Experiment 1–Longevity data

## Analysis of longevity data

Done at the same time as the flight cylinder assay (previous script). Same-sex groups of 10 moths (all male or all female) in a cage were examined on a daily basis, and the number dead since the previous interval and the number of survivors were recorded. These data are summarized using Kaplan-Meier survivorship plots and Cox proportional hazard analysis. The data were recorded as groups, but the packages used here require one record per moth.

Since around a third of the moths examined were alive at the end of the observation period, a mean longevity has limited meaning while an estimate of the median longevity is arguably more informative. The package used to generate the survival curve is also used to get an estimate of median survival for each dose with a 95% confidence interval.

library(readxl)  
library(dplyr)  
df\_lngvt <- read\_excel("Dataset.xlsx", sheet = "nowcox")  
df\_lngvt <- df\_lngvt %>%   
 rename(day = dor,  
 gy = dose)  
df\_lngvt

## # A tibble: 420 × 7  
## rep day sex gy mort alive total  
## <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 1 1 f 0 0 10 10  
## 2 1 2 f 0 1 9 10  
## 3 1 3 f 0 1 8 10  
## 4 1 4 f 0 0 8 10  
## 5 1 5 f 0 2 6 10  
## 6 1 6 f 0 3 3 10  
## 7 1 7 f 0 3 0 10  
## 8 2 1 f 0 0 10 10  
## 9 2 2 f 0 0 10 10  
## 10 2 3 f 0 2 8 10  
## # ℹ 410 more rows

## Recode from aggregated to case format

First using the dplyr and tidyr package to creat 1 row for each for death records

library(tidyr)  
  
# For rows with deaths, create an observation per death using tidyr's uncount()  
events <- df\_lngvt %>%  
 filter(mort > 0) %>% # only rows where at least one death occurred  
 mutate(time = day, event = 1) %>% # assign time and indicate event = 1 (death)  
 select(rep, day, sex, gy, time, mort, event) %>%   
 uncount(weights = mort) # expand each row 'mort' times  
  
# Preview event data:  
events

## # A tibble: 429 × 6  
## rep day sex gy time event  
## <dbl> <dbl> <chr> <dbl> <dbl> <dbl>  
## 1 1 2 f 0 2 1  
## 2 1 3 f 0 3 1  
## 3 1 5 f 0 5 1  
## 4 1 5 f 0 5 1  
## 5 1 6 f 0 6 1  
## 6 1 6 f 0 6 1  
## 7 1 6 f 0 6 1  
## 8 1 7 f 0 7 1  
## 9 1 7 f 0 7 1  
## 10 1 7 f 0 7 1  
## # ℹ 419 more rows

All moths still alive on the last day of observation (day 7) are recorded as censored on that day.

censored <- df\_lngvt %>%  
 group\_by(rep, sex, gy) %>%  
 slice\_tail(n = 1) %>% # select the final monitoring interval for each group  
 ungroup() %>%   
 filter(alive > 0) %>% # only if there are survivors  
 mutate(time = day, event = 0) %>% # assign time and indicate event = 0 (censored)  
 select(rep, day, sex, gy, time, event, alive) %>%  
 uncount(weights = alive) # expand each row 'alive' times  
  
# Preview censored data:  
censored

## # A tibble: 171 × 6  
## rep day sex gy time event  
## <dbl> <dbl> <chr> <dbl> <dbl> <dbl>  
## 1 1 7 f 100 7 0  
## 2 1 7 f 100 7 0  
## 3 1 7 f 150 7 0  
## 4 1 7 f 250 7 0  
## 5 1 7 f 250 7 0  
## 6 1 7 f 250 7 0  
## 7 1 7 f 250 7 0  
## 8 1 7 m 0 7 0  
## 9 1 7 m 0 7 0  
## 10 1 7 m 0 7 0  
## # ℹ 161 more rows

Now combine the events and censored observations

# Combine events and censored observations  
indiv\_data <- bind\_rows(events, censored)  
  
# (Optional) Order the data, if desired:  
indiv\_data <- indiv\_data %>% arrange(rep, sex, gy, time)  
  
# Make Sex and Dose (Gy) explicitly factors  
indiv\_data$sex <- as.factor(indiv\_data$sex)  
indiv\_data$gy <- as.factor(indiv\_data$gy)  
  
