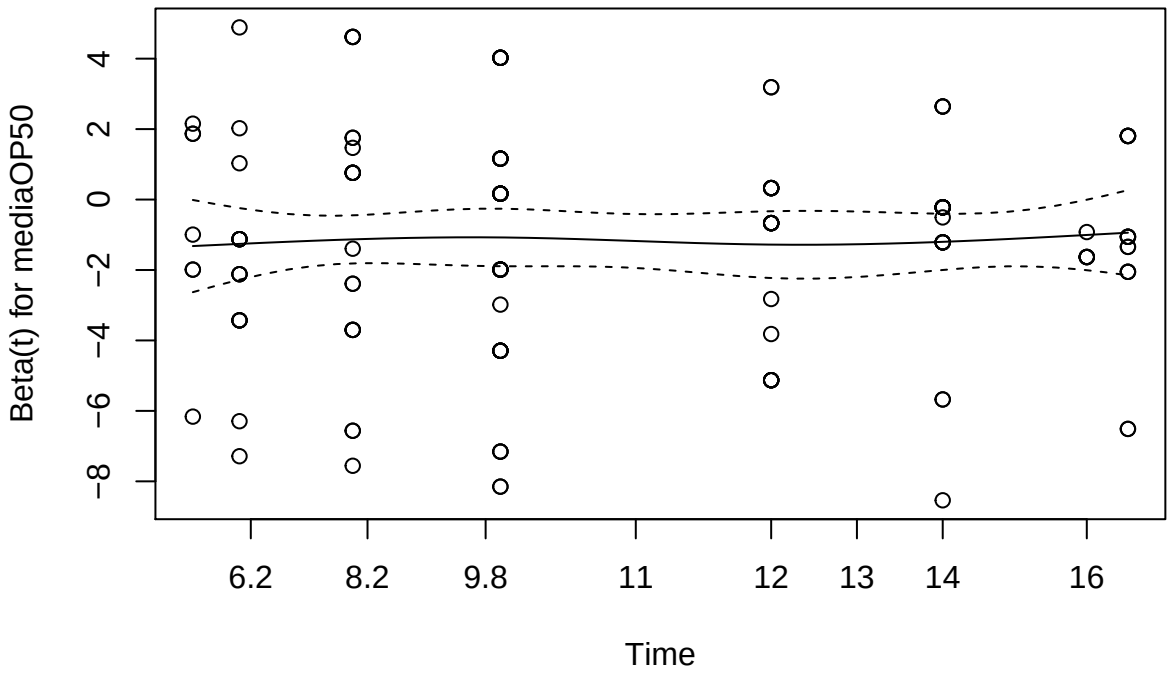
```
## [1] 1304.882
```

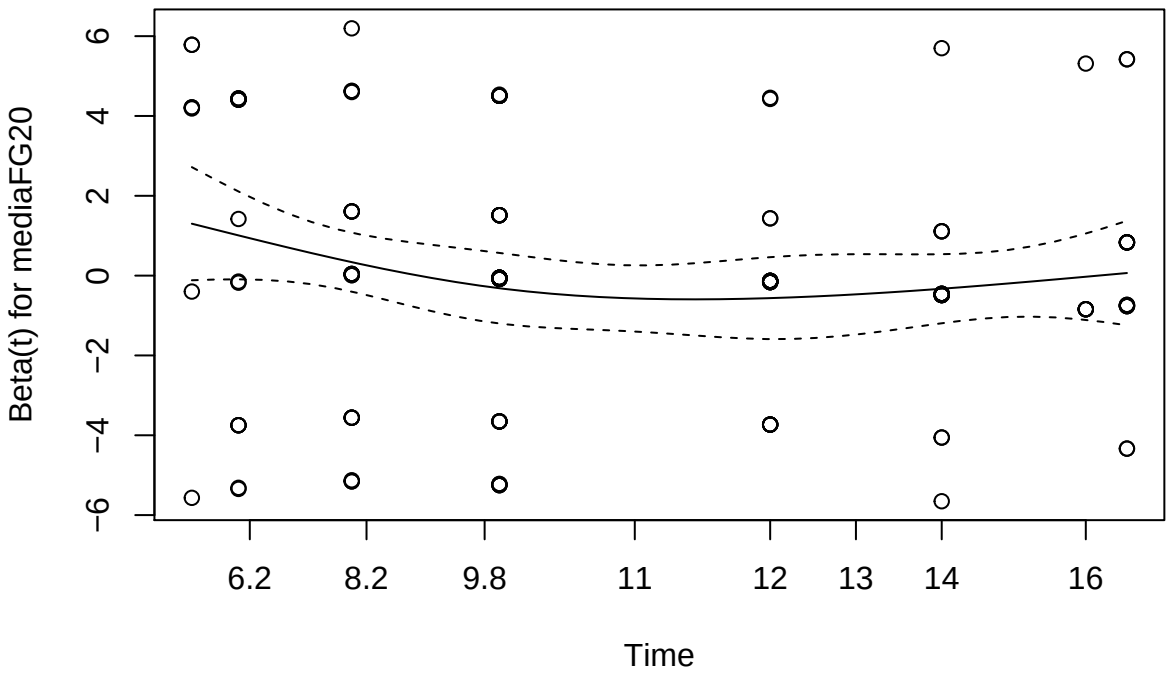
```
AICtable <- bbmle::AICtab(fit_mut, fit_med, fit_both,fit_int, weights = TRUE, sort = FALSE,base = TRUE,
AICtable
```

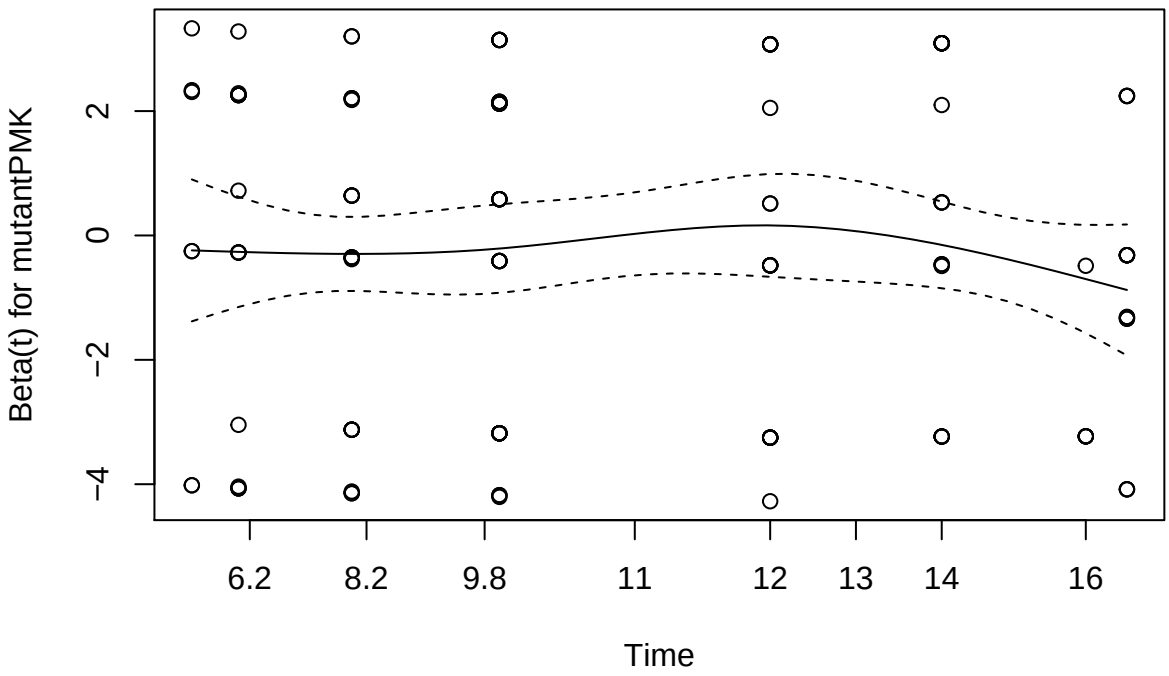
```
##          logLik AIC    dLogLik dAIC   df weight
## fit_mut  -650.4 1304.9    0.0    31.7 2  <0.001
## fit_med  -640.5 1284.9   10.0    11.7 2  0.0018
## fit_both -632.6 1273.2   17.8     0.0 4  0.6436
## fit_int  -629.2 1274.4   21.2     1.2 8  0.3546
```

