Cox model

Fuyu Guo; Jueyi Liu; Xinyi qian

4/27/2022

This RMD file is corresponding to the Cox models in the main analysis following the second conceptual framework (the study is simplied as a cross-sectional study and time scale is time since entry)

Load data and packages

```
library(tidyverse)
```

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

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
```

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

```
library(survival)
dta <- readRDS("clean_data_0425.rds")


dta$fu_time <- ifelse(!is.na(dta$DIE_YR),  dta$DIE_YR - dta$EXAM_YR,
                      dta$LAST_YR-dta$EXAM_YR)
# exclude people whose follow-up time is 0
dta <- dta %>% filter(fu_time > 0)
```

Main analysis

Model 1: Crude analysis

```
dta$BOOZE_ord <- dta$BOOZE_q %>% as.numeric()
# categorical version
fit1_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q, data = dta)
fit1_q %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q, data = dta)
##
##   n= 9190, number of events= 552
##
##                    coef exp(coef) se(coef)      z Pr(>|z|)
```

```
## BOOZE_q(0,0.5] -0.05525   0.94625  0.15213 -0.363    0.716
## BOOZE_q(0.5,2] -0.17802   0.83693  0.12506 -1.423    0.155
## BOOZE_q(2,77]   0.14322   1.15398  0.09963  1.437    0.151
##
##                exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]    0.9462     1.0568    0.7023     1.275
## BOOZE_q(0.5,2]    0.8369     1.1948    0.6550     1.069
## BOOZE_q(2,77]     1.1540     0.8666    0.9493     1.403
##
## Concordance= 0.532  (se = 0.012 )
## Likelihood ratio test= 6.57  on 3 df,   p=0.09
## Wald test            = 6.5  on 3 df,   p=0.09
## Score (logrank) test = 6.53  on 3 df,   p=0.09
```

```
# continuous version
fit1_con <- coxph(Surv(fu_time, cancer_death) ~ BOOZE, data = dta)
fit1_con %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE, data = dta)
##
##   n= 9190, number of events= 552
##
##           coef exp(coef) se(coef)     z Pr(>|z|)
## BOOZE 0.029295  1.029728 0.006454 4.539 5.66e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##       exp(coef) exp(-coef) lower .95 upper .95
## BOOZE      1.03     0.9711     1.017     1.043
##
## Concordance= 0.515  (se = 0.013 )
## Likelihood ratio test= 15.91  on 1 df,   p=7e-05
## Wald test            = 20.6  on 1 df,   p=6e-06
## Score (logrank) test = 20.32  on 1 df,   p=7e-06
```

```
# ordinal version for p-value for trend
fit1_ord <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_ord, data = dta)
fit1_ord %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_ord, data = dta)
##
##   n= 9190, number of events= 552
##
##              coef exp(coef) se(coef)     z Pr(>|z|)
## BOOZE_ord 0.03206   1.03258  0.03323 0.965    0.335
##
##           exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_ord     1.033     0.9684    0.9675     1.102
##
## Concordance= 0.509  (se = 0.012 )
## Likelihood ratio test= 0.93  on 1 df,   p=0.3
## Wald test            = 0.93  on 1 df,   p=0.3
## Score (logrank) test = 0.93  on 1 df,   p=0.3
```

Model 2: Adjusting for age

```
# categorical version
fit2_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS, data = dta)
fit2_q %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS,
##     data = dta)
##
##   n= 9190, number of events= 552
##
##                     coef exp(coef) se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5] 0.105880  1.111688 0.152333  0.695    0.487
## BOOZE_q(0.5,2] 0.123222  1.131135 0.125812  0.979    0.327
## BOOZE_q(2,77]  0.472044  1.603267 0.101003  4.674 2.96e-06 ***
## AGEYRS         0.070455  1.072996 0.004475 15.744  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]     1.112     0.8995    0.8247     1.498
## BOOZE_q(0.5,2]     1.131     0.8841    0.8839     1.447
## BOOZE_q(2,77]      1.603     0.6237    1.3153     1.954
## AGEYRS             1.073     0.9320    1.0636     1.082
##
## Concordance= 0.715  (se = 0.01 )
## Likelihood ratio test= 332.5  on 4 df,   p=<2e-16
## Wald test            = 255.1  on 4 df,   p=<2e-16
## Score (logrank) test = 291.1  on 4 df,   p=<2e-16
```

```
# continuous version
fit2_con <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS, data = dta)
fit2_con %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS,
##     data = dta)
##
##   n= 9190, number of events= 552
##
##             coef exp(coef) se(coef)      z Pr(>|z|)
## BOOZE   0.038740  1.039500 0.006129  6.321  2.6e-10 ***
## AGEYRS 0.069202  1.071653 0.004416 15.671  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##        exp(coef) exp(-coef) lower .95 upper .95
## BOOZE      1.040     0.9620     1.027     1.052
## AGEYRS     1.072     0.9331     1.062     1.081
##
## Concordance= 0.719  (se = 0.01 )
## Likelihood ratio test= 339.9  on 2 df,   p=<2e-16
## Wald test            = 269.9  on 2 df,   p=<2e-16
## Score (logrank) test = 306.5  on 2 df,   p=<2e-16
```

```r
# ordinal version for p-value for trend
fit2_ord <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_ord + AGEYRS, data = dta)
fit2_ord %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_ord + AGEYRS,
##     data = dta)
##
##   n= 9190, number of events= 552
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## BOOZE_ord 0.14660   1.15789  0.03342  4.386 1.15e-05 ***
## AGEYRS    0.07063   1.07319  0.00447 15.800  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_ord     1.158     0.8636     1.084     1.236
## AGEYRS        1.073     0.9318     1.064     1.083
##
## Concordance= 0.715  (se = 0.01 )
## Likelihood ratio test= 329.9  on 2 df,   p=<2e-16
## Wald test            = 252.7  on 2 df,   p=<2e-16
## Score (logrank) test = 288.9  on 2 df,   p=<2e-16
```

Model 3: MV-adjusted analysis

```r
# categorical version
fit3_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS +
                      SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) + AVGSMK + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
fit3_q %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS +
##     SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                         coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5]     0.062381  1.064368  0.153886  0.405   0.6852
## BOOZE_q(0.5,2]    -0.022538  0.977714  0.130281 -0.173   0.8627
## BOOZE_q(2,77]      0.216482  1.241701  0.111603  1.940   0.0524 .
## AGEYRS             0.078181  1.081318  0.004890 15.988  < 2e-16 ***
## SEX               -0.533047  0.586814  0.101146 -5.270 1.36e-07 ***
## as.factor(RACE)2  -0.065050  0.937021  0.154848 -0.420   0.6744
## as.factor(RACE)3  -0.466283  0.627330  0.416720 -1.119   0.2632
## GRADES            -0.014233  0.985868  0.012522 -1.137   0.2557
## as.factor(SIZE)2   0.163250  1.177331  0.168056  0.971   0.3313
## as.factor(SIZE)3   0.353475  1.424007  0.166891  2.118   0.0342 *
## as.factor(SIZE)4   0.006977  1.007001  0.194056  0.036   0.9713
## as.factor(SIZE)5  -0.095880  0.908573  0.322831 -0.297   0.7665
```

