

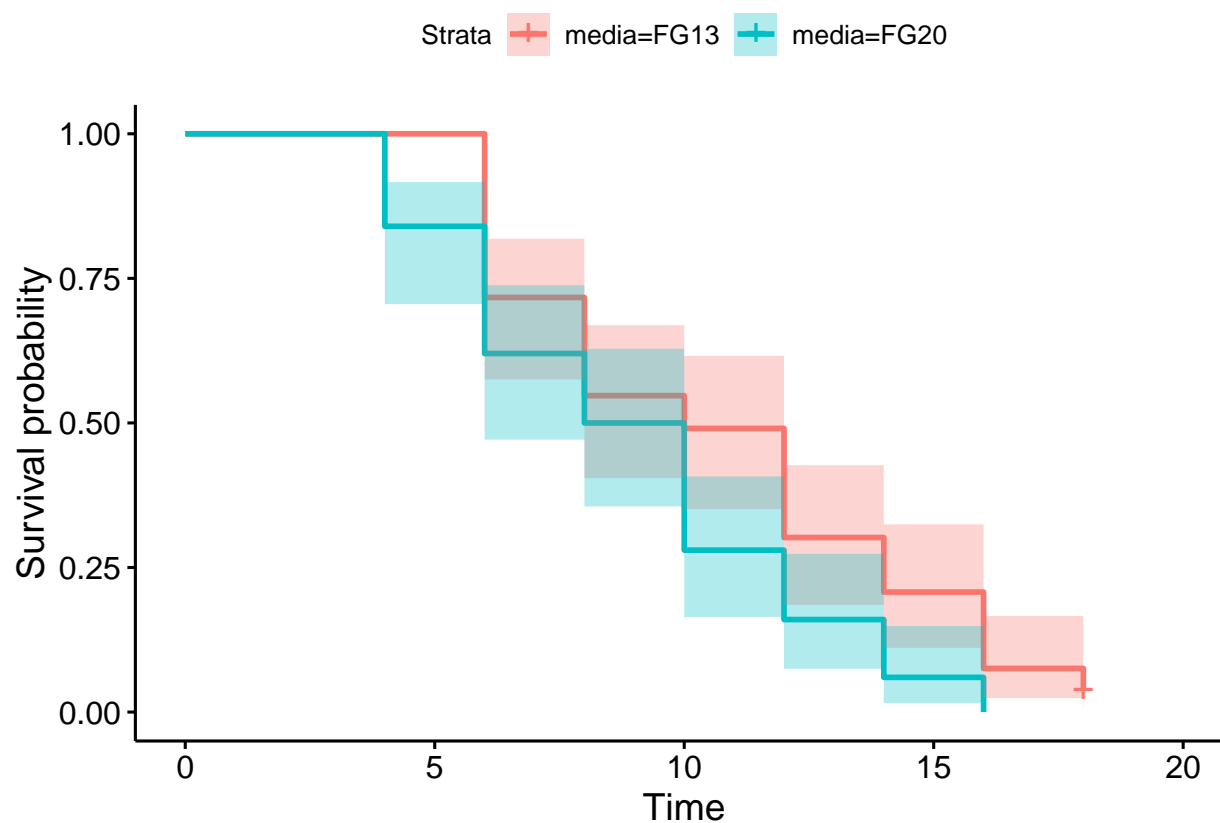
# Metabolite

*Jonas Gehrlein*

*29 nov 2018*

```
data <- read.table(here('data','Worm_4days.csv'), header = TRUE,
                  sep = ';')
data_2 <- read.table(here('data','Worm_second.csv'), header = TRUE,
                   sep = ';')
colnames(data_2) <- c('Time', 'media', 'Metabolite', 'Status')
data_met <- subset(data_2, Metabolite == 1)
d_OP50 <- subset(data_2, media == 'OP50')
```

```
data_met$Survobj <- with(data_met, Surv(data_met$Time, event = data_met$Status))
km <- survfit(Survobj ~ media, data = data_met, conf.type = "log-log", error = "greenwood")
s_km <- summary(km)
ggsurvplot(km, conf.int = TRUE)
```