# One last look  
indiv\_data

## # A tibble: 600 × 6  
## rep day sex gy time event  
## <dbl> <dbl> <fct> <fct> <dbl> <dbl>  
## 1 1 2 f 0 2 1  
## 2 1 3 f 0 3 1  
## 3 1 5 f 0 5 1  
## 4 1 5 f 0 5 1  
## 5 1 6 f 0 6 1  
## 6 1 6 f 0 6 1  
## 7 1 6 f 0 6 1  
## 8 1 7 f 0 7 1  
## 9 1 7 f 0 7 1  
## 10 1 7 f 0 7 1  
## # ℹ 590 more rows

Note here that event = 1 mean “event happened” (moth died), and event = 0 means moth was last seen alive

## Use the Cox Proportional Hazards model to evaluate impact

The R package ‘Survival’ will be used to implement the Cox model and evaluate the impact of sex and irradiation

library(survival)  
  
# Create the survival object  
surv\_obj <- with(indiv\_data, Surv(time = time, event = event))  
  
# Fit the Cox model (adjust predictors as appropriate)  
cox\_fit <- coxph(surv\_obj ~ sex + gy, data = indiv\_data)  
  
# View parameter estimates for the printed model  
summary(cox\_fit)

## Call:  
## coxph(formula = surv\_obj ~ sex + gy, data = indiv\_data)  
##   
## n= 600, number of events= 429   
##   
## coef exp(coef) se(coef) z Pr(>|z|)   
## sexm -0.183926 0.831997 0.096956 -1.897 0.0578 .  
## gy100 0.081795 1.085233 0.168802 0.485 0.6280   
## gy150 -0.008396 0.991639 0.170342 -0.049 0.9607   
## gy200 0.093471 1.097978 0.169682 0.551 0.5817   
## gy250 0.145092 1.156146 0.169137 0.858 0.3910   
## gy300 0.352740 1.422961 0.163446 2.158 0.0309 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## sexm 0.8320 1.2019 0.6880 1.006  
## gy100 1.0852 0.9215 0.7795 1.511  
## gy150 0.9916 1.0084 0.7102 1.385  
## gy200 1.0980 0.9108 0.7873 1.531  
## gy250 1.1561 0.8649 0.8299 1.611  
## gy300 1.4230 0.7028 1.0329 1.960  
##   
## Concordance= 0.555 (se = 0.016 )  
## Likelihood ratio test= 10.3 on 6 df, p=0.1  
## Wald test = 10.57 on 6 df, p=0.1  
## Score (logrank) test = 10.63 on 6 df, p=0.1

The Likelihood ratio and the Wald test provide similar results, but the former would be considered more robust for small data sets while the latter would be considered more informative for large data sets. The P value of 0.1 and the concordance of 0.555 indicate that the model has not detected important effects. We can also examine the model using a type II ANOVA.

library(car)  
print(Anova(cox\_fit))

## Analysis of Deviance Table (Type II tests)  
## LR Chisq Df Pr(>Chisq)   
## sex 3.6044 1 0.05763 .  
## gy 6.4243 5 0.26709   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The type II analysis provides a simple and robust parameter estimates indicating “not significant” (or an almost significant effect of sex, if one wishes to make that argument).

Another way to examine the current model is to generate estimates of median longevity. When the sexes are examined separately, the upper confidence cannot be estimated in some cases.

sfit\_2way <- survfit(surv\_obj ~ sex + gy, data = indiv\_data)  
print(sfit\_2way)

## Call: survfit(formula = surv\_obj ~ sex + gy, data = indiv\_data)  
##   
## n events median 0.95LCL 0.95UCL  
## sex=f, gy=0 50 43 6 6 7  
## sex=f, gy=100 50 33 7 6 NA  
## sex=f, gy=150 50 36 6 6 7  
## sex=f, gy=200 50 38 6 5 7  
## sex=f, gy=250 50 34 6 5 7  
## sex=f, gy=300 50 41 6 5 7  
## sex=m, gy=0 50 27 7 7 NA  
## sex=m, gy=100 50 38 6 6 7  
## sex=m, gy=150 50 32 7 6 NA  
## sex=m, gy=200 50 31 7 6 NA  
## sex=m, gy=250 50 36 6 6 7  
## sex=m, gy=300 50 40 6 5 7

## Examination of irradiation only (sexes pooled)

The primary purpose here is to get a more robust estimate of the 95% confidence interval for the median estimates.

cox\_fit\_dose\_only <- coxph(surv\_obj ~ gy, data = indiv\_data)  
Anova(cox\_fit\_dose\_only)

## Analysis of Deviance Table  
## Cox model: response is surv\_obj  
## Terms added sequentially (first to last)  
##   
## loglik Chisq Df Pr(>|Chi|)  
## NULL -2530.6   
## gy -2527.2 6.6951 5 0.2443

Still not significant, in case anyone wondered.