```
## as.factor(SIZE)6   0.048661  1.049864  0.306838  0.159   0.8740
## as.factor(SIZE)7  -0.049226  0.951966  0.261103 -0.189   0.8505
## as.factor(SIZE)8   1.769602  5.868518  0.738832  2.395   0.0166 *
## as.factor(MARRY)3  0.024005  1.024296  0.129198  0.186   0.8526
## as.factor(MARRY)4  0.231038  1.259907  0.182961  1.263   0.2067
## as.factor(MARRY)5 -0.131390  0.876875  0.328963 -0.399   0.6896
## as.factor(MARRY)6 -0.157045  0.854666  0.216598 -0.725   0.4684
## as.factor(MARRY)8  0.086953  1.090846  0.712647  0.122   0.9029
## AVGSMK             0.025014  1.025329  0.002750  9.097  < 2e-16 ***
## as.factor(SMSA)2  -0.198918  0.819617  0.134333 -1.481   0.1387
## as.factor(SMSA)4  -0.077507  0.925421  0.226029 -0.343   0.7317
## URBAN             -1.775186  0.169452  0.745544 -2.381   0.0173 *
## HTN_REP           -0.053813  0.947609  0.091773 -0.586   0.5576
## RBC                0.004711  1.004723  0.061802  0.076   0.9392
## DRMI              -0.095295  0.909104  0.171634 -0.555   0.5787
## DIAB              -0.118480  0.888269  0.185637 -0.638   0.5233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]       1.0644     0.9395    0.7872    1.4391
## BOOZE_q(0.5,2]       0.9777     1.0228    0.7574    1.2621
## BOOZE_q(2,77]        1.2417     0.8053    0.9977    1.5453
## AGEYRS               1.0813     0.9248    1.0710    1.0917
## SEX                  0.5868     1.7041    0.4813    0.7155
## as.factor(RACE)2     0.9370     1.0672    0.6917    1.2693
## as.factor(RACE)3     0.6273     1.5941    0.2772    1.4197
## GRADES               0.9859     1.0143    0.9620    1.0104
## as.factor(SIZE)2     1.1773     0.8494    0.8469    1.6366
## as.factor(SIZE)3     1.4240     0.7022    1.0267    1.9750
## as.factor(SIZE)4     1.0070     0.9930    0.6884    1.4730
## as.factor(SIZE)5     0.9086     1.1006    0.4826    1.7106
## as.factor(SIZE)6     1.0499     0.9525    0.5754    1.9156
## as.factor(SIZE)7     0.9520     1.0505    0.5707    1.5881
## as.factor(SIZE)8     5.8685     0.1704    1.3792   24.9703
## as.factor(MARRY)3    1.0243     0.9763    0.7952    1.3195
## as.factor(MARRY)4    1.2599     0.7937    0.8802    1.8033
## as.factor(MARRY)5    0.8769     1.1404    0.4602    1.6709
## as.factor(MARRY)6    0.8547     1.1700    0.5590    1.3067
## as.factor(MARRY)8    1.0908     0.9167    0.2699    4.4093
## AVGSMK               1.0253     0.9753    1.0198    1.0309
## as.factor(SMSA)2     0.8196     1.2201    0.6299    1.0665
## as.factor(SMSA)4     0.9254     1.0806    0.5942    1.4412
## URBAN                0.1695     5.9014    0.0393    0.7306
## HTN_REP              0.9476     1.0553    0.7916    1.1343
## RBC                  1.0047     0.9953    0.8901    1.1341
## DRMI                 0.9091     1.1000    0.6494    1.2726
## DIAB                 0.8883     1.1258    0.6173    1.2781
##
## Concordance= 0.753  (se = 0.01 )
## Likelihood ratio test= 471.8  on 28 df,   p=<2e-16
## Wald test            = 402.1  on 28 df,   p=<2e-16
## Score (logrank) test = 428.7  on 28 df,   p=<2e-16
```

```r
# continuous version
fit3_con <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS +
                      SEX + as.factor(RACE) + GRADES + as.factor(SIZE)+ as.factor(MARRY) + AVGSMK + as.
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
fit3_con %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS +
##     SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                         coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE                0.025297  1.025620  0.007071  3.578 0.000347 ***
## AGEYRS               0.078123  1.081255  0.004866 16.056  < 2e-16 ***
## SEX                 -0.509585  0.600745  0.100083 -5.092 3.55e-07 ***
## as.factor(RACE)2    -0.058130  0.943527  0.154639 -0.376 0.706985
## as.factor(RACE)3    -0.450551  0.637277  0.416578 -1.082 0.279451
## GRADES              -0.013815  0.986280  0.012375 -1.116 0.264277
## as.factor(SIZE)2     0.149436  1.161180  0.168014  0.889 0.373773
## as.factor(SIZE)3     0.343525  1.409909  0.166630  2.062 0.039245 *
## as.factor(SIZE)4     0.004909  1.004921  0.193847  0.025 0.979797
## as.factor(SIZE)5    -0.112964  0.893183  0.322738 -0.350 0.726326
## as.factor(SIZE)6     0.040942  1.041791  0.306999  0.133 0.893908
## as.factor(SIZE)7    -0.068461  0.933830  0.260874 -0.262 0.792990
## as.factor(SIZE)8     1.823447  6.193169  0.737773  2.472 0.013453 *
## as.factor(MARRY)3    0.016787  1.016929  0.129056  0.130 0.896505
## as.factor(MARRY)4    0.234022  1.263673  0.182686  1.281 0.200191
## as.factor(MARRY)5   -0.128498  0.879415  0.328669 -0.391 0.695822
## as.factor(MARRY)6   -0.159279  0.852758  0.216308 -0.736 0.461514
## as.factor(MARRY)8    0.095722  1.100453  0.712576  0.134 0.893140
## AVGSMK               0.024881  1.025193  0.002741  9.077  < 2e-16 ***
## as.factor(SMSA)2    -0.196148  0.821890  0.134305 -1.460 0.144162
## as.factor(SMSA)4    -0.059435  0.942297  0.225652 -0.263 0.792250
## URBAN               -1.844398  0.158121  0.744804 -2.476 0.013273 *
## HTN_REP             -0.057840  0.943801  0.091702 -0.631 0.528210
## RBC                  0.009358  1.009401  0.061458  0.152 0.878984
## DRMI                -0.086226  0.917387  0.171515 -0.503 0.615154
## DIAB                -0.111931  0.894106  0.185001 -0.605 0.545160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## BOOZE                  1.0256     0.9750   1.01150    1.0399
## AGEYRS                 1.0813     0.9249   1.07099    1.0916
## SEX                    0.6007     1.6646   0.49374    0.7309
## as.factor(RACE)2       0.9435     1.0599   0.69683    1.2776
## as.factor(RACE)3       0.6373     1.5692   0.28167    1.4418
## GRADES                 0.9863     1.0139   0.96265    1.0105
## as.factor(SIZE)2       1.1612     0.8612   0.83538    1.6140
## as.factor(SIZE)3       1.4099     0.7093   1.01708    1.9545
## as.factor(SIZE)4       1.0049     0.9951   0.68727    1.4694
```

```
## as.factor(SIZE)5     0.8932     1.1196   0.47449    1.6813
## as.factor(SIZE)6     1.0418     0.9599   0.57077    1.9015
## as.factor(SIZE)7     0.9338     1.0709   0.56003    1.5571
## as.factor(SIZE)8     6.1932     0.1615   1.45854   26.2970
## as.factor(MARRY)3    1.0169     0.9834   0.78966    1.3096
## as.factor(MARRY)4    1.2637     0.7913   0.88335    1.8077
## as.factor(MARRY)5    0.8794     1.1371   0.46178    1.6748
## as.factor(MARRY)6    0.8528     1.1727   0.55809    1.3030
## as.factor(MARRY)8    1.1005     0.9087   0.27229    4.4475
## AVGSMK               1.0252     0.9754   1.01970    1.0307
## as.factor(SMSA)2     0.8219     1.2167   0.63167    1.0694
## as.factor(SMSA)4     0.9423     1.0612   0.60550    1.4664
## URBAN                0.1581     6.3243   0.03673    0.6807
## HTN_REP              0.9438     1.0595   0.78854    1.1296
## RBC                  1.0094     0.9907   0.89485    1.1386
## DRMI                 0.9174     1.0901   0.65548    1.2839
## DIAB                 0.8941     1.1184   0.62218    1.2849
##
## Concordance= 0.755  (se = 0.01 )
## Likelihood ratio test= 477.7  on 26 df,   p=<2e-16
## Wald test            = 411.4  on 26 df,   p=<2e-16
## Score (logrank) test = 438.2  on 26 df,   p=<2e-16
```