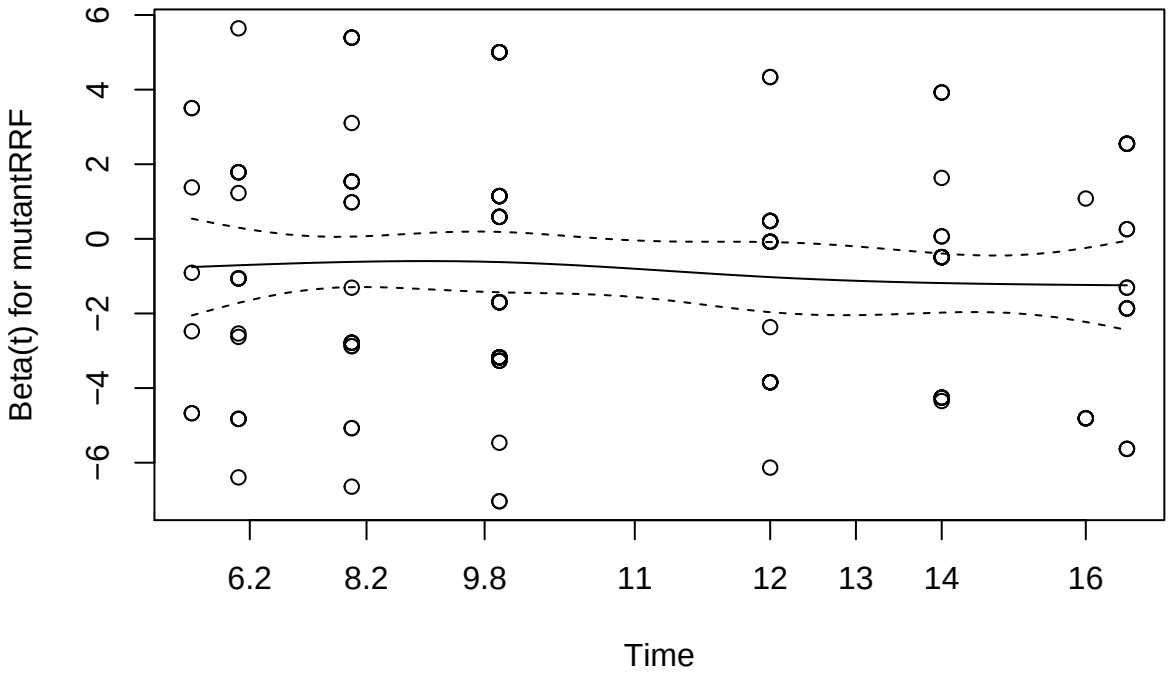
Best model must include both media and mutant strain but not necessarily an interaction between them. This could be driven by few observations for FG20 in RRF-3 Then we do a check on the proportional hazard assumption

```
prop_both <- cox.zph(fit_both)
prop_int <- cox.zph(fit_int)
plot(prop_both)
```

