

## Solution to Exercise 4: Multivariable analysis in R part 2: Cox proportional hazard model

At the end of this exercise you should be able to:

- Use stepwise logistic regression to determine risk factors for a bacteriologically adverse outcome among patients with initial fluoroquinolone resistance who neither defaulted nor died
- Summarize your findings in a table

### *Solution:*

The `e_ex04.r` script from Exercise 4:

```
# Exercise 4: Multivariable analysis in R part 2: Cox proportional hazard
model

library(survival)
rm(list=ls())
e_ex04a <- read.table("e_ex02_02.dat")
names(e_ex04a)

e_ex04b <- e_ex04a[c(1:4, 8)]

attach(e_ex04b)
e_ex04b$fq03[fq04 == "Susceptible"] <- "1-Susceptible"
e_ex04b$fq03[fq04 == "Low-level resistance"] <- "2-Low-level resistance"
e_ex04b$fq03[fq04 == "High-level resistance"] <- "3-High-level resistance"
e_ex04b$out02[outcome02 == "Success"] <- 0
e_ex04b$out02[outcome02 == "Failure"] <- 1
detach(e_ex04b)
names(e_ex04b)
e_ex04c <- e_ex04b[c(1, 3:4, 6:7)]
attach(e_ex04c)
names(e_ex04c)

# Reproducing the Kaplan-Meier lifetable analysis from EpiData Analysis
fit1 <- survfit(formula=Surv(totobstime, out02) ~ fq03, data=e_ex04c)
plot(fit1, col = c(1, 2, 4), ymin=0.45)
# colors: 1=black; 2=red, 4=blue
summary(fit1, times = seq(0, 2000, 30))

# Cox proportional hazard model
mdrcox <- coxph(Surv(totobstime, out02) ~ factor(fq03) + factor(sex) + age,
  data=e_ex04c)
summary(mdrcox)
cox.zph(mdrcox)

mdrcox <- coxph(Surv(totobstime, out02) ~ strata(fq03) + factor(sex) + age,
  data=e_ex04c)
summary(mdrcox)
cox.zph(mdrcox)
plot(cox.zph(mdrcox))
```

```

p <- plot(survfit(mdrcox), ylim=c(.45, 1), xlab="Days since treatment start",
  mark.time=F, ylab="Proportion without adverse event", col=c(1, 2, 4),
  main="Cox proportional hazard model by initial fluoroquinolone
  resistance")
print(p)

mdrcox2 <- survfit(mdrcox)
summary(mdrcox2, times = seq(0, 3000, 30))
summary(mdrcox2, times = seq(0, 5, 1))

```

The `e_ex04_solution.r` script for the task:

```

# Exercise 4 Solution: Logistic regression to identify risk factors for
  failure

library(survival)
rm(list=ls())
e_ex04a_task <- read.table("e_ex02_03.dat")
names(e_ex04a_task)

e_ex04b_task <- e_ex04a_task[c(1:3, 8:11)]

attach(e_ex04b_task)
e_ex04b_task$fq03[fq04 == "Low-level resistance"] <- "1-Low-level
  resistance"
e_ex04b_task$fq03[fq04 == "High-level resistance"] <- "2-High-level
  resistance"
e_ex04b_task$out02[outcome02 == "Success"] <- 0
e_ex04b_task$out02[outcome02 == "Failure"] <- 1
detach(e_ex04b_task)
names(e_ex04b_task)
e_ex04c_task <- e_ex04b_task[c(1, 3, 5:9)]
attach(e_ex04c_task)
names(e_ex04c_task)

# Logistic regression
mylogit<- glm(formula = out02 ~ factor(fq03) + factor(sex) + factor(pza02) +
  factor(kmy02) + factor(pth02) + age, data=e_ex04c_task, family=binomial)
summary(mylogit)

# Remove KMY02
mylogit<- glm(formula = out02 ~ factor(fq03) + factor(sex) + factor(pza02) +
  factor(pth02) + age, data=e_ex04c_task, family=binomial)
summary(mylogit)

# Remove PTH02
mylogit<- glm(formula = out02 ~ factor(fq03) + factor(sex) + factor(pza02) +
  age, data=e_ex04c_task, family=binomial)
summary(mylogit)

# Remove SEX
mylogit<- glm(formula = out02 ~ factor(fq03) + factor(pza02) + age,
  data=e_ex04c_task, family=binomial)
summary(mylogit)

# Remove AGE

```

```

mylogit<- glm(formula = out02 ~ factor(fq03) + factor(pza02),
  data=e_ex04c_task, family=binomial)
summary(mylogit)

# => PZA just not significant with p=0.5468, leave it in
# Final model
mylogit<- glm(formula = out02 ~ factor(fq03) + factor(pza02),
  data=e_ex04c_task, family=binomial)
summary(mylogit)
lreg.or <-exp(cbind(OR = coef(mylogit), confint(mylogit)))
round(lreg.or, digits=4)

load("c:/epidata_course/woolf.test.r")
tab <- table(fq03, out02, pza02)
woolf.test(tab)
# P=0.936 => no interaction

# Univariate analysis
load("c:/epidata_course/tab2by2.r")
tab2by2(fq03, out02)
tab2by2(pza02, out02)

```

	Successful		Not successful		Total	Crude odds ratio			Adjusted odds ratio*		
	n	n	n	%		Point	CI low	CI high	Point	CI low	CI high
Total	44	9	17.0	53							
Fluoroquinolone resistance											
Low level	30	1	3.2	31	Ref						
High level	14	8	36.4	22	17.1	2.0	150.7	13.0	1.9	262.1	0.025
Pyrazinamide resistance											
Not known *	27	1	3.6	28	Ref						
Resistant	17	8	32.0	25	12.7	1.5	110.8	9.2	1.3	187.0	0.055

CI       95% confidence interval  
\*        By logistic regression  
\*\*       Susceptible or not tested