```
# ordinal version for p-value for trend
fit3_ord <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_ord + AGEYRS +
                    SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) + AVGSMK + as
                    HTN_REP + RBC + DRMI + DIAB, data = dta)
fit3_ord %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_ord + AGEYRS +
##     SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                       coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE_ord          0.062151  1.064122  0.036847  1.687   0.0917 .
## AGEYRS             0.078365  1.081517  0.004886 16.039  < 2e-16 ***
## SEX               -0.538485  0.583632  0.100911 -5.336 9.49e-08 ***
## as.factor(RACE)2  -0.066634  0.935537  0.154802 -0.430   0.6669
## as.factor(RACE)3  -0.457329  0.632972  0.416624 -1.098   0.2723
## GRADES            -0.013534  0.986557  0.012497 -1.083   0.2788
## as.factor(SIZE)2   0.167929  1.182852  0.168016  0.999   0.3176
## as.factor(SIZE)3   0.360193  1.433606  0.166776  2.160   0.0308 *
## as.factor(SIZE)4   0.014382  1.014486  0.193901  0.074   0.9409
## as.factor(SIZE)5  -0.105015  0.900311  0.322890 -0.325   0.7450
## as.factor(SIZE)6   0.042319  1.043227  0.307031  0.138   0.8904
## as.factor(SIZE)7  -0.050068  0.951165  0.261082 -0.192   0.8479
## as.factor(SIZE)8   1.793805  6.012284  0.738198  2.430   0.0151 *
## as.factor(MARRY)3  0.025838  1.026174  0.129184  0.200   0.8415
## as.factor(MARRY)4  0.238907  1.269860  0.182786  1.307   0.1912
## as.factor(MARRY)5 -0.127184  0.880572  0.328802 -0.387   0.6989
## as.factor(MARRY)6 -0.159988  0.852154  0.216474 -0.739   0.4599
```

```
## as.factor(MARRY)8  0.086470  1.090319  0.712627  0.121   0.9034
## AVGSMK              0.024900  1.025212  0.002740  9.087  < 2e-16 ***
## as.factor(SMSA)2  -0.192976  0.824501  0.134187 -1.438   0.1504
## as.factor(SMSA)4  -0.064359  0.937668  0.225840 -0.285   0.7757
## URBAN             -1.800171  0.165271  0.744872 -2.417   0.0157 *
## HTN_REP           -0.051133  0.950153  0.091712 -0.558   0.5772
## RBC                0.005406  1.005421  0.061781  0.088   0.9303
## DRMI              -0.105163  0.900178  0.171519 -0.613   0.5398
## DIAB              -0.111261  0.894706  0.185551 -0.600   0.5488
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_ord            1.0641     0.9397   0.98998    1.1438
## AGEYRS               1.0815     0.9246   1.07121    1.0919
## SEX                  0.5836     1.7134   0.47890    0.7113
## as.factor(RACE)2     0.9355     1.0689   0.69071    1.2672
## as.factor(RACE)3     0.6330     1.5798   0.27974    1.4322
## GRADES               0.9866     1.0136   0.96269    1.0110
## as.factor(SIZE)2     1.1829     0.8454   0.85097    1.6442
## as.factor(SIZE)3     1.4336     0.6975   1.03388    1.9879
## as.factor(SIZE)4     1.0145     0.9857   0.69374    1.4835
## as.factor(SIZE)5     0.9003     1.1107   0.47813    1.6953
## as.factor(SIZE)6     1.0432     0.9586   0.57152    1.9042
## as.factor(SIZE)7     0.9512     1.0513   0.57019    1.5867
## as.factor(SIZE)8     6.0123     0.1663   1.41477   25.5502
## as.factor(MARRY)3    1.0262     0.9745   0.79663    1.3219
## as.factor(MARRY)4    1.2699     0.7875   0.88750    1.8170
## as.factor(MARRY)5    0.8806     1.1356   0.46226    1.6774
## as.factor(MARRY)6    0.8522     1.1735   0.55751    1.3025
## as.factor(MARRY)8    1.0903     0.9172   0.26975    4.4070
## AVGSMK               1.0252     0.9754   1.01972    1.0307
## as.factor(SMSA)2     0.8245     1.2129   0.63383    1.0725
## as.factor(SMSA)4     0.9377     1.0665   0.60230    1.4598
## URBAN                0.1653     6.0507   0.03838    0.7116
## HTN_REP              0.9502     1.0525   0.79383    1.1373
## RBC                  1.0054     0.9946   0.89076    1.1348
## DRMI                 0.9002     1.1109   0.64318    1.2599
## DIAB                 0.8947     1.1177   0.62193    1.2871
##
## Concordance= 0.753  (se = 0.01 )
## Likelihood ratio test= 469.8  on 26 df,   p=<2e-16
## Wald test            = 401.1  on 26 df,   p=<2e-16
## Score (logrank) test = 427.2  on 26 df,   p=<2e-16
```

Exploring effect modification by sex

Sex interaction terms in the Cox model

```
dta$SEX <- if_else(dta$SEX == 1, "Male", "Female")

# categorical version
fit4_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + BOOZE_q:SEX + AGEYRS +
```

```
                      SEX + as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) + AVGSMK + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
fit4_q %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + BOOZE_q:SEX +
##     AGEYRS + SEX + as.factor(RACE) + GRADES + as.factor(MARRY) +
##     as.factor(SIZE) + AVGSMK + as.factor(SMSA) + URBAN + HTN_REP +
##     RBC + DRMI + DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                              coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5]           0.302619  1.353399  0.207658  1.457   0.1450
## BOOZE_q(0.5,2]           0.286139  1.331278  0.190068  1.505   0.1322
## BOOZE_q(2,77]            0.463128  1.589037  0.183904  2.518   0.0118 *
## AGEYRS                   0.078086  1.081215  0.004893 15.959  < 2e-16 ***
## SEXMale                  0.783906  2.190009  0.139428  5.622 1.88e-08 ***
## as.factor(RACE)2        -0.066640  0.935532  0.154850 -0.430   0.6669
## as.factor(RACE)3        -0.467228  0.626737  0.416778 -1.121   0.2623
## GRADES                  -0.014812  0.985297  0.012532 -1.182   0.2372
## as.factor(MARRY)3        0.030567  1.031039  0.129223  0.237   0.8130
## as.factor(MARRY)4        0.225202  1.252575  0.183173  1.229   0.2189
## as.factor(MARRY)5       -0.134948  0.873761  0.328997 -0.410   0.6817
## as.factor(MARRY)6       -0.156478  0.855150  0.216471 -0.723   0.4698
## as.factor(MARRY)8        0.023324  1.023598  0.713275  0.033   0.9739
## as.factor(SIZE)2         0.167515  1.182363  0.168075  0.997   0.3189
## as.factor(SIZE)3         0.361694  1.435759  0.166878  2.167   0.0302 *
## as.factor(SIZE)4         0.014946  1.015058  0.194024  0.077   0.9386
## as.factor(SIZE)5        -0.105367  0.899994  0.323012 -0.326   0.7443
## as.factor(SIZE)6         0.071535  1.074155  0.307017  0.233   0.8158
## as.factor(SIZE)7        -0.056982  0.944611  0.260965 -0.218   0.8272
## as.factor(SIZE)8         1.756008  5.789279  0.738936  2.376   0.0175 *
## AVGSMK                   0.025052  1.025369  0.002761  9.073  < 2e-16 ***
## as.factor(SMSA)2        -0.198166  0.820234  0.134216 -1.476   0.1398
## as.factor(SMSA)4        -0.067451  0.934773  0.225930 -0.299   0.7653
## URBAN                   -1.766685  0.170899  0.745551 -2.370   0.0178 *
## HTN_REP                 -0.043962  0.956990  0.091853 -0.479   0.6322
## RBC                      0.009946  1.009996  0.061852  0.161   0.8722
## DRMI                    -0.090652  0.913335  0.171547 -0.528   0.5972
## DIAB                    -0.112334  0.893746  0.185810 -0.605   0.5455
## BOOZE_q(0,0.5]:SEXMale -0.493872  0.610259  0.307189 -1.608   0.1079
## BOOZE_q(0.5,2]:SEXMale -0.555087  0.574023  0.252243 -2.201   0.0278 *
## BOOZE_q(2,77]:SEXMale  -0.405682  0.666522  0.218054 -1.860   0.0628 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                         exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]             1.3534     0.7389   0.90088    2.0332
## BOOZE_q(0.5,2]             1.3313     0.7512   0.91724    1.9322
## BOOZE_q(2,77]              1.5890     0.6293   1.10814    2.2786
## AGEYRS                     1.0812     0.9249   1.07090    1.0916
## SEXMale                    2.1900     0.4566   1.66634    2.8782
## as.factor(RACE)2           0.9355     1.0689   0.69064    1.2673
```