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

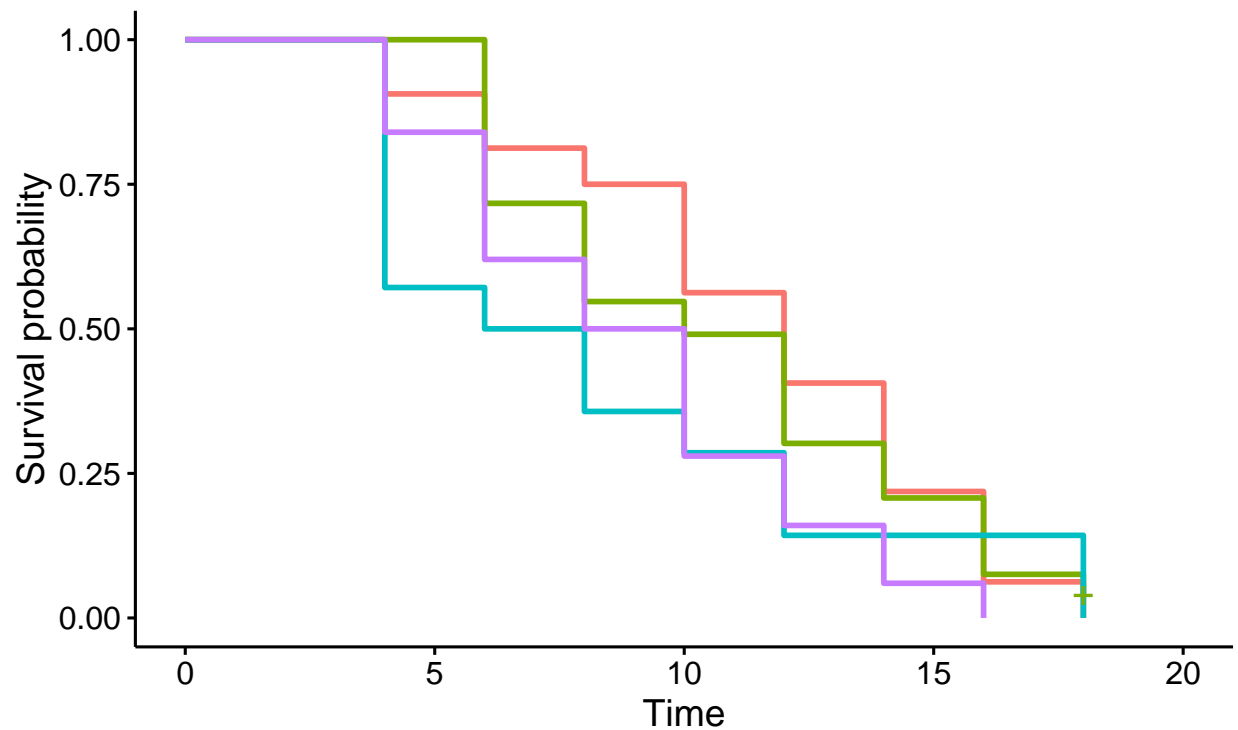
```
## Call:
## coxph(formula = Survobj ~ media, data = data_met)
```

```
##
## n= 103, number of events= 101
##
##      coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG20 0.5178    1.6784   0.2061 2.513   0.012 *
## mediaOP50    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
## mediaFG20      1.678      0.5958      1.121      2.514
## mediaOP50      NA         NA         NA         NA
##
## Concordance= 0.571 (se = 0.031 )
## Rsquare= 0.059 (max possible= 0.999 )
## Likelihood ratio test= 6.28 on 1 df,  p=0.01
## Wald test              = 6.31 on 1 df,  p=0.01
## Score (logrank) test = 6.44 on 1 df,  p=0.01
```

Also lower survival for FG20 if only including survival and higher survival for OP50 which is higher than if we had compared with fly gut which isn't only metabolites. If we include metabolites in a cox model without OP50

```
data_2_fly <- subset(data_2, media != 'OP50' )
data_2_fly$Survobj <- with(data_2_fly, Surv(data_2_fly$Time, event = data_2_fly$Status))
km_fly <- survfit(Survobj ~ media + Metabolite, data = data_2_fly, conf.type = "log-log", error = "green")
ggsurvplot(km_fly)
```

media=FG13, Metabolite=0 + media=FG13, Metabolite=1 + media=FG20, Metabolite=0 + media=FG20, Metabolite=1 +