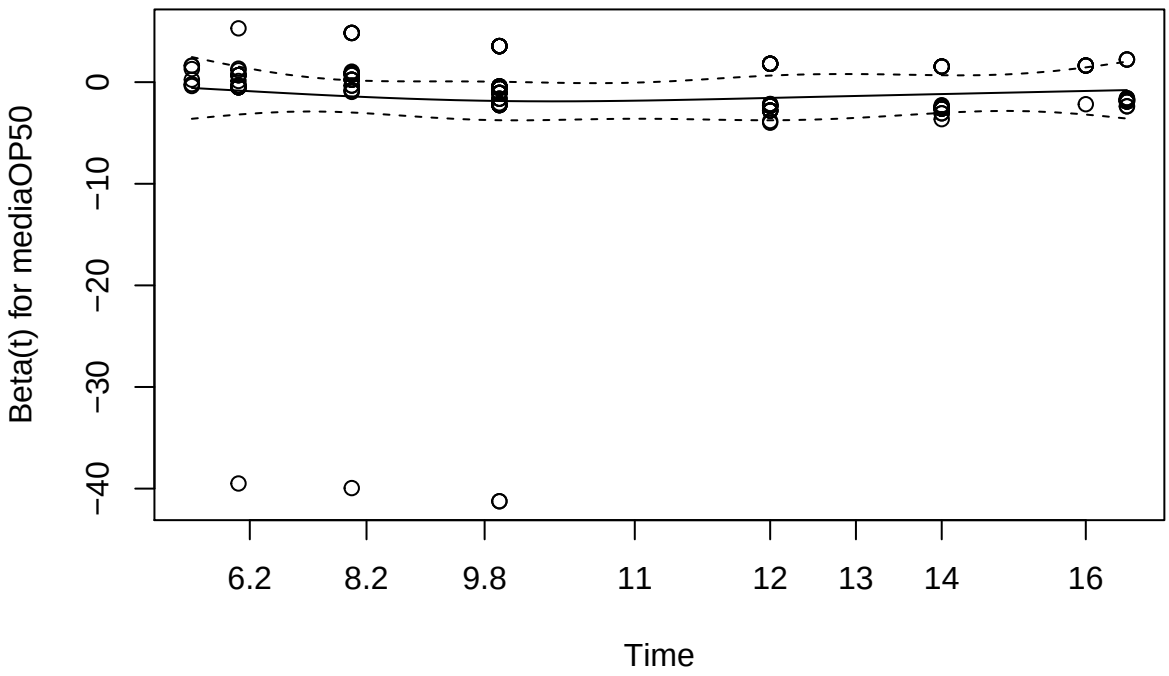


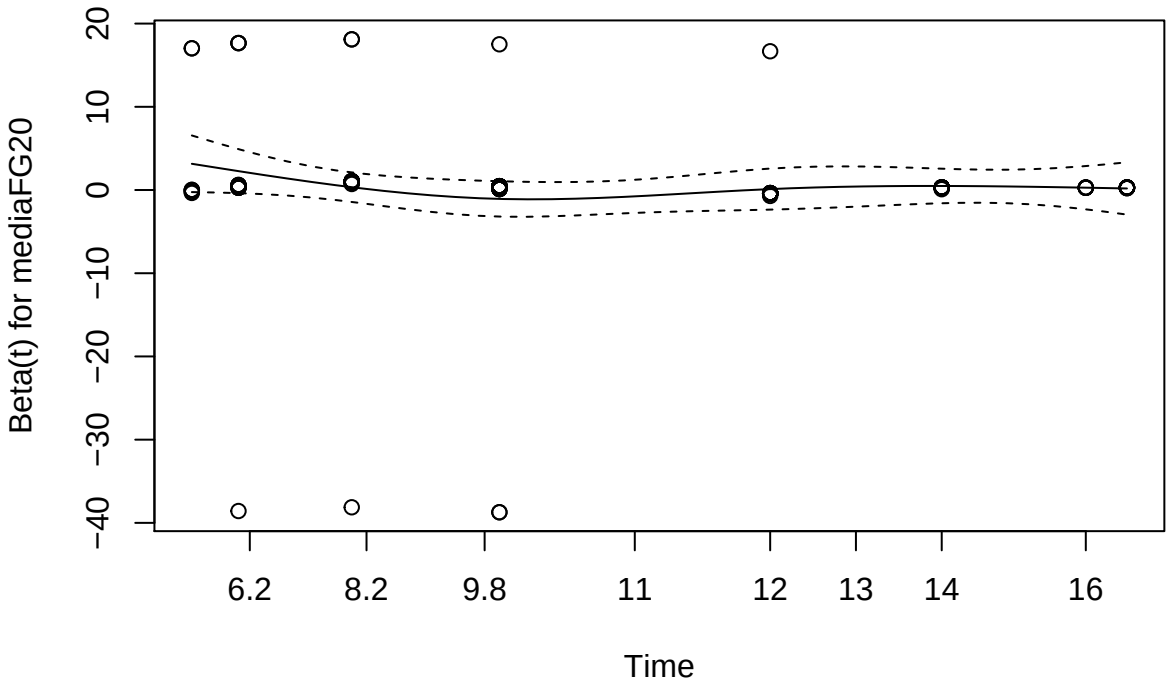

Beta(t) for mediaFG20
6
4
2
0
−2
−4
−6
6.2
8.2
9.8
11
12
13
14
16
Time

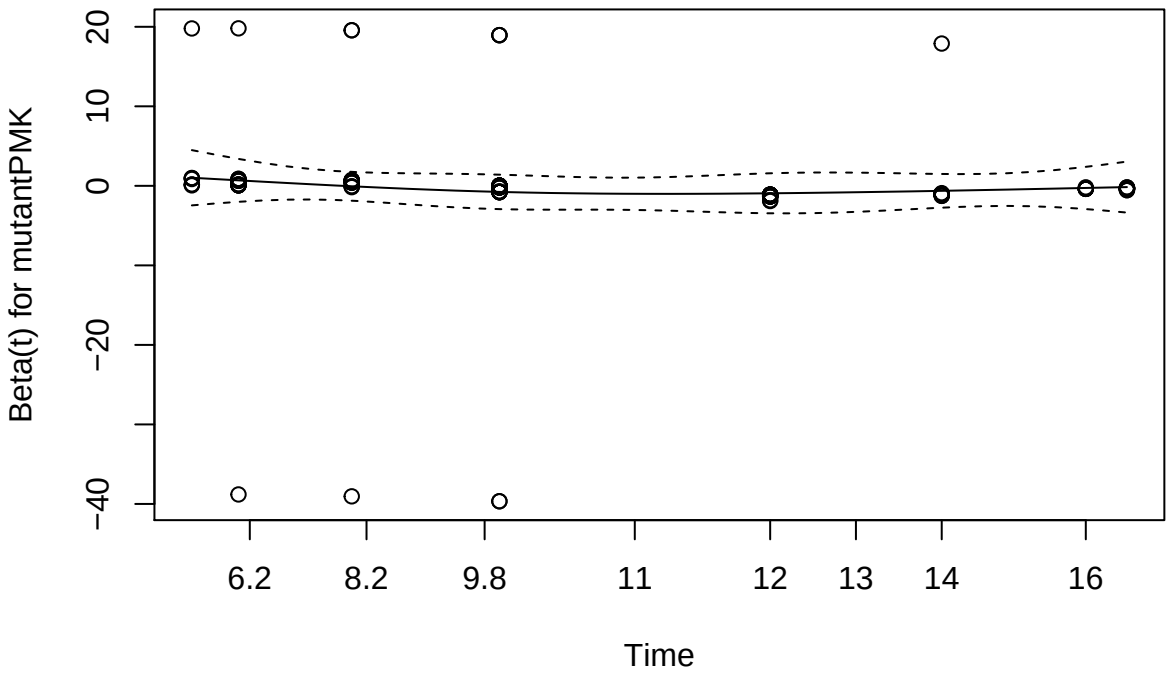

Beta(t) for mutantPMK
2
0
−2
−4
6.2
8.2
9.8
11
12
13
14
16
Time

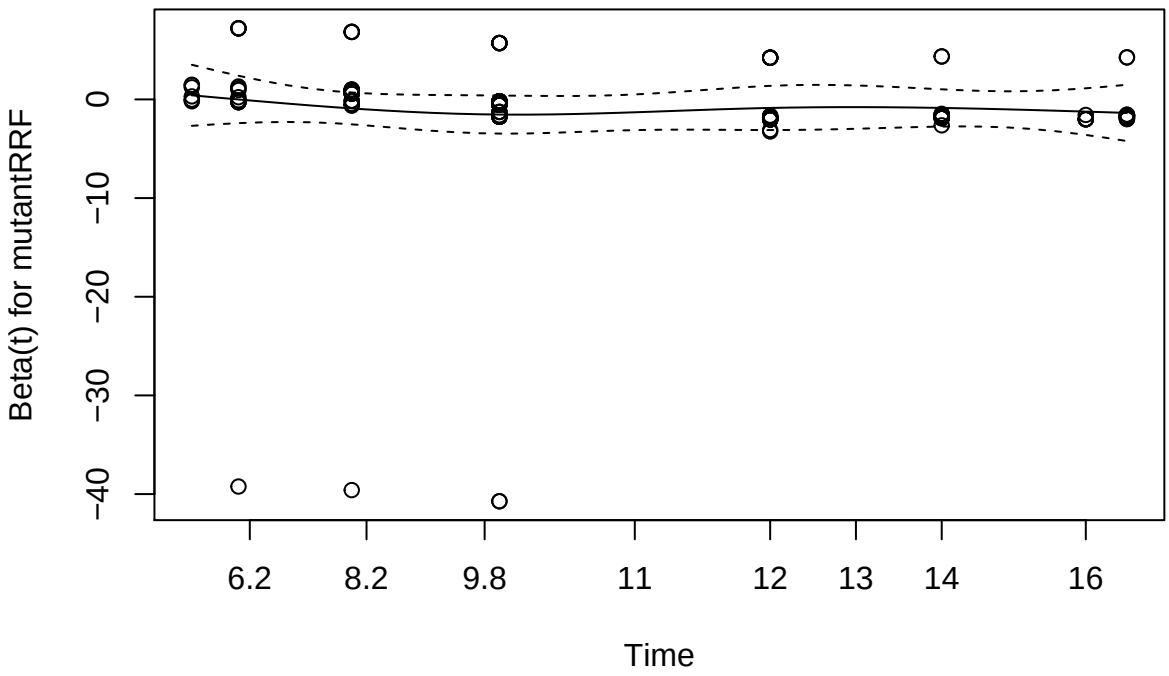


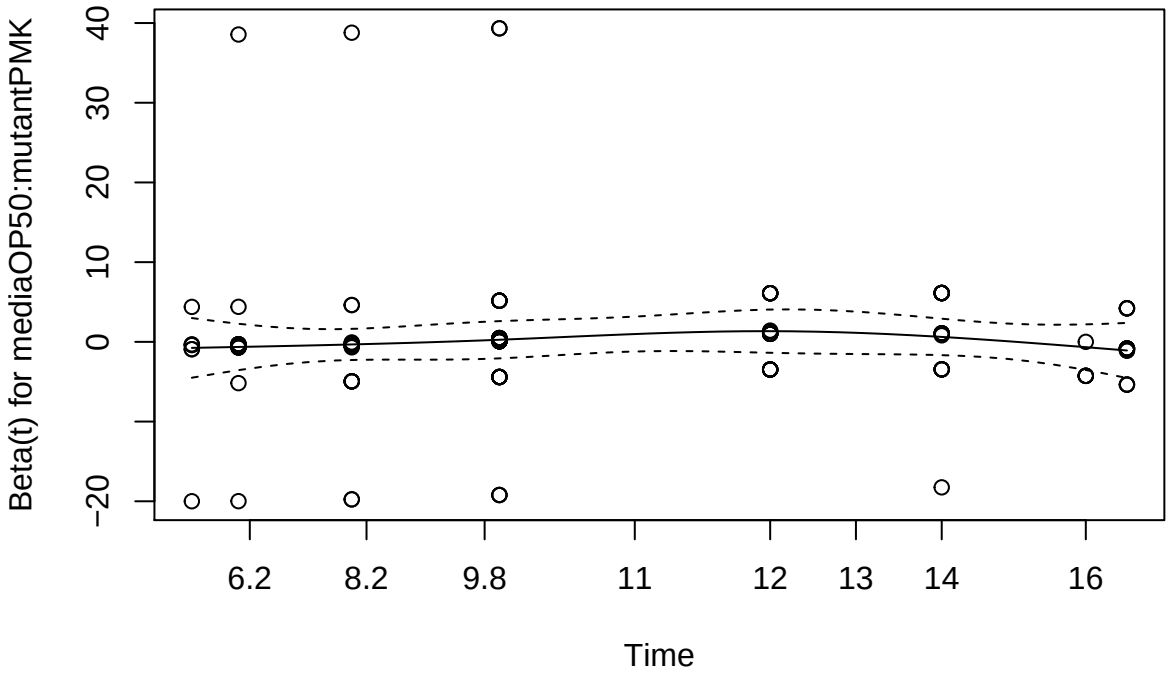

```
plot(prop_int)
```

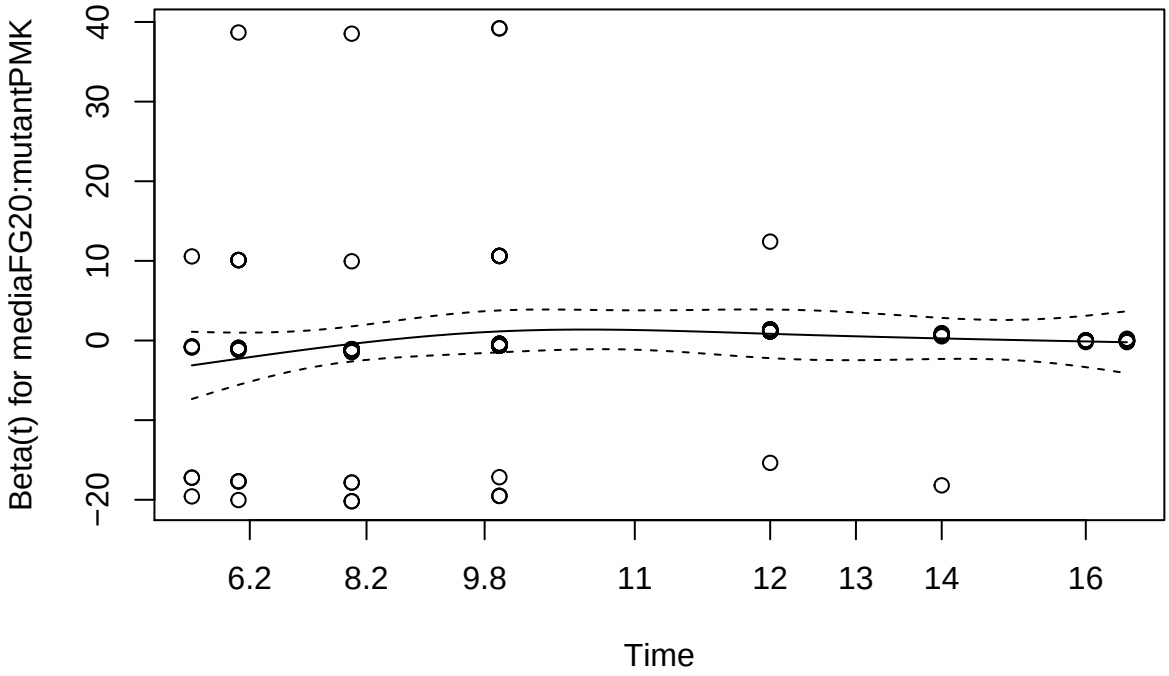

Beta(t) for mediaOP50
0
−10
−20
−30
−40
6.2
8.2
9.8
11
12
13
14
16