```
## as.factor(RACE)3          0.6267     1.5956   0.27690    1.4186
## GRADES                    0.9853     1.0149   0.96139    1.0098
## as.factor(MARRY)3         1.0310     0.9699   0.80035    1.3282
## as.factor(MARRY)4         1.2526     0.7984   0.87476    1.7936
## as.factor(MARRY)5         0.8738     1.1445   0.45851    1.6651
## as.factor(MARRY)6         0.8551     1.1694   0.55948    1.3071
## as.factor(MARRY)8         1.0236     0.9769   0.25292    4.1426
## as.factor(SIZE)2          1.1824     0.8458   0.85052    1.6437
## as.factor(SIZE)3          1.4358     0.6965   1.03522    1.9913
## as.factor(SIZE)4          1.0151     0.9852   0.69396    1.4847
## as.factor(SIZE)5          0.9000     1.1111   0.47785    1.6951
## as.factor(SIZE)6          1.0742     0.9310   0.58848    1.9607
## as.factor(SIZE)7          0.9446     1.0586   0.56640    1.5754
## as.factor(SIZE)8          5.7893     0.1727   1.36032   24.6381
## AVGSMK                    1.0254     0.9753   1.01983    1.0309
## as.factor(SMSA)2          0.8202     1.2192   0.63051    1.0670
## as.factor(SMSA)4          0.9348     1.0698   0.60034    1.4555
## URBAN                     0.1709     5.8514   0.03964    0.7368
## HTN_REP                   0.9570     1.0449   0.79932    1.1458
## RBC                       1.0100     0.9901   0.89469    1.1402
## DRMI                      0.9133     1.0949   0.65254    1.2784
## DIAB                      0.8937     1.1189   0.62094    1.2864
## BOOZE_q(0,0.5]:SEXMale    0.6103     1.6386   0.33422    1.1143
## BOOZE_q(0.5,2]:SEXMale    0.5740     1.7421   0.35012    0.9411
## BOOZE_q(2,77]:SEXMale     0.6665     1.5003   0.43472    1.0219
##
## Concordance= 0.754  (se = 0.01 )
## Likelihood ratio test= 478.9  on 31 df,   p=<2e-16
## Wald test            = 408.7  on 31 df,   p=<2e-16
## Score (logrank) test = 437.8  on 31 df,   p=<2e-16
```

```
coef <- coef(fit4_q)[c(3, 31)] %>% as.matrix(ncol = 1)
vcov <- vcov(fit4_q)[c(3, 31), c(3, 31)]
var <- t(c(1, 1)) %*% vcov %*% c(1, 1)
sd <- sqrt(var)
mean_male <- coef[2,] + coef[1,]
exp(mean_male)
```

```
## BOOZE_q(2,77]:SEXMale
##              1.059128
```

```
c(mean_male - 1.96*sd, mean_male+ 1.96*sd) %>% exp()
```

```
## [1] 0.8187146 1.3701391
```

```
# Continuous version
fit4_con <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + BOOZE:SEX  + AGEYRS +
                      SEX + as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) + AVGSMK + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
fit4_con %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + BOOZE:SEX +
##     AGEYRS + SEX + as.factor(RACE) + GRADES + as.factor(MARRY) +
##     as.factor(SIZE) + AVGSMK + as.factor(SMSA) + URBAN + HTN_REP +
##     RBC + DRMI + DIAB, data = dta)
```

```
##
##   n= 9190, number of events= 552
##
##                        coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE              0.030800  1.031279  0.019053  1.617   0.1060
## AGEYRS             0.078126  1.081259  0.004866 16.054  < 2e-16 ***
## SEXMale            0.522100  1.685564  0.108219  4.825  1.4e-06 ***
## as.factor(RACE)2  -0.058404  0.943269  0.154645 -0.378   0.7057
## as.factor(RACE)3  -0.449974  0.637645  0.416591 -1.080   0.2801
## GRADES            -0.014027  0.986071  0.012392 -1.132   0.2577
## as.factor(MARRY)3  0.018414  1.018584  0.129168  0.143   0.8866
## as.factor(MARRY)4  0.234040  1.263695  0.182684  1.281   0.2002
## as.factor(MARRY)5 -0.127087  0.880657  0.328696 -0.387   0.6990
## as.factor(MARRY)6 -0.159358  0.852691  0.216306 -0.737   0.4613
## as.factor(MARRY)8  0.094250  1.098834  0.712587  0.132   0.8948
## as.factor(SIZE)2   0.150162  1.162023  0.168022  0.894   0.3715
## as.factor(SIZE)3   0.345427  1.412592  0.166742  2.072   0.0383 *
## as.factor(SIZE)4   0.007147  1.007173  0.193984  0.037   0.9706
## as.factor(SIZE)5  -0.113895  0.892352  0.322744 -0.353   0.7242
## as.factor(SIZE)6   0.041269  1.042133  0.306998  0.134   0.8931
## as.factor(SIZE)7  -0.069331  0.933017  0.260922 -0.266   0.7905
## as.factor(SIZE)8   1.825465  6.205679  0.737802  2.474   0.0134 *
## AVGSMK             0.024850  1.025161  0.002744  9.057  < 2e-16 ***
## as.factor(SMSA)2  -0.196562  0.821551  0.134291 -1.464   0.1433
## as.factor(SMSA)4  -0.056827  0.944757  0.225800 -0.252   0.8013
## URBAN             -1.846863  0.157731  0.744872 -2.479   0.0132 *
## HTN_REP           -0.056597  0.944975  0.091777 -0.617   0.5374
## RBC                0.009765  1.009813  0.061460  0.159   0.8738
## DRMI              -0.086738  0.916918  0.171529 -0.506   0.6131
## DIAB              -0.110376  0.895497  0.185070 -0.596   0.5509
## BOOZE:SEXMale     -0.006266  0.993753  0.020302 -0.309   0.7576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE                1.0313     0.9697   0.99348    1.0705
## AGEYRS               1.0813     0.9248   1.07099    1.0916
## SEXMale              1.6856     0.5933   1.36342    2.0838
## as.factor(RACE)2     0.9433     1.0601   0.69663    1.2772
## as.factor(RACE)3     0.6376     1.5683   0.28182    1.4427
## GRADES               0.9861     1.0141   0.96241    1.0103
## as.factor(MARRY)3    1.0186     0.9818   0.79077    1.3120
## as.factor(MARRY)4    1.2637     0.7913   0.88337    1.8078
## as.factor(MARRY)5    0.8807     1.1355   0.46240    1.6772
## as.factor(MARRY)6    0.8527     1.1728   0.55805    1.3029
## as.factor(MARRY)8    1.0988     0.9101   0.27188    4.4411
## as.factor(SIZE)2     1.1620     0.8606   0.83598    1.6152
## as.factor(SIZE)3     1.4126     0.7079   1.01879    1.9586
## as.factor(SIZE)4     1.0072     0.9929   0.68863    1.4731
## as.factor(SIZE)5     0.8924     1.1206   0.47404    1.6798
## as.factor(SIZE)6     1.0421     0.9596   0.57096    1.9021
## as.factor(SIZE)7     0.9330     1.0718   0.55949    1.5559
## as.factor(SIZE)8     6.2057     0.1611   1.46141   26.3516
## AVGSMK               1.0252     0.9755   1.01966    1.0307
```