No big difference on the plot between metabolite or without if we fit a model of both media and metabolite

```
fit_fly_meta <- coxph(Survobj ~ media + Metabolite , data = data_2_fly)
summary(fit_fly_meta)
```

```
## Call:
## coxph(formula = Survobj ~ media + Metabolite, data = data_2_fly)
##
##   n= 149, number of events= 147
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG20  0.49618   1.64243  0.17022  2.915  0.00356 **
## mediaOP50      NA         NA  0.00000   NA      NA
## Metabolite  0.08789   1.09186  0.18127  0.485  0.62778
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## mediaFG20      1.642      0.6089   1.1765   2.293
## mediaOP50       NA         NA      NA      NA
## Metabolite      1.092      0.9159   0.7654   1.558
##
## Concordance= 0.59 (se = 0.027 )
## Rsquare= 0.059 (max possible= 1 )
## Likelihood ratio test= 9.07 on 2 df,  p=0.01
## Wald test           = 9.31 on 2 df,  p=0.01
```

```
## Score (logrank) test = 9.5 on 2 df, p=0.009
```

```
AIC(fit_fly_meta)
```

```
## [1] 1193.565
```

```
fit_fly_med <- coxph(Survobj ~ media, data = data_2_fly)
summary(fit_fly_med)
```

```
## Call:
## coxph(formula = Survobj ~ media, data = data_2_fly)
##
## n= 149, number of events= 147
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## mediaFG20 0.5080    1.6619  0.1684 3.016  0.00256 **
## mediaOP50      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
## mediaFG20      1.662      0.6017    1.195    2.312
## mediaOP50      NA         NA         NA         NA
##
## Concordance= 0.586 (se = 0.025 )
## Rsquare= 0.058 (max possible= 1 )
## Likelihood ratio test= 8.83 on 1 df, p=0.003
## Wald test = 9.1 on 1 df, p=0.003
## Score (logrank) test = 9.28 on 1 df, p=0.002
```

```
AIC(fit_fly_med)
```

```
## [1] 1191.803
```

It certainly isn't a good predictor in the model. and is centered around 0. But very uncertain about the exact size on the size of the predictor. The only real difference is for FG13 where metabolite has worse survival

```
data_bac <- subset(data_2_fly, Metabolite == 0)
data_bac$Survobj <- with(data_bac, Surv(data_bac$Time, event = data_bac$Status))
km_bac <- survfit(Survobj ~ media, data = data_bac, conf.type = "log-log", error = "greenwood")
s_km_bac <- summary(km_bac)

df_13_bac <- data.frame(c(0, s_km_bac$time[1:8]), c(1, s_km_bac$surv[1:8]),
                        c(0, s_km_bac$std.err[1:8]))
colnames(df_13_bac) <- c('Time', 'Surv', 'Std.error')
df_20_bac <- data.frame(c(0, s_km_bac$time[9:14]), c(1, s_km_bac$surv[9:14]),
                        c(0, s_km_bac$std.err[9:14]))
colnames(df_20_bac) <- c('Time', 'Surv', 'Std.error')

data_meta <- subset(data_met, media != 'OP50')
```

```
km_meta <- survfit(Survobj ~ media, data = data_meta, conf.type = "log-log", error = "greenwood")
s_km_meta <- summary(km_meta)
```

```
df_13_meta <- data.frame(c(0, s_km_meta$time[1:7]), c(1, s_km_meta$urv[1:7]),
                        c(0, s_km_meta$std.err[1:7]))
colnames(df_13_meta) <- c('Time', 'Surv', 'Std.error')
df_20_meta <- data.frame(c(0, s_km_meta$time[8:14]), c(1, s_km_meta$urv[8:14]),
                        c(0, s_km_meta$std.err[8:14]))
colnames(df_20_meta) <- c('Time', 'Surv', 'Std.error')
```