Time

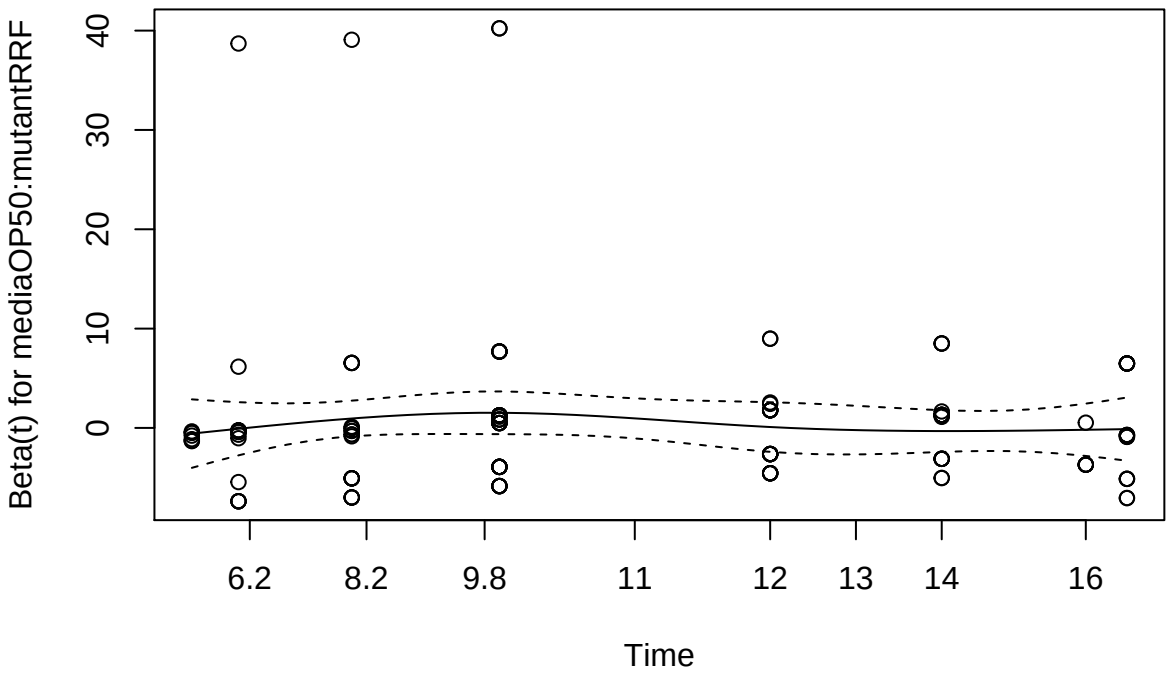

Beta(t) for mediaFG20
20
10
0
−10
−20
−30
−40
6.2
8.2
9.8
11
12
13
14
16
Time

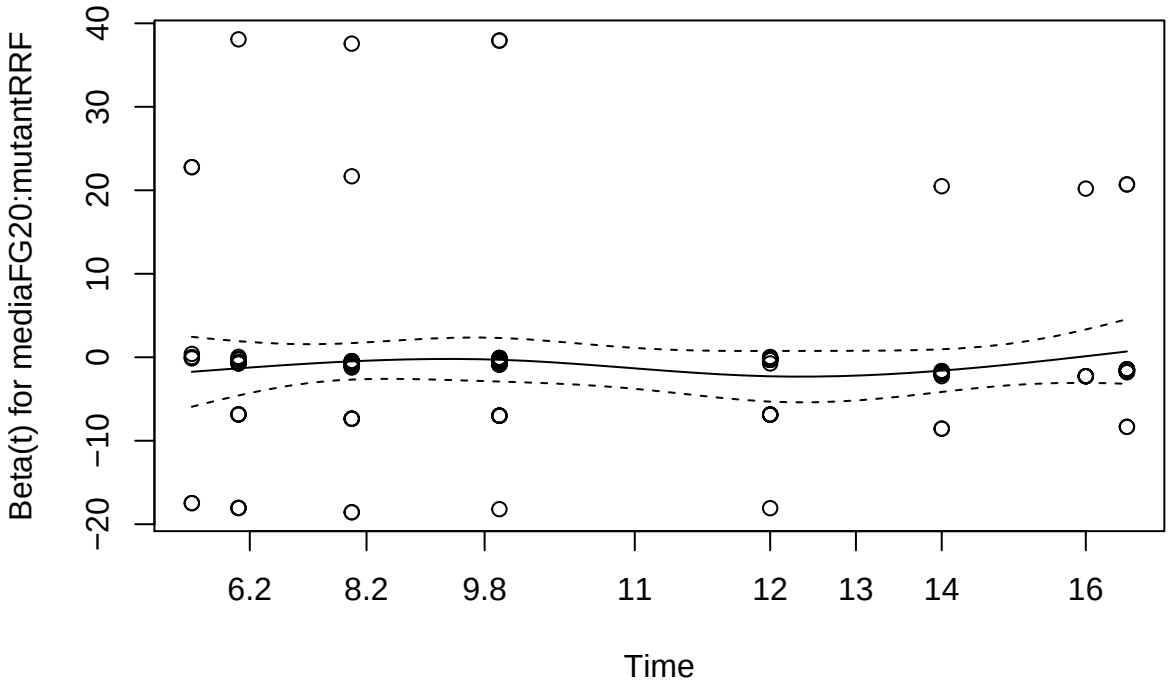

Beta(t) for mutantPMK
20
10
0
−20
−40
6.2
8.2
9.8
11
12
13
14
16
Time


Beta(t) for mutantRRF
0
−10
−20
−30
−40
6.2
8.2
9.8
11
12
13
14
16
Time


Beta(t) for mediaOP50:mutantPMK
40
30
20
10
0
−20
6.2
8.2
9.8
11
12
13
14
16
Time


Beta(t) for mediaFG20:mutantPMK
40
30
20
10
0
−20
6.2
8.2
9.8
11
12
13
14
16
Time