```
## as.factor(SMSA)2     0.8216     1.2172   0.63143    1.0689
## as.factor(SMSA)4     0.9448     1.0585   0.60690    1.4707
## URBAN                0.1577     6.3399   0.03663    0.6791
## HTN_REP              0.9450     1.0582   0.78940    1.1312
## RBC                  1.0098     0.9903   0.89521    1.1391
## DRMI                 0.9169     1.0906   0.65513    1.2833
## DIAB                 0.8955     1.1167   0.62306    1.2871
## BOOZE:SEXMale        0.9938     1.0063   0.95499    1.0341
##
## Concordance= 0.755  (se = 0.01 )
## Likelihood ratio test= 477.8  on 27 df,   p=<2e-16
## Wald test            = 411  on 27 df,   p=<2e-16
## Score (logrank) test = 438.4  on 27 df,   p=<2e-16
```

```
coef <- coef(fit4_con)[c(1, 27)] %>% as.matrix(ncol = 1)
vcov <- vcov(fit4_con)[c(1, 27), c(1, 27)]
var <- t(c(1, 1)) %*% vcov %*% c(1, 1)
sd <- sqrt(var)
mean_male <- coef[2,] + coef[1,]
exp(mean_male)
```

```
## BOOZE:SEXMale
##      1.024837
```

```
c(mean_male - 1.96*sd, mean_male+ 1.96*sd) %>% exp()
```

```
## [1] 1.009782 1.040117
```

Stratified Cox model by sex, and adding an interaction term by SEX and alcohol consumption

```
# Categorical version
fit5_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + strata(SEX) + BOOZE_q:strata(SEX) + AGEYRS  + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
fit5_q %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + strata(SEX) +
##     BOOZE_q:strata(SEX) + AGEYRS + as.factor(RACE) + GRADES +
##     as.factor(MARRY) + as.factor(SIZE) + AVGSMK + as.factor(SMSA) +
##     URBAN + HTN_REP + RBC + DRMI + DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                                    coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5]                 0.300374  1.350364  0.207677  1.446   0.1481
## BOOZE_q(0.5,2]                 0.280642  1.323980  0.190103  1.476   0.1399
## BOOZE_q(2,77]                  0.460000  1.584074  0.183936  2.501   0.0124
## AGEYRS                         0.077855  1.080966  0.004891 15.918   <2e-16
## as.factor(RACE)2              -0.066293  0.935857  0.154866 -0.428   0.6686
## as.factor(RACE)3              -0.466262  0.627343  0.416794 -1.119   0.2633
## GRADES                        -0.014681  0.985426  0.012538 -1.171   0.2416
## as.factor(MARRY)3              0.032210  1.032734  0.129181  0.249   0.8031
## as.factor(MARRY)4              0.223209  1.250082  0.183193  1.218   0.2231
## as.factor(MARRY)5             -0.133737  0.874820  0.328993 -0.407   0.6844
```

```
## as.factor(MARRY)6              -0.158490  0.853432  0.216451 -0.732   0.4640
## as.factor(MARRY)8               0.028605  1.029018  0.713344  0.040   0.9680
## as.factor(SIZE)2                0.166094  1.180684  0.168084  0.988   0.3231
## as.factor(SIZE)3                0.362502  1.436921  0.166874  2.172   0.0298
## as.factor(SIZE)4                0.016400  1.016535  0.194023  0.085   0.9326
## as.factor(SIZE)5               -0.104615  0.900671  0.323020 -0.324   0.7460
## as.factor(SIZE)6                0.072781  1.075495  0.307076  0.237   0.8126
## as.factor(SIZE)7               -0.056924  0.944666  0.260976 -0.218   0.8273
## as.factor(SIZE)8                1.724272  5.608434  0.739043  2.333   0.0196
## AVGSMK                          0.024968  1.025282  0.002762  9.039   <2e-16
## as.factor(SMSA)2               -0.196934  0.821245  0.134203 -1.467   0.1423
## as.factor(SMSA)4               -0.069065  0.933266  0.225912 -0.306   0.7598
## URBAN                          -1.733064  0.176742  0.745653 -2.324   0.0201
## HTN_REP                        -0.044148  0.956812  0.091822 -0.481   0.6307
## RBC                             0.012230  1.012305  0.061772  0.198   0.8431
## DRMI                           -0.094820  0.909537  0.171556 -0.553   0.5805
## DIAB                           -0.113823  0.892416  0.185809 -0.613   0.5402
## BOOZE_q(0,0.5]:strata(SEX)Male -0.491957  0.611428  0.307200 -1.601   0.1093
## BOOZE_q(0.5,2]:strata(SEX)Male -0.546088  0.579211  0.252285 -2.165   0.0304
## BOOZE_q(2,77]:strata(SEX)Male  -0.398870  0.671078  0.218088 -1.829   0.0674
##
## BOOZE_q(0,0.5]
## BOOZE_q(0.5,2]
## BOOZE_q(2,77]                  *
## AGEYRS                         ***
## as.factor(RACE)2
## as.factor(RACE)3
## GRADES
## as.factor(MARRY)3
## as.factor(MARRY)4
## as.factor(MARRY)5
## as.factor(MARRY)6
## as.factor(MARRY)8
## as.factor(SIZE)2
## as.factor(SIZE)3               *
## as.factor(SIZE)4
## as.factor(SIZE)5
## as.factor(SIZE)6
## as.factor(SIZE)7
## as.factor(SIZE)8               *
## AVGSMK                         ***
## as.factor(SMSA)2
## as.factor(SMSA)4
## URBAN                          *
## HTN_REP
## RBC
## DRMI
## DIAB
## BOOZE_q(0,0.5]:strata(SEX)Male
## BOOZE_q(0.5,2]:strata(SEX)Male *
## BOOZE_q(2,77]:strata(SEX)Male  .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
##                                 exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]                     1.3504     0.7405   0.89883    2.0287
## BOOZE_q(0.5,2]                     1.3240     0.7553   0.91215    1.9218
## BOOZE_q(2,77]                      1.5841     0.6313   1.10461    2.2717
## AGEYRS                             1.0810     0.9251   1.07065    1.0914
## as.factor(RACE)2                   0.9359     1.0685   0.69086    1.2677
## as.factor(RACE)3                   0.6273     1.5940   0.27716    1.4200
## GRADES                             0.9854     1.0148   0.96151    1.0099
## as.factor(MARRY)3                  1.0327     0.9683   0.80173    1.3303
## as.factor(MARRY)4                  1.2501     0.7999   0.87298    1.7901
## as.factor(MARRY)5                  0.8748     1.1431   0.45907    1.6671
## as.factor(MARRY)6                  0.8534     1.1717   0.55837    1.3044
## as.factor(MARRY)8                  1.0290     0.9718   0.25423    4.1651
## as.factor(SIZE)2                   1.1807     0.8470   0.84930    1.6414
## as.factor(SIZE)3                   1.4369     0.6959   1.03607    1.9929
## as.factor(SIZE)4                   1.0165     0.9837   0.69498    1.4869
## as.factor(SIZE)5                   0.9007     1.1103   0.47820    1.6964
## as.factor(SIZE)6                   1.0755     0.9298   0.58915    1.9633
## as.factor(SIZE)7                   0.9447     1.0586   0.56642    1.5755
## as.factor(SIZE)8                   5.6084     0.1783   1.31755   23.8735
## AVGSMK                             1.0253     0.9753   1.01975    1.0308
## as.factor(SMSA)2                   0.8212     1.2177   0.63130    1.0683
## as.factor(SMSA)4                   0.9333     1.0715   0.59939    1.4531
## URBAN                              0.1767     5.6580   0.04099    0.7622
## HTN_REP                            0.9568     1.0451   0.79922    1.1455
## RBC                                1.0123     0.9878   0.89687    1.1426
## DRMI                               0.9095     1.0995   0.64982    1.2731
## DIAB                               0.8924     1.1206   0.62002    1.2845
## BOOZE_q(0,0.5]:strata(SEX)Male     0.6114     1.6355   0.33485    1.1164
## BOOZE_q(0.5,2]:strata(SEX)Male     0.5792     1.7265   0.35326    0.9497
## BOOZE_q(2,77]:strata(SEX)Male      0.6711     1.4901   0.43766    1.0290
##
## Concordance= 0.738  (se = 0.01 )
## Likelihood ratio test= 429.9  on 30 df,   p=<2e-16
## Wald test            = 353.9  on 30 df,   p=<2e-16
## Score (logrank) test = 383.4  on 30 df,   p=<2e-16
```

