# Clinical Trials & Survival Analysis

## Micro Session 3, Friday 27 April, 10–12, Computer room 71

This session is an opportunity to try some of the practical exercises on Survival Data contained on the task sheets and more substantial exercises. Some of the data sets are on the course web page at http://nickfieller.staff.shef.ac.uk/tampere/index.html and you can download them (recommended) or open them directly in **R** or SPSS.

## Regression with Survival Data in SPSS

The facilities for regression with survival data in SPSS are found through the menus in Analyze>Survival>Cox Regression. SPSS is only be able to perform Cox proportional hazards regression. However, it is still useful as a first step to produce some Kaplan-Meier plots of the survival curves for different levels of factors you expect to be important (e.g. treatment groups). However if there are other covariates in addition to the important factor the survival curve produced for separate groups is in some sense 'averaged' over the other covariates.

1. Download the file `prostatic.sav` from the course web page and read it into SPSS.
2. Produce Kaplan-Meier plots for the two treatment groups & note which of the treatment groups seems to have better survival prospects.
3. Next open the Analyze>Survival>Cox Regression dialogues box. Enter Survival Time and Status in the usual way. Enter the other variables into the covariates box.
4. Next, open the sub-dialogue box Categorical and enter Treatment as a categorical covariate. Note that you have the option of using the first or the last category as the reference category and I leave it to you to experiment with the effect of changing them.
5. Open the sub-dialogue box Plots and tick some of these, perhaps Survival and log minus log. You may also ask for separate lines for some of the covariates. First try separate lines for Treatment and then you can experiment with choosing others. Compare the survival plots for different treatment levels with those obtained in step 2.
6. Another survival data set with several covariates is `methrex.sav` and you can also use Cox Regression when you just have two groups. but do not spend too much time on this if you are happy with the SPSS facilities and instead move on to the section on **R**.

## Regression with Survival Data inR

Start **R** and change the working director to the folder where you will store the Rdata files (see the note and solutions for the first Micro Lab session). Download file `prostatic.Rdata`. The facilities for survival analysis in **R** are in a special library called survival so the first step is always to open this library.

1. Type `getwd()` to make sure your working directory is where your data are stored. Use Windows Explorer to check the data file `prostatic.Rdata` is there (or you can use `dir()` from within **R**).
2. Type `library(survival)`
3. Type `load("prostatic.Rdata")` and then `attach(prostatic)` to make the individual variables accessible to **R** and then `prostatic[1:5,]` to list the first five lines of the data to discover the names of the variables
4. Next, create a 'survival object' with the command `Surv()` (note the first capital letter) and call the object `prostatic.sv` (or a name of your choice): `prostatic.sv<- Surv(Survival.Time,Status)`
5. Type `prost.coxph<-coxph(prostatic.sv~Treatment+Tumour.Size)` to perform a Cox regression and store the results in `prost.coxph`.
6. Type `summary(prost.coxph)` to list the results of the analysis. Try including more of the covariates and trying dropping some of them.
7. Noting the use of the up arrow ↑ to recall previous commands and then editing them type `prost.exp<-survreg(prostatic.sv~Treatment+Tumour.Size, dist="exponential")` to perform an exponential regression and store the results in `prost.exp` followed by `summary(prost.exp)` to examine the results of the analysis.
8. Type `prost.wei<-survreg(prostatic.sv~Treatment+Tumour.Size, dist="weibull")` to perform a Weibull regression and store the results in prost.wei followed by `summary(prost.wei)` to examine the results of the analysis.
9. Look at the results of the three different regression models in terms of the p-values for the key covariates. Are they very different?
10. For the model `prost.coxph` try producing log minus log plots using the `strata()` command as described in §4.4.7.3 on P105. If you want to add titles and axis labels to the plot use `title()` (type `help(title)` to find out how to do this.