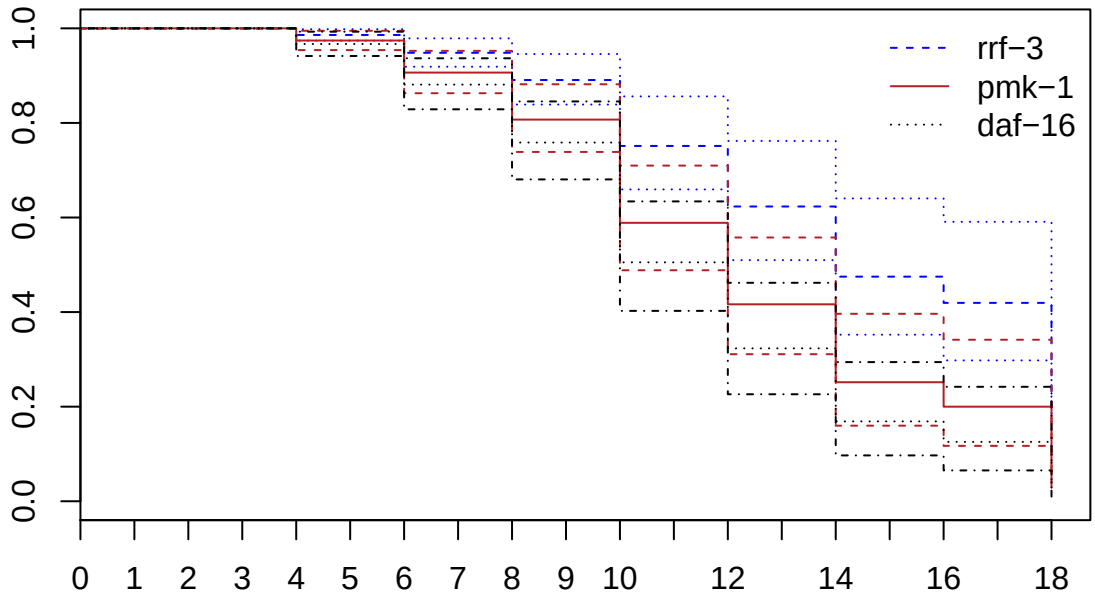

Beta(t) for mediaOP50:mutantRRF
0
10
20
30
40
6.2
8.2
9.8
11
12
13
14
16
Time



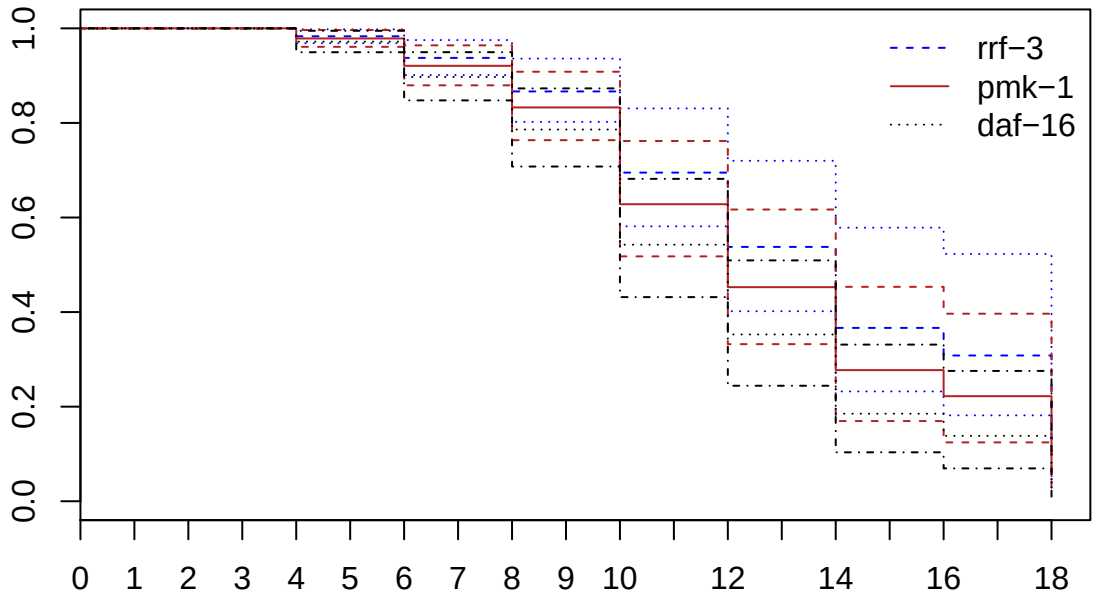
Both schoenfield residuals and the test is not significant so we proceed as if the assumptions is true for our dataset.

Then we plot predictions for both models

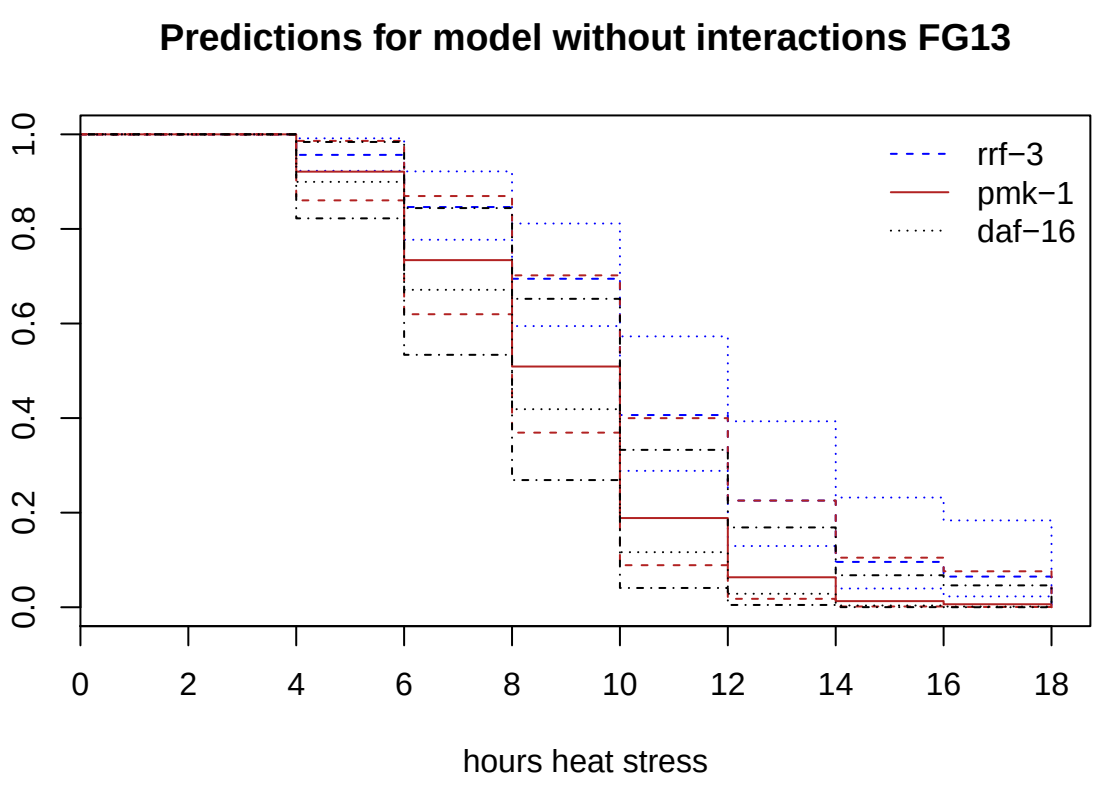
```
plot(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'OP50')), lty = 2, xaxp = c(0,18,18
     main = 'Predictions for model without interactions OP50')
lines(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'OP50')), col = 'firebrick')
lines(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'OP50')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n')
```



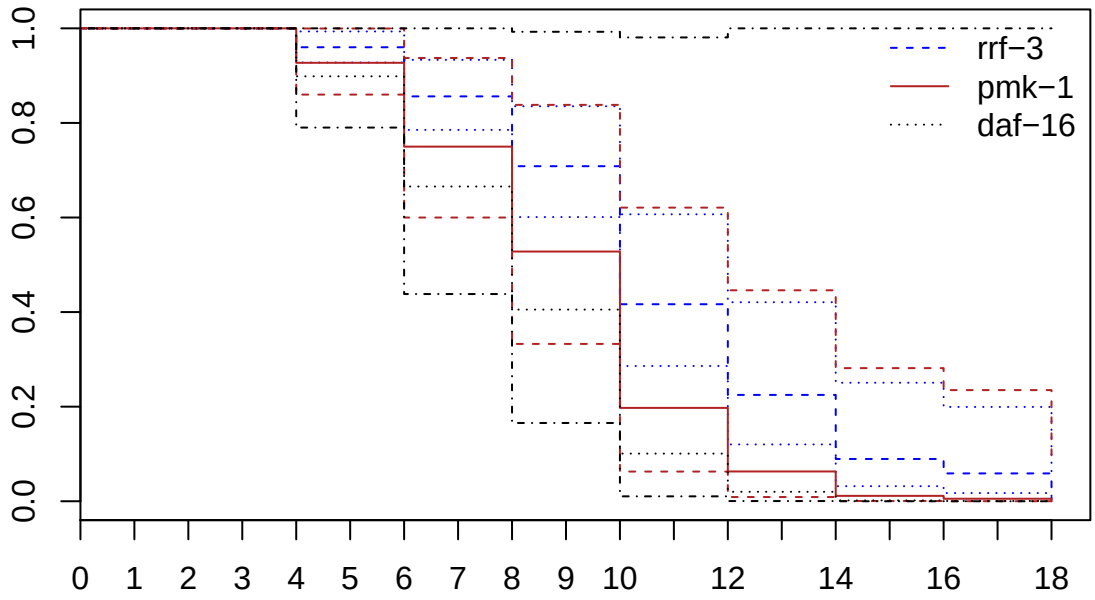

```
plot(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'OP50')), lty = 2, xaxp = c(0,18,18)
     main = 'Predictions for model with interactions OP50')
lines(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'OP50')), col = 'firebrick')
lines(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'OP50')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n')
```



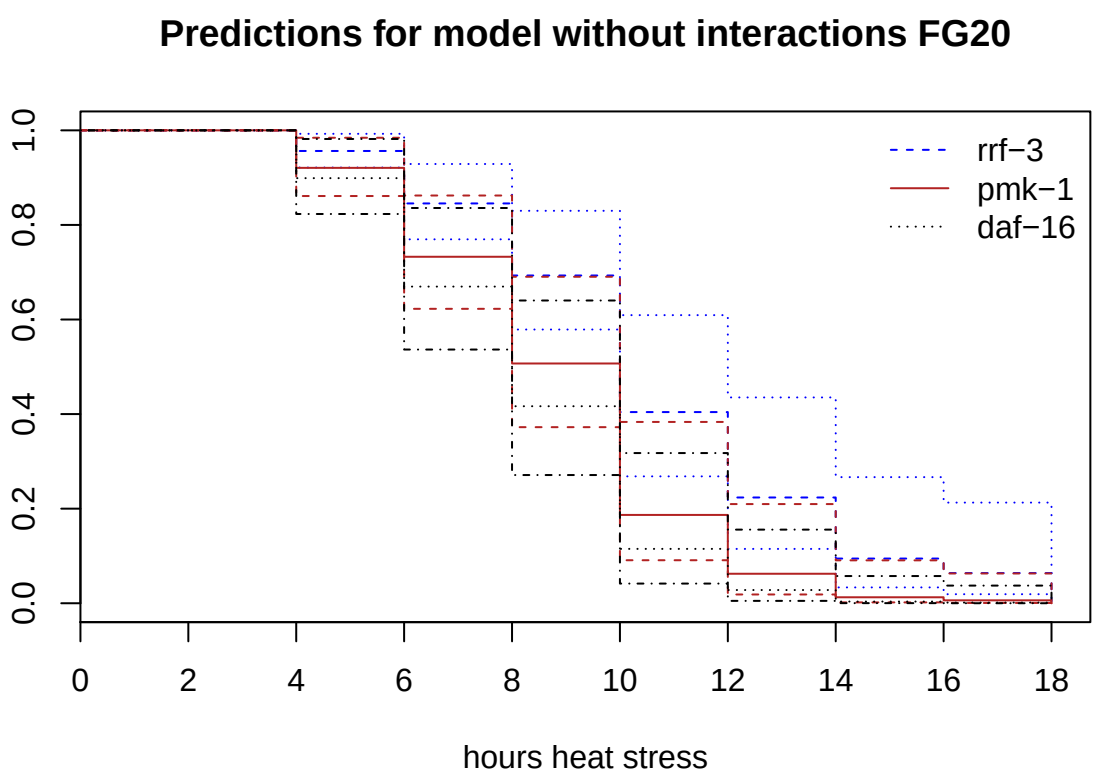

```
plot(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'FG13')), lty = 2, xaxp = c(0,18,9)
     main = 'Predictions for model without interactions FG13',
     xlab = 'hours heat stress')
lines(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'FG13')), col = 'firebrick')
lines(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'FG13')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
        lty = c(2,1,3), col = c('blue','firebrick','black'),
        bty = 'n')
```