```
coef <- coef(fit5_q)[c(3, 30)] %>% as.matrix(ncol = 1)
vcov <- vcov(fit5_q)[c(3, 30), c(3, 30)]
var <- t(c(1, 1)) %*% vcov %*% c(1, 1)
sd <- sqrt(var)
mean_male <- coef[2,] + coef[1,]
exp(mean_male)
```

```
## BOOZE_q(2,77]:strata(SEX)Male
##                      1.063037
```

```
c(mean_male - 1.96*sd, mean_male+ 1.96*sd) %>% exp()
```

```
## [1] 0.8217616 1.3751522
```

```
# Continuous version
fit5_cont <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + strata(SEX) + BOOZE:strata(SEX) + AGEYRS  + as
                     HTN_REP + RBC + DRMI + DIAB, data = dta)
fit5_cont %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + strata(SEX) +
##     BOOZE:strata(SEX) + AGEYRS + as.factor(RACE) + GRADES + as.factor(MARRY) +
##     as.factor(SIZE) + AVGSMK + as.factor(SMSA) + URBAN + HTN_REP +
##     RBC + DRMI + DIAB, data = dta)
##
##   n= 9190, number of events= 552
##
##                             coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE                   0.030383  1.030849  0.019053  1.595   0.1108
## AGEYRS                  0.077876  1.080989  0.004865 16.009   <2e-16 ***
## as.factor(RACE)2       -0.058446  0.943229  0.154663 -0.378   0.7055
## as.factor(RACE)3       -0.449498  0.637948  0.416599 -1.079   0.2806
## GRADES                 -0.013846  0.986249  0.012398 -1.117   0.2641
## as.factor(MARRY)3       0.020074  1.020277  0.129126  0.155   0.8765
## as.factor(MARRY)4       0.231805  1.260874  0.182702  1.269   0.2045
## as.factor(MARRY)5      -0.125861  0.881738  0.328694 -0.383   0.7018
## as.factor(MARRY)6      -0.161532  0.850839  0.216294 -0.747   0.4552
## as.factor(MARRY)8       0.099779  1.104927  0.712624  0.140   0.8886
## as.factor(SIZE)2        0.148899  1.160556  0.168026  0.886   0.3755
## as.factor(SIZE)3        0.345704  1.412985  0.166733  2.073   0.0381 *
## as.factor(SIZE)4        0.008757  1.008795  0.193967  0.045   0.9640
## as.factor(SIZE)5       -0.113205  0.892968  0.322759 -0.351   0.7258
## as.factor(SIZE)6        0.042694  1.043619  0.307055  0.139   0.8894
## as.factor(SIZE)7       -0.069086  0.933246  0.260948 -0.265   0.7912
## as.factor(SIZE)8        1.794065  6.013851  0.737925  2.431   0.0150 *
## AVGSMK                  0.024763  1.025072  0.002745  9.020   <2e-16 ***
## as.factor(SMSA)2       -0.195066  0.822781  0.134280 -1.453   0.1463
## as.factor(SMSA)4       -0.058536  0.943144  0.225792 -0.259   0.7954
## URBAN                  -1.813968  0.163006  0.744990 -2.435   0.0149 *
## HTN_REP                -0.056634  0.944940  0.091743 -0.617   0.5370
## RBC                     0.012133  1.012207  0.061377  0.198   0.8433
## DRMI                   -0.090597  0.913386  0.171533 -0.528   0.5974
## DIAB                   -0.111247  0.894718  0.185062 -0.601   0.5478
## BOOZE:strata(SEX)Male -0.005884  0.994133  0.020297 -0.290   0.7719
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                       exp(coef) exp(-coef) lower .95 upper .95
## BOOZE                    1.0308     0.9701   0.99306     1.070
## AGEYRS                   1.0810     0.9251   1.07073     1.091
## as.factor(RACE)2         0.9432     1.0602   0.69658     1.277
## as.factor(RACE)3         0.6379     1.5675   0.28195     1.443
## GRADES                   0.9862     1.0139   0.96257     1.011
## as.factor(MARRY)3        1.0203     0.9801   0.79215     1.314
## as.factor(MARRY)4        1.2609     0.7931   0.88136     1.804
## as.factor(MARRY)5        0.8817     1.1341   0.46297     1.679
## as.factor(MARRY)6        0.8508     1.1753   0.55685     1.300
## as.factor(MARRY)8        1.1049     0.9050   0.27337     4.466
## as.factor(SIZE)2         1.1606     0.8617   0.83491     1.613
## as.factor(SIZE)3         1.4130     0.7077   1.01909     1.959
## as.factor(SIZE)4         1.0088     0.9913   0.68976     1.475
## as.factor(SIZE)5         0.8930     1.1199   0.47435     1.681
## as.factor(SIZE)6         1.0436     0.9582   0.57171     1.905
```

```
## as.factor(SIZE)7         0.9332     1.0715   0.55960     1.556
## as.factor(SIZE)8         6.0139     0.1663   1.41589    25.543
## AVGSMK                   1.0251     0.9755   1.01957     1.031
## as.factor(SMSA)2         0.8228     1.2154   0.63239     1.070
## as.factor(SMSA)4         0.9431     1.0603   0.60588     1.468
## URBAN                    0.1630     6.1347   0.03785     0.702
## HTN_REP                  0.9449     1.0583   0.78943     1.131
## RBC                      1.0122     0.9879   0.89748     1.142
## DRMI                     0.9134     1.0948   0.65260     1.278
## DIAB                     0.8947     1.1177   0.62253     1.286
## BOOZE:strata(SEX)Male    0.9941     1.0059   0.95536     1.034
##
## Concordance= 0.739  (se = 0.01 )
## Likelihood ratio test= 428.9  on 26 df,   p=<2e-16
## Wald test            = 355.5  on 26 df,   p=<2e-16
## Score (logrank) test = 384.2  on 26 df,   p=<2e-16
```

```
coef <- coef(fit5_cont)[c(1, 26)] %>% as.matrix(ncol = 1)
vcov <- vcov(fit5_cont)[c(1, 26), c(1, 26)]
var <- t(c(1, 1)) %*% vcov %*% c(1, 1)
sd <- sqrt(var)
mean_male <- coef[2,] + coef[1,]
exp(mean_male)
```