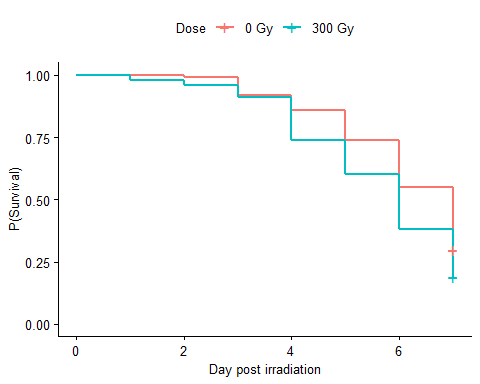
km\_fit <- survfit(surv\_obj ~ gy, data = indiv\_data)  
print(km\_fit)

## Call: survfit(formula = surv\_obj ~ gy, data = indiv\_data)  
##   
## n events median 0.95LCL 0.95UCL  
## gy=0 100 70 7 6 7  
## gy=100 100 71 6 6 7  
## gy=150 100 68 7 6 7  
## gy=200 100 69 6 6 7  
## gy=250 100 70 6 6 7  
## gy=300 100 81 6 6 6

Using the pooled sexes provides a larger sample and allows an upper 95% CL for all doses, so this is probably better for reporting.

For visualization, presenting survival curves for only the 0 and 300 Gy levels seems to offer a cleaner plot.

library(survminer) # Required for surv\_fit() function  
# Create set with only 0 and 300 Gy  
indiv\_data2 <- indiv\_data %>%   
 filter(gy == 0 | gy == 300)  
# Drop now unused levels  
indiv\_data2$gy <- droplevels(indiv\_data2$gy)  
  
# Create another survival object  
surv\_obj2 <- with(indiv\_data2, Surv(time = time, event = event))  
  
# Fit Kaplan-Meier Curve  
km\_fit2 <- surv\_fit(surv\_obj2 ~ gy, data = indiv\_data2)  
  
# Plot survival curves  
p1 <- ggsurvplot(km\_fit2, data = indiv\_data2, pval = FALSE,  
 legend.title = "Dose", legend.labs = c("0 Gy","300 Gy"))   
  
# Format the image to be saved  
p1$plot <- p1$plot +   
 xlab("Day post irradiation") +  
 ylab("P(Survival)") +  
 theme(axis.text.x = element\_text(color = "black", size = 10),   
 axis.text.y = element\_text(color = "black", size = 10),  
 axis.title.x = element\_text(color = "black", size = 10),  
 axis.title.y = element\_text(color = "black", size = 10),  
 legend.title = element\_text(color = "black", size = 10),  
 legend.text = element\_text(color = "black", size = 10))  
  
# Save the image to local storage, then display the image  
ggsave(filename = "survival\_curve.jpg", plot = p1$plot,  
 width = 2.83, height = 1.89, dpi = 300)  
  
p1



## Examination of irradiation only (sexes pooled)

In case we want to visualize the data in this manner, we can repeat the process in the previous section

# Since there are only two sexes, the earlier data set and surv\_obj will work  
  
# Fit Kaplan-Meier Curve  
km\_fit3 <- surv\_fit(surv\_obj ~ sex, data = indiv\_data)  
  
# Plot survival curves  
p2 <- ggsurvplot(km\_fit3, data = indiv\_data2, pval = FALSE,  
 legend.title = "Sex", legend.labs = c("Female","Male"))   
  
# Format the image to be saved  
p2$plot <- p2$plot +   
 xlab("Day post irradiation") +  
 ylab("P(Survival)") +  
 theme(axis.text.x = element\_text(color = "black", size = 10),   
 axis.text.y = element\_text(color = "black", size = 10),  
 axis.title.x = element\_text(color = "black", size = 10),  
 axis.title.y = element\_text(color = "black", size = 10),  
 legend.title = element\_text(color = "black", size = 10),  
 legend.text = element\_text(color = "black", size = 10))  
  
# Save the image to local storage, then display the image  
ggsave(filename = "survival\_curve\_sex.jpg", plot = p2$plot,  
 width = 2.83, height = 1.89, dpi = 300)  
  
p2