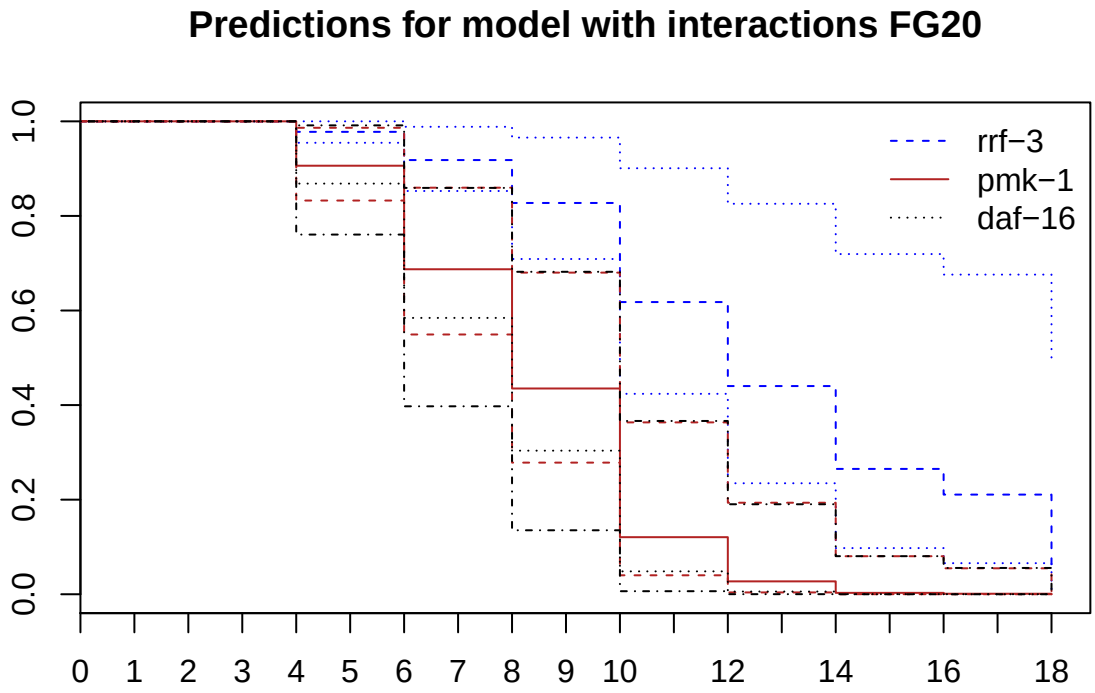

```
plot(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'FG13')), lty = 2, xaxp = c(0,18,18)
     main = 'Predictions for model with interactions FG13')
lines(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'FG13')), col = 'firebrick')
lines(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'FG13')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n')
```




```
plot(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'FG20')), lty = 2, xaxp = c(0,18,9)
     main = 'Predictions for model without interactions FG20',
     xlab = 'hours heat stress')
lines(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'FG20')), col = 'firebrick')
lines(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'FG20')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
        lty = c(2,1,3), col = c('blue','firebrick','black'),
        bty = 'n')
```

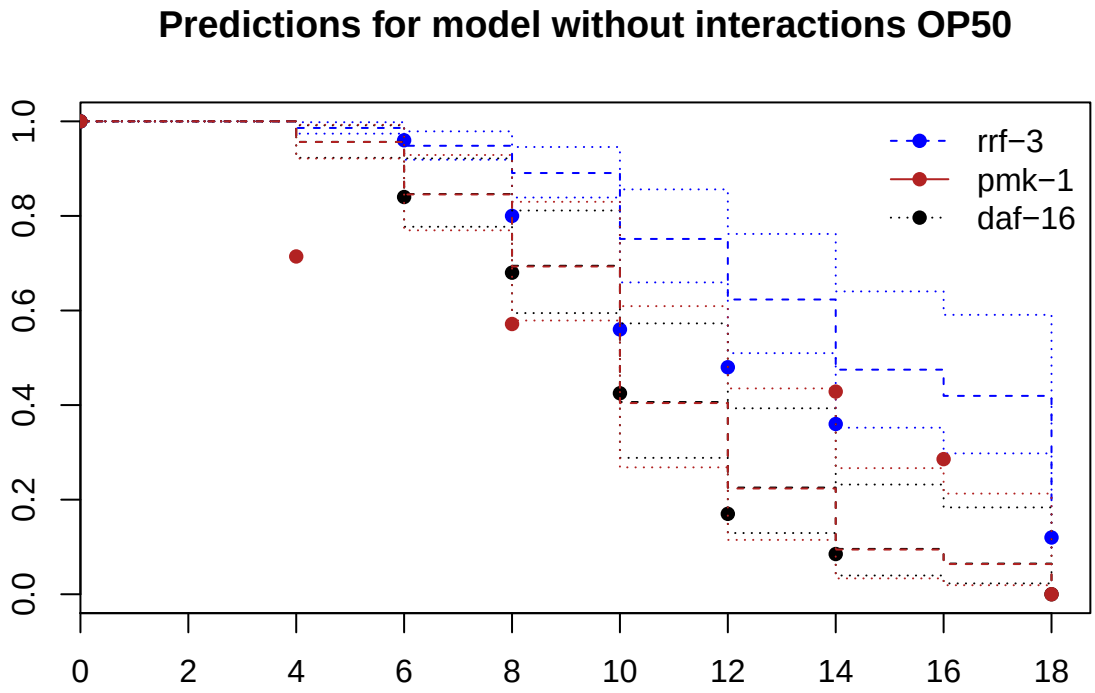



```
plot(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'FG20')), lty = 2, xaxp = c(0,18,18)
     main = 'Predictions for model with interactions FG20')
lines(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'FG20')), col = 'firebrick')
lines(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'FG20')), lty =3)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n')
```



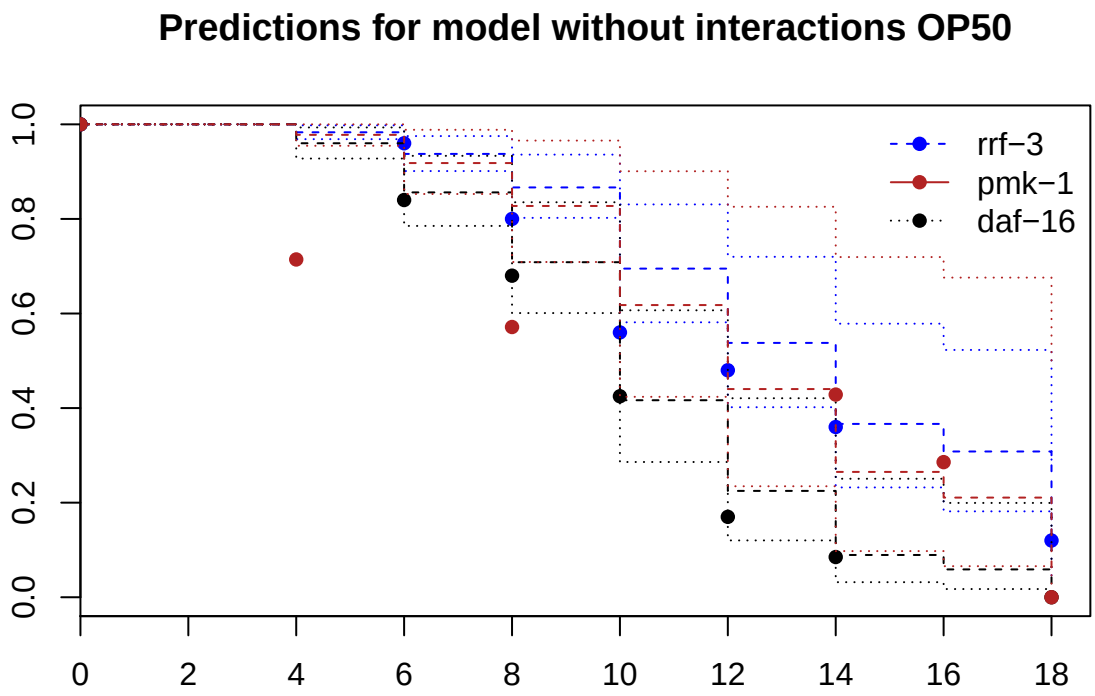
We check how well predictions line up with reality

```
plot(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'OP50')), lty = 2, xaxp = c(0,18,9)
     main = 'Predictions for model without interactions OP50')
points(df_OP50_r$Time,df_OP50_r$Surv, col = 'blue', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'FG13')), lty = 2, col = 'black',)
points(df_fly_13_r$Time,df_fly_13_r$Surv, col = 'black', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'RRF', media = 'FG20')), lty = 2, col = 'firebric
points(df_fly_20_r$Time,df_fly_20_r$Surv, col = 'firebrick', pch = 16)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n', pch = 16)
```