```
## BOOZE:strata(SEX)Male
##              1.024801
```

```
c(mean_male - 1.96*sd, mean_male+ 1.96*sd) %>% exp()
```

```
## [1] 1.009775 1.040050
```

Subgroup analysis

```
# categorical version
fit6_q_m <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS  + as.factor(RACE) + GRADES + as.factor
                    HTN_REP + RBC + DRMI + DIAB, data = dta[dta$SEX == "Male",])
fit6_q_m %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS +
##     as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta[dta$SEX == "Male", ])
##
##   n= 4314, number of events= 329
##
##                          coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5]    -0.119329  0.887516  0.229390 -0.520   0.6029
## BOOZE_q(0.5,2]    -0.219839  0.802648  0.174367 -1.261   0.2074
## BOOZE_q(2,77]      0.136033  1.145720  0.135379  1.005   0.3150
## AGEYRS             0.088349  1.092370  0.006724 13.140  < 2e-16 ***
## as.factor(RACE)2   0.014507  1.014613  0.190425  0.076   0.9393
## as.factor(RACE)3  -0.752925  0.470987  0.589271 -1.278   0.2013
## GRADES            -0.033238  0.967308  0.015547 -2.138   0.0325 *
## as.factor(MARRY)3 -0.040408  0.960397  0.227778 -0.177   0.8592
```

```
## as.factor(MARRY)4  0.114662  1.121494  0.271474  0.422   0.6728
## as.factor(MARRY)5 -0.269597  0.763687  0.424699 -0.635   0.5256
## as.factor(MARRY)6  0.083859  1.087475  0.264647  0.317   0.7513
## as.factor(MARRY)8 -0.065804  0.936315  1.015043 -0.065   0.9483
## as.factor(SIZE)2   0.219714  1.245721  0.223155  0.985   0.3248
## as.factor(SIZE)3   0.471345  1.602148  0.214932  2.193   0.0283 *
## as.factor(SIZE)4  -0.008270  0.991764  0.256648 -0.032   0.9743
## as.factor(SIZE)5  -0.032918  0.967618  0.417673 -0.079   0.9372
## as.factor(SIZE)6  -0.582896  0.558279  0.493430 -1.181   0.2375
## as.factor(SIZE)7  -0.066800  0.935382  0.356182 -0.188   0.8512
## as.factor(SIZE)8   1.639899  5.154651  0.768539  2.134   0.0329 *
## AVGSMK             0.024089  1.024381  0.003367  7.155 8.35e-13 ***
## as.factor(SMSA)2  -0.173825  0.840444  0.176303 -0.986   0.3242
## as.factor(SMSA)4   0.213062  1.237462  0.318649  0.669   0.5037
## URBAN             -1.956302  0.141380  0.789058 -2.479   0.0132 *
## HTN_REP           -0.110896  0.895032  0.122633 -0.904   0.3658
## RBC               -0.118771  0.888011  0.094023 -1.263   0.2065
## DRMI              -0.049896  0.951328  0.193459 -0.258   0.7965
## DIAB              -0.044193  0.956769  0.238210 -0.186   0.8528
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]       0.8875     1.1267   0.56613    1.3913
## BOOZE_q(0.5,2]       0.8026     1.2459   0.57030    1.1297
## BOOZE_q(2,77]        1.1457     0.8728   0.87871    1.4939
## AGEYRS               1.0924     0.9154   1.07807    1.1069
## as.factor(RACE)2     1.0146     0.9856   0.69857    1.4736
## as.factor(RACE)3     0.4710     2.1232   0.14840    1.4949
## GRADES               0.9673     1.0338   0.93828    0.9972
## as.factor(MARRY)3    0.9604     1.0412   0.61456    1.5008
## as.factor(MARRY)4    1.1215     0.8917   0.65875    1.9093
## as.factor(MARRY)5    0.7637     1.3094   0.33221    1.7556
## as.factor(MARRY)6    1.0875     0.9196   0.64737    1.8268
## as.factor(MARRY)8    0.9363     1.0680   0.12806    6.8459
## as.factor(SIZE)2     1.2457     0.8027   0.80440    1.9292
## as.factor(SIZE)3     1.6021     0.6242   1.05136    2.4415
## as.factor(SIZE)4     0.9918     1.0083   0.59972    1.6401
## as.factor(SIZE)5     0.9676     1.0335   0.42676    2.1940
## as.factor(SIZE)6     0.5583     1.7912   0.21225    1.4685
## as.factor(SIZE)7     0.9354     1.0691   0.46538    1.8801
## as.factor(SIZE)8     5.1547     0.1940   1.14292   23.2478
## AVGSMK               1.0244     0.9762   1.01764    1.0312
## as.factor(SMSA)2     0.8404     1.1898   0.59489    1.1873
## as.factor(SMSA)4     1.2375     0.8081   0.66267    2.3108
## URBAN                0.1414     7.0731   0.03011    0.6638
## HTN_REP              0.8950     1.1173   0.70381    1.1382
## RBC                  0.8880     1.1261   0.73856    1.0677
## DRMI                 0.9513     1.0512   0.65112    1.3900
## DIAB                 0.9568     1.0452   0.59985    1.5261
##
## Concordance= 0.772  (se = 0.012 )
## Likelihood ratio test= 333  on 27 df,   p=<2e-16
## Wald test            = 262.3  on 27 df,   p=<2e-16
```

```
## Score (logrank) test = 294.4  on 27 df,   p=<2e-16
```

```
fit6_q_f <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS  + as.factor(RACE) + GRADES + as.factor
                      HTN_REP + RBC + DRMI + DIAB, data = dta[dta$SEX == "Female",])
```

```
## Warning in coxph.fit(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 19,23 ; coefficient may be infinite.
```

```
fit6_q_f %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS +
##     as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta[dta$SEX == "Female", ])
##
##   n= 4876, number of events= 223
##
##                          coef  exp(coef)   se(coef)      z Pr(>|z|)
## BOOZE_q(0,0.5]     2.346e-01  1.264e+00  2.101e-01  1.117   0.2641
## BOOZE_q(0.5,2]     1.785e-01  1.195e+00  1.959e-01  0.911   0.3621
## BOOZE_q(2,77]      3.028e-01  1.354e+00  1.940e-01  1.561   0.1186
## AGEYRS             6.231e-02  1.064e+00  7.383e-03  8.440  < 2e-16 ***
## as.factor(RACE)2  -3.387e-01  7.127e-01  2.776e-01 -1.220   0.2224
## as.factor(RACE)3  -1.188e-02  9.882e-01  5.903e-01 -0.020   0.9839
## GRADES             3.107e-02  1.032e+00  2.223e-02  1.398   0.1622
## as.factor(MARRY)3  1.736e-01  1.190e+00  1.624e-01  1.068   0.2854
## as.factor(MARRY)4  2.914e-01  1.338e+00  2.519e-01  1.157   0.2473
## as.factor(MARRY)5  1.062e-01  1.112e+00  5.257e-01  0.202   0.8400
## as.factor(MARRY)6 -5.246e-01  5.918e-01  3.937e-01 -1.332   0.1828
## as.factor(MARRY)8  5.692e-02  1.059e+00  1.010e+00  0.056   0.9551
## as.factor(SIZE)2   7.061e-02  1.073e+00  2.568e-01  0.275   0.7834
## as.factor(SIZE)3   1.723e-01  1.188e+00  2.672e-01  0.645   0.5190
## as.factor(SIZE)4   1.997e-02  1.020e+00  2.987e-01  0.067   0.9467
## as.factor(SIZE)5  -4.074e-01  6.654e-01  5.554e-01 -0.734   0.4632
## as.factor(SIZE)6   6.107e-01  1.842e+00  4.059e-01  1.505   0.1324
## as.factor(SIZE)7  -1.391e-01  8.701e-01  3.996e-01 -0.348   0.7277
## as.factor(SIZE)8  -1.101e+01  1.656e-05  1.123e+03 -0.010   0.9922
## AVGSMK             2.856e-02  1.029e+00  5.019e-03  5.691 1.27e-08 ***
## as.factor(SMSA)2  -2.101e-01  8.105e-01  2.083e-01 -1.009   0.3131
## as.factor(SMSA)4  -4.197e-01  6.572e-01  3.284e-01 -1.278   0.2012
## URBAN              1.138e+01  8.721e+04  1.123e+03  0.010   0.9919
## HTN_REP            1.116e-01  1.118e+00  1.408e-01  0.793   0.4281
## RBC                1.427e-01  1.153e+00  7.628e-02  1.871   0.0613 .
## DRMI              -3.667e-01  6.930e-01  3.878e-01 -0.946   0.3443
## DIAB              -2.153e-01  8.063e-01  3.017e-01 -0.713   0.4756
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE_q(0,0.5]    1.264e+00  7.909e-01    0.8377     1.909
## BOOZE_q(0.5,2]    1.195e+00  8.365e-01    0.8143     1.755
## BOOZE_q(2,77]     1.354e+00  7.388e-01    0.9255     1.980
## AGEYRS            1.064e+00  9.396e-01    1.0490     1.080
## as.factor(RACE)2  7.127e-01  1.403e+00    0.4137     1.228
```

