FHS_RMST.Rmd

Grace Rade

7/5/2022

```
library(survival)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## Warning: package 'tidyr' was built under R version 4.0.5

## Warning: package 'readr' was built under R version 4.0.5

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(rio)
library(survminer)
```

```
## Loading required package: ggpubr
```

```
library(haven)
library(survRM2)



all_data <- read_sas('alldata2022.sas7bdat copy')
cvd_status <- read_dta("FHS 2154 - datasets/dr281_vr_survcvd_2014_a_1023s.dta") %>% select(IDTYPE, ranic
last_contact <- read_dta("FHS 2154 - datasets/dr281_vr_survdth_2014_a_1025s.dta")
all_data2 <- all_data %>%
  inner_join(last_contact, by = c("IDTYPE", "ranid"))
all_data2 <- all_data2 %>%
  inner_join(cvd_status, by = c("IDTYPE", "ranid"))

all_data2$LASTCON[is.na(all_data2$LASTCON)] <- all_data2$DATEDTH[is.na(all_data2$LASTCON)] ## replacing

all_data2 <- all_data2 %>%
  mutate(dead = if_else(LASTCON > 5000, 1, 0)) ## alive = 1, dead = 0
```

```r
all_data2$LASTCON[all_data2$LASTCON > 5000] <- 5000


quantile(all_data2$TXB2_M) ## finding the quantiles for thromboxane levels
```

```
##        0%       25%       50%       75%      100%
##   156.300   919.375  1997.950  4303.825 20000.000
```

```r
data_rmst <- all_data2 %>%
  select(IDTYPE, ranid, TXB2_M, TXB2_M_I, H010, EVENT, chfyn, chfdays, chfdate, afibyn, afibdate, afibda
  mutate(TXB2_quart = if_else(TXB2_M >= 4303.825, "Q4", NULL))

data_rmst$TXB2_quart[data_rmst$TXB2_M < 4303.825] <- "Q3"
data_rmst$TXB2_quart[data_rmst$TXB2_M < 1997.950] <- "Q2"
data_rmst$TXB2_quart[data_rmst$TXB2_M < 919.375] <- "Q1"

data_rmst <- data_rmst %>%
  mutate(TXB2_quant = if_else(TXB2_M >= 4303.825, "Q4", "Q1-Q3"), events = if_else(CHDDEATH == 0, 1, 0)
  rename(aspirin_use = H010)

data_rmst$events[data_rmst$CVDDEATH == 0] <- 2
data_rmst$events[data_rmst$chfyn == 0] <- 3
data_rmst$events[data_rmst$LASTCON == 5000] <- 0

export(data_rmst, "rmst_data.csv")
## creating a new dataset with thromboxane quartile variables (one with categories of all 4 quartiles,

## as far as i can tell, there is no cancer variable
```

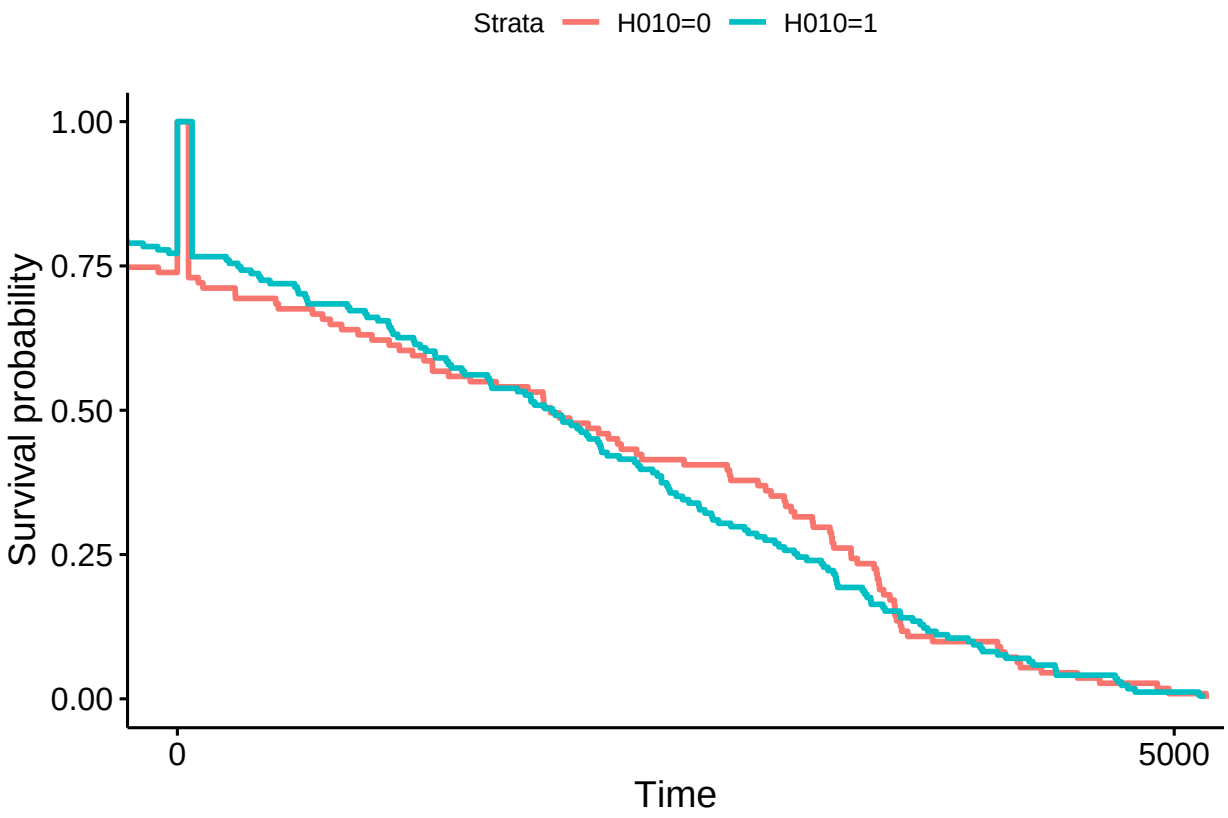
H010:

0 = no aspirin use

1 = aspirin use ## RMST

```r
model1 <- survfit(Surv(chfdays) ~ H010, data = all_data2)

ggsurvplot(model1, data = all_data2)
```




```
# trying to get R to calculate RMST is not something I've been successful at, disregard this curve

# the bump at the beginning of both curves confuses me
```