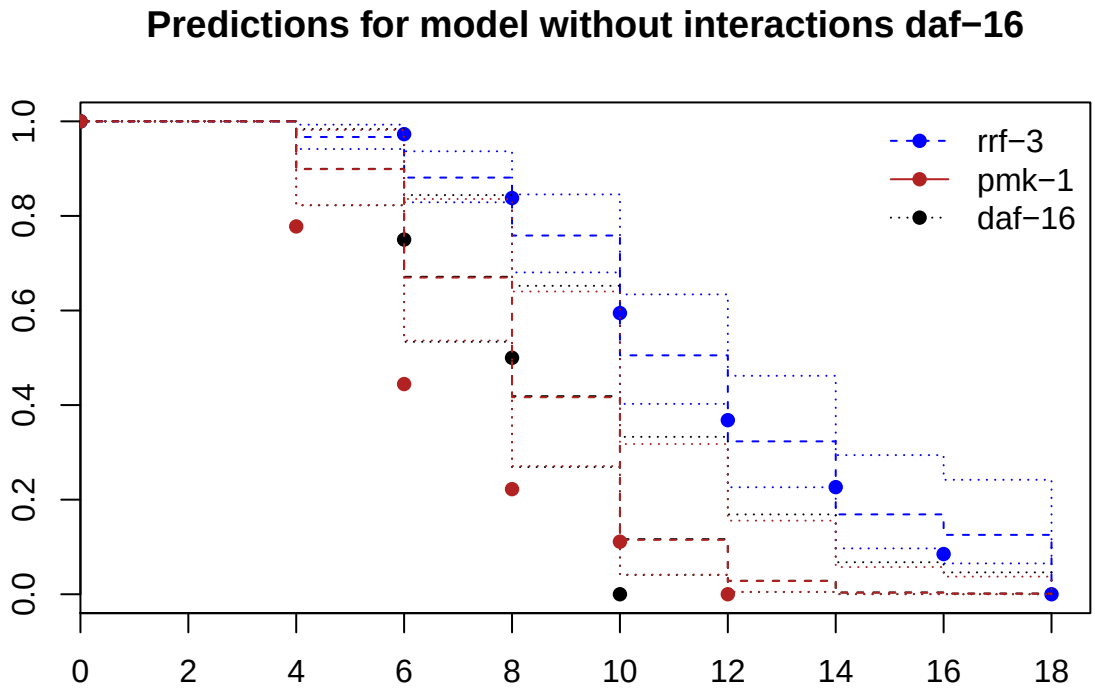
Fit is very bad as is clear both concordance and R^2 from the model. But there is far too little uncertainty in the confidence intervals. Which overestimates survival heavily for this group. Lets try the other model

```
plot(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'OP50')), lty = 2, xaxp = c(0,18,9),
     main = 'Predictions for model without interactions OP50')
points(df_OP50_r$Time,df_OP50_r$Surv, col = 'blue', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'FG13')), lty = 2, col = 'black')
points(df_fly_13_r$Time,df_fly_13_r$Surv, col = 'black', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'RRF', media = 'FG20')), lty = 2, col = 'firebrick
points(df_fly_20_r$Time,df_fly_20_r$Surv, col = 'firebrick', pch = 16)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
        lty = c(2,1,3), col = c('blue','firebrick','black'),
        bty = 'n', pch = 16)
```



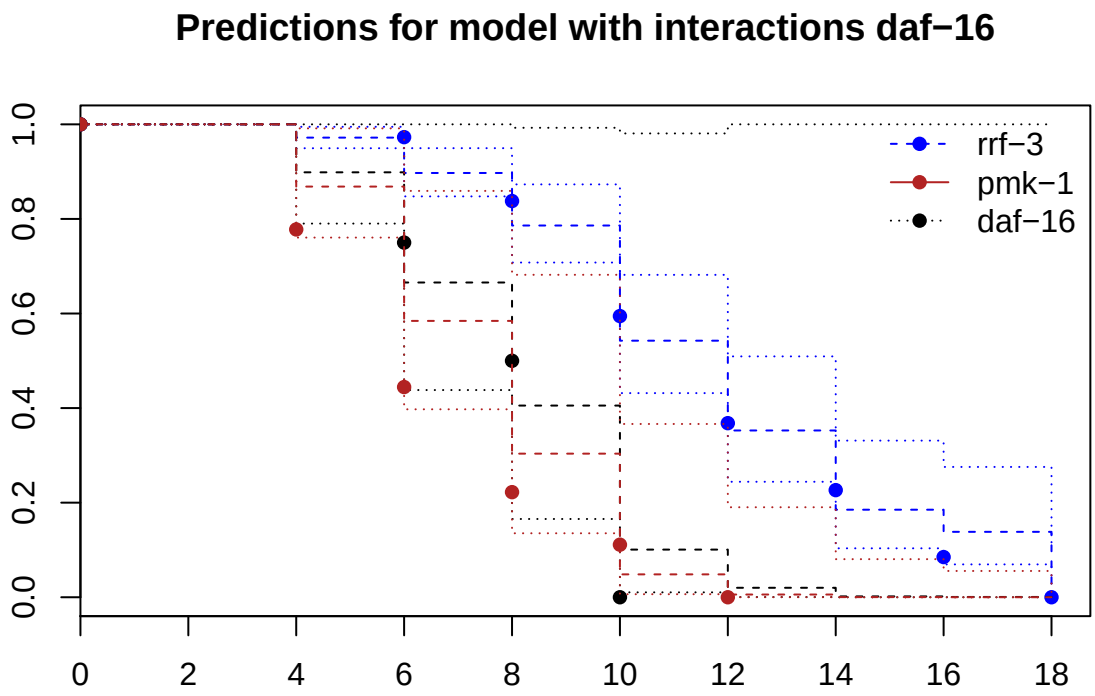
Still on the low end but a lot better. Seems that without interactions it overestimates survival on OP50.

```
plot(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'OP50')), lty = 2, xaxp = c(0,18,9)
     main = 'Predictions for model without interactions daf-16')
points(df_OP50_d$Time,df_OP50_d$Surv, col = 'blue', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'FG13')), lty = 2, col = 'black',)
points(df_fly_13_d$Time,df_fly_13_d$Surv, col = 'black', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'DAF', media = 'FG20')), lty = 2, col = 'firebric
points(df_fly_20_d$Time,df_fly_20_d$Surv, col = 'firebrick', pch = 16)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n', pch = 16)
```



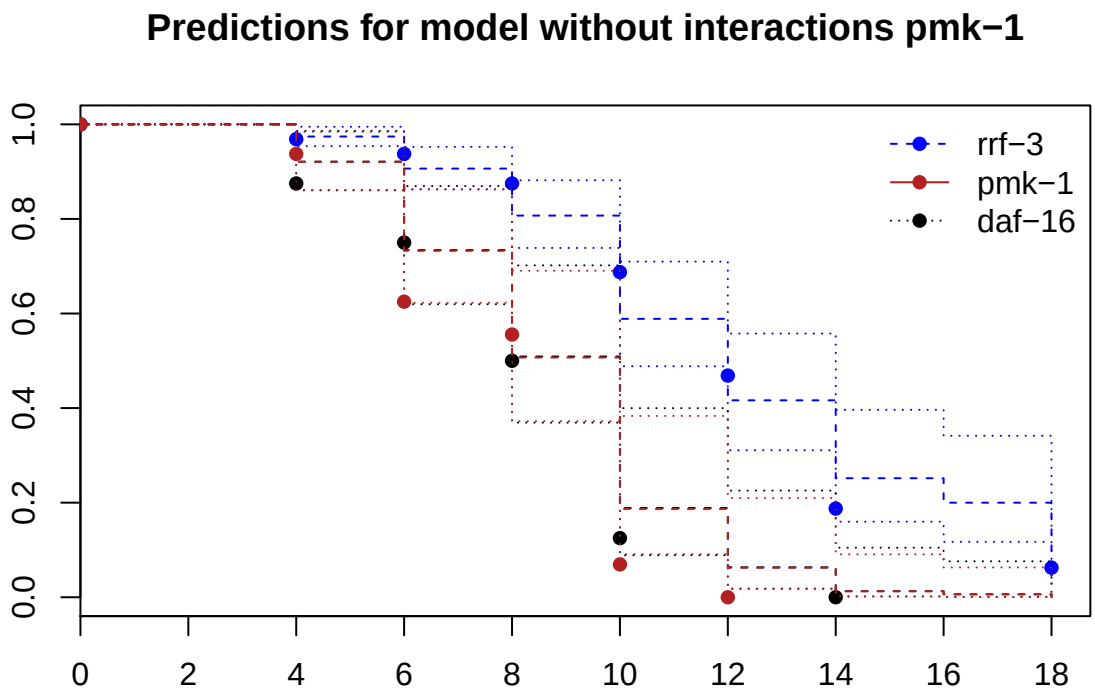
Fits notably better

```
plot(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'OP50')), lty = 2, xaxp = c(0,18,9),
     main = 'Predictions for model with interactions daf-16')
points(df_OP50_d$Time,df_OP50_d$Surv, col = 'blue', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'FG13')), lty = 2, col = 'black',)
points(df_fly_13_d$Time,df_fly_13_d$Surv, col = 'black', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'DAF', media = 'FG20')), lty = 2, col = 'firebrick
points(df_fly_20_d$Time,df_fly_20_d$Surv, col = 'firebrick', pch = 16)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
        lty = c(2,1,3), col = c('blue','firebrick','black'),
        bty = 'n', pch = 16)
```