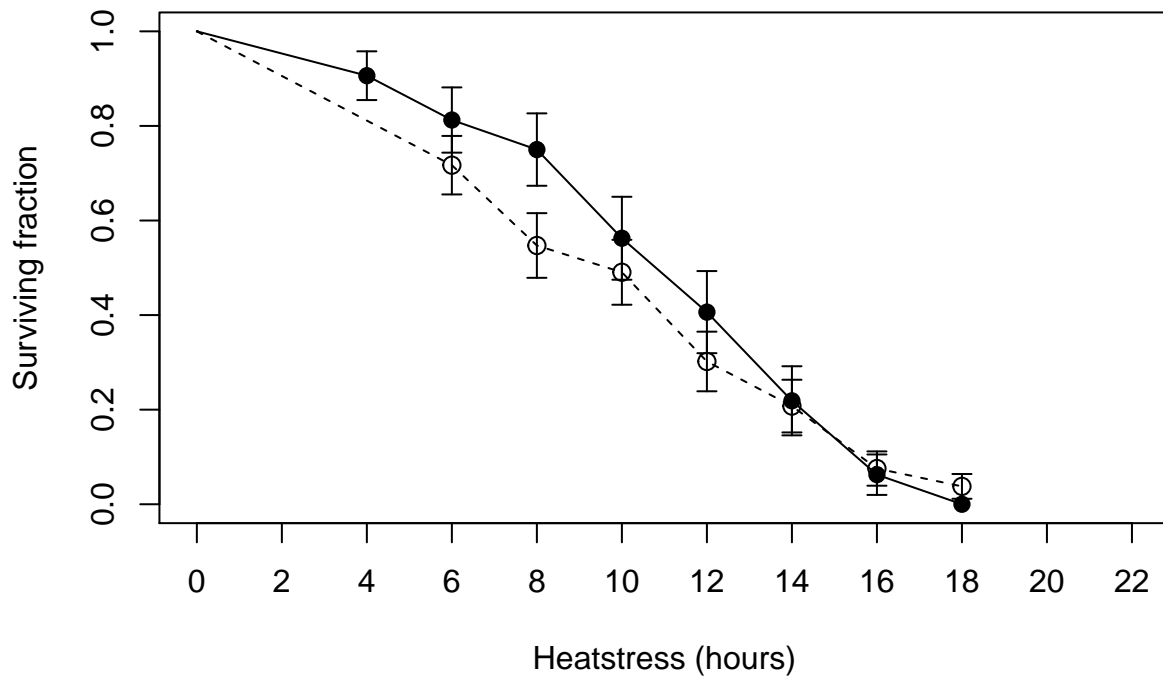
```
plot(df_13_bac$Time[2:9], df_13_bac$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Heatstress (hours)',
     ylab = 'Surviving fraction',
     main = expression('Comparison between fly gut 20 without and with living bacteria'),
     xlim = c(0, 22), ylim = c(0, 1), xaxp = c(0, 22, 11))
lines(df_13_bac$Time, df_13_bac$Surv)
arrows(df_13_bac$Time, df_13_bac$Surv - df_13_bac$Std.error, df_13_bac$Time,
       df_13_bac$Surv + df_13_bac$Std.error, length = 0.05, angle = 90, code = 3, col = 'black')
```

```
## Warning in arrows(df_13_bac$Time, df_13_bac$Surv - df_13_bac$Std.error, :
## zero-length arrow is of indeterminate angle and so skipped
```

```
points(df_13_meta$Time[2:9], df_13_meta$Surv[2:9], pch = 1, cex = 1.2)
lines(df_13_meta$Time, df_13_meta$Surv, lty = 2)
arrows(df_13_meta$Time, df_13_meta$Surv - df_13_meta$Std.error, df_13_meta$Time, df_13_meta$Surv + df_13_me
```

```
## Warning in arrows(df_13_meta$Time, df_13_meta$Surv -
## df_13_meta$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

## Comparison between fly gut 20 without and with living bacteria



```
plot(df_20_bac$Time[2:9],df_20_bac$Surv[2:9], pch = 16, cex = 1.2, xlab = 'Heatstress (hours)',
     ylab = 'Surviving fraction',
     main = expression('Comparison between fly gut 20 without and with living bacteria'),
     xlim = c(0,22), ylim = c(0,1), xaxp = c(0,22,11))
lines(df_20_bac$Time,df_20_bac$Surv)
arrows(df_20_bac$Time[2:9], df_20_bac$Surv[2:9]-df_20_bac$Std.error[2:9], df_20_bac$Time[2:9],
       df_20_bac$Surv[2:9]+df_20_bac$Std.error[2:9], length=0.05, angle=90, code=3, col = 'black')
points(df_20_meta$Time[2:9],df_20_meta$Surv[2:9], pch = 1, cex = 1.2)
lines(df_20_meta$Time,df_20_meta$Surv, lty = 2)
arrows(df_20_meta$Time, df_20_meta$Surv-df_20_meta$Std.error, df_20_meta$Time, df_20_meta$Surv+df_20_me
```

```
## Warning in arrows(df_20_meta$Time, df_20_meta$Surv -
## df_20_meta$Std.error, : zero-length arrow is of indeterminate angle and so
## skipped
```

Comparison between fly gut 20 without and with living bacteria