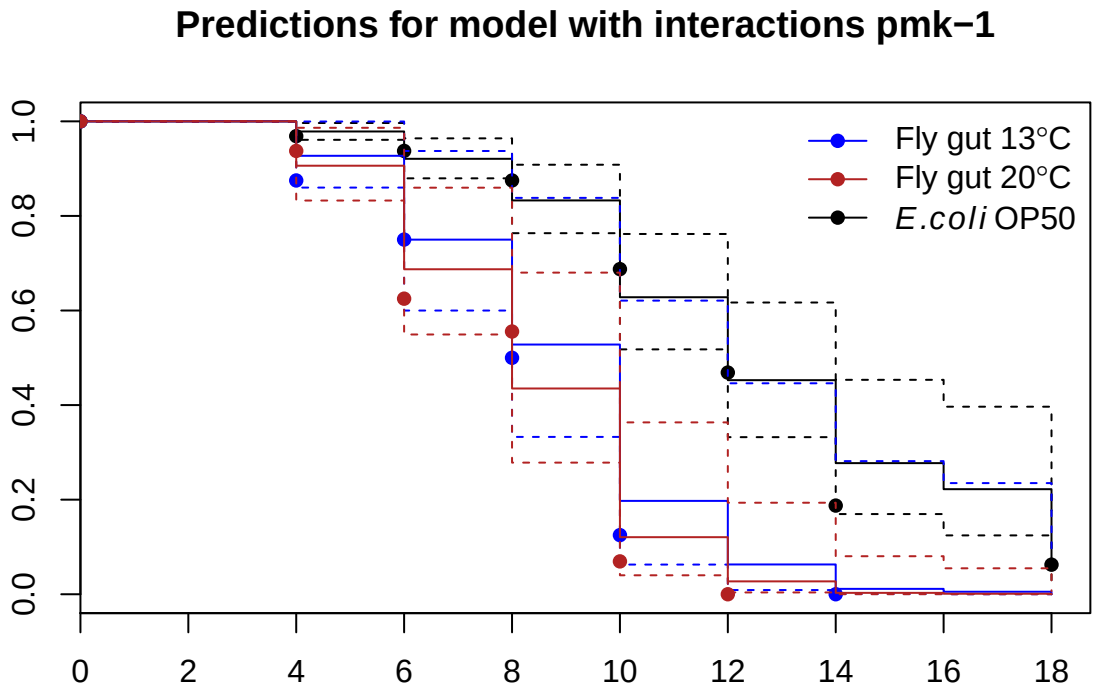
Change opacity on some points to fix overlap

```
plot(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'OP50')), lty = 2, xaxp = c(0,18,9)
     main = 'Predictions for model without interactions pmk-1')
points(df_OP50_P$Time,df_OP50_P$Surv, col = 'blue', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'FG13')), lty = 2, col = 'black',)
points(df_fly_13_P$Time,df_fly_13_P$Surv, col = 'black', pch = 16)
lines(survfit(fit_both, newdata = data.frame(mutant = 'PMK', media = 'FG20')), lty = 2, col = 'firebric
points(df_fly_20_P$Time,df_fly_20_P$Surv, col = 'firebrick', pch = 16)
legend('topright', legend = c('rrf-3','pmk-1','daf-16'),
       lty = c(2,1,3), col = c('blue','firebrick','black'),
       bty = 'n', pch = 16)
```



Fitsvery bad with far too much uncertainty

```
plot(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'OP50')), lty = 1, xaxp = c(0,18,9),
     main = 'Predictions for model with interactions pmk-1')
points(df_OP50_P$Time,df_OP50_P$Surv, col = 'black', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'FG13')), lty = 1, col = 'blue')
points(df_fly_13_P$Time,df_fly_13_P$Surv, col = 'blue', pch = 16)
lines(survfit(fit_int, newdata = data.frame(mutant = 'PMK', media = 'FG20')), lty = 1, col = 'firebrick
points(df_fly_20_P$Time,df_fly_20_P$Surv, col = 'firebrick', pch = 16)
legend('topright', pch = c(16,16,16),lty = c(1,1,1),
       col = c('blue','firebrick','black'),
       legend = c(expression('Fly gut 13'*degree*C),
                  expression('Fly gut 20'*degree*C),
                  expression(italic(E.coli) ~ OP50)), bty = 'n')
```



The model has a problem with catching that all worms die around the same time and may therefore overestimate survival as it fits okay until the worms start dying. Some other model structure is maybe necessary or not all important covariates of survival are measured.

```
d_1350_r <- subset(d_RRF, media == 'FG13' | media == 'OP50')
d_2050_r <- subset(d_RRF, media == 'FG20' | media == 'OP50')
d_1320_r <- subset(d_RRF, media == 'FG13' | media == 'FG20')
survdiff(Survobj ~media ,  data = d_1350_r)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1350_r)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 25       24     17.6      2.32      5.64
## media=OP50 25       22     28.4      1.44      5.64
##
##  Chisq= 5.6  on 1 degrees of freedom, p= 0.02
```

```
survdiff(Survobj ~media ,  data = d_2050_r)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_2050_r)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG20  7        7     5.62     0.342     0.599
```

```
## media=OP50 25       22    23.38     0.082     0.599
##
##  Chisq= 0.6  on 1 degrees of freedom, p= 0.4
```

```
survdiff(Survobj ~media ,  data = d_1320_r)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1320_r)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 25       24    21.92     0.197      1.16
## media=FG20  7        7     9.08     0.476      1.16
##
##  Chisq= 1.2  on 1 degrees of freedom, p= 0.3
```

```
d_1350_p <- subset(d_PMK, media == 'FG13' | media == 'OP50')
d_2050_p <- subset(d_PMK, media == 'FG20' | media == 'OP50')
d_1320_p <- subset(d_PMK, media == 'FG13' | media == 'FG20')
survdiff(Survobj ~media ,  data = d_1350_p)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1350_p)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13  8        8     3.49      5.83      8.89
## media=OP50 32       30    34.51      0.59      8.89
##
##  Chisq= 8.9  on 1 degrees of freedom, p= 0.003
```

```
survdiff(Survobj ~media ,  data = d_2050_p)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_2050_p)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG20 16       15     6.55     10.89      18.7
## media=OP50 32       30    38.45      1.86      18.7
##
##  Chisq= 18.7  on 1 degrees of freedom, p= 2e-05
```

```
survdiff(Survobj ~media ,  data = d_1320_p)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1320_p)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13  8        8     8.61    0.0431     0.138
## media=FG20 16       15    14.39    0.0258     0.138
##
##  Chisq= 0.1  on 1 degrees of freedom, p= 0.7
```

```r
d_1350_d <- subset(d_DAF, media == 'FG13' | media == 'OP50')
d_2050_d <- subset(d_DAF, media == 'FG20' | media == 'OP50')
d_1320_d <- subset(d_DAF, media == 'FG13' | media == 'FG20')
survdiff(Survobj ~media ,  data = d_1350_d)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1350_d)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13  4        4     1.32     5.414      7.37
## media=OP50 37       36    38.68     0.185      7.37
##
##  Chisq= 7.4  on 1 degrees of freedom, p= 0.007
```

```r
survdiff(Survobj ~media ,  data = d_2050_d)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_2050_d)
##
##             N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG20  9        9     2.74    14.274      19.7
## media=OP50 37       36    42.26     0.927      19.7
##
##  Chisq= 19.7  on 1 degrees of freedom, p= 9e-06
```

```r
survdiff(Survobj ~media ,  data = d_1320_d)
```

```
## Call:
## survdiff(formula = Survobj ~ media, data = d_1320_d)
##
##            N Observed Expected (O-E)^2/E (O-E)^2/V
## media=FG13 4        4     4.86    0.1508     0.412
## media=FG20 9        9     8.14    0.0899     0.412
##
##  Chisq= 0.4  on 1 degrees of freedom, p= 0.5
```