```
## as.factor(RACE)3  9.882e-01  1.012e+00    0.3107     3.143
## GRADES            1.032e+00  9.694e-01    0.9876     1.077
## as.factor(MARRY)3 1.190e+00  8.407e-01    0.8652     1.636
## as.factor(MARRY)4 1.338e+00  7.472e-01    0.8169     2.193
## as.factor(MARRY)5 1.112e+00  8.993e-01    0.3969     3.116
## as.factor(MARRY)6 5.918e-01  1.690e+00    0.2735     1.280
## as.factor(MARRY)8 1.059e+00  9.447e-01    0.1461     7.668
## as.factor(SIZE)2  1.073e+00  9.318e-01    0.6487     1.775
## as.factor(SIZE)3  1.188e+00  8.417e-01    0.7037     2.006
## as.factor(SIZE)4  1.020e+00  9.802e-01    0.5681     1.832
## as.factor(SIZE)5  6.654e-01  1.503e+00    0.2241     1.976
## as.factor(SIZE)6  1.842e+00  5.430e-01    0.8313     4.080
## as.factor(SIZE)7  8.701e-01  1.149e+00    0.3976     1.904
## as.factor(SIZE)8  1.656e-05  6.040e+04    0.0000       Inf
## AVGSMK            1.029e+00  9.718e-01    1.0189     1.039
## as.factor(SMSA)2  8.105e-01  1.234e+00    0.5388     1.219
## as.factor(SMSA)4  6.572e-01  1.522e+00    0.3453     1.251
## URBAN             8.721e+04  1.147e-05    0.0000       Inf
## HTN_REP           1.118e+00  8.944e-01    0.8484     1.473
## RBC               1.153e+00  8.670e-01    0.9932     1.339
## DRMI              6.930e-01  1.443e+00    0.3241     1.482
## DIAB              8.063e-01  1.240e+00    0.4463     1.457
##
## Concordance= 0.717  (se = 0.016 )
## Likelihood ratio test= 141.4  on 27 df,   p=<2e-16
## Wald test            = 124.5  on 27 df,   p=2e-14
## Score (logrank) test = 132.7  on 27 df,   p=6e-16
```

```
z <- (coef(fit6_q_f)[3] - coef(fit6_q_m)[3])/sqrt(vcov(fit6_q_f)[3,3] + vcov(fit6_q_m)[3,3])
1 - pnorm(abs(z))
```

```
## BOOZE_q(2,77]
##     0.2404372
```

```
# continuous version
fit6_cont_m <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS  + as.factor(RACE) + GRADES + as.facto
                     HTN_REP + RBC + DRMI + DIAB, data = dta[dta$SEX == "Male",])
fit6_cont_m %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS +
##     as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta[dta$SEX == "Male", ])
##
##   n= 4314, number of events= 329
##
##                        coef exp(coef)  se(coef)      z Pr(>|z|)
## BOOZE              0.027028  1.027396  0.007505  3.601 0.000316 ***
## AGEYRS             0.089025  1.093108  0.006717 13.254  < 2e-16 ***
## as.factor(RACE)2   0.032191  1.032715  0.190032  0.169 0.865485
## as.factor(RACE)3  -0.711462  0.490926  0.589152 -1.208 0.227200
## GRADES            -0.034877  0.965724  0.015428 -2.261 0.023780 *
## as.factor(MARRY)3 -0.071713  0.930798  0.228175 -0.314 0.753303
## as.factor(MARRY)4  0.124238  1.132285  0.270507  0.459 0.646034
```

```
## as.factor(MARRY)5 -0.253290  0.776243  0.423828 -0.598 0.550091
## as.factor(MARRY)6  0.083721  1.087326  0.264406  0.317 0.751517
## as.factor(MARRY)8  0.021485  1.021718  1.014861  0.021 0.983110
## as.factor(SIZE)2   0.200740  1.222307  0.223233  0.899 0.368525
## as.factor(SIZE)3   0.467983  1.596770  0.215059  2.176 0.029550 *
## as.factor(SIZE)4   0.002177  1.002180  0.256571  0.008 0.993229
## as.factor(SIZE)5  -0.068238  0.934038  0.417371 -0.163 0.870129
## as.factor(SIZE)6  -0.612824  0.541819  0.493718 -1.241 0.214516
## as.factor(SIZE)7  -0.081216  0.921994  0.355378 -0.229 0.819230
## as.factor(SIZE)8   1.760029  5.812603  0.767788  2.292 0.021886 *
## AVGSMK             0.023583  1.023863  0.003348  7.044 1.87e-12 ***
## as.factor(SMSA)2  -0.164745  0.848110  0.176423 -0.934 0.350402
## as.factor(SMSA)4   0.265983  1.304713  0.318330  0.836 0.403404
## URBAN             -2.085212  0.124281  0.788630 -2.644 0.008191 **
## HTN_REP           -0.113310  0.892873  0.122585 -0.924 0.355308
## RBC               -0.105150  0.900190  0.093337 -1.127 0.259927
## DRMI              -0.050182  0.951056  0.193419 -0.259 0.795288
## DIAB              -0.014320  0.985783  0.237405 -0.060 0.951903
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## BOOZE                1.0274     0.9733   1.01239    1.0426
## AGEYRS               1.0931     0.9148   1.07881    1.1076
## as.factor(RACE)2     1.0327     0.9683   0.71158    1.4988
## as.factor(RACE)3     0.4909     2.0370   0.15471    1.5578
## GRADES               0.9657     1.0355   0.93696    0.9954
## as.factor(MARRY)3    0.9308     1.0743   0.59516    1.4557
## as.factor(MARRY)4    1.1323     0.8832   0.66635    1.9240
## as.factor(MARRY)5    0.7762     1.2883   0.33825    1.7814
## as.factor(MARRY)6    1.0873     0.9197   0.64759    1.8257
## as.factor(MARRY)8    1.0217     0.9787   0.13979    7.4676
## as.factor(SIZE)2     1.2223     0.8181   0.78916    1.8932
## as.factor(SIZE)3     1.5968     0.6263   1.04757    2.4339
## as.factor(SIZE)4     1.0022     0.9978   0.60611    1.6571
## as.factor(SIZE)5     0.9340     1.0706   0.41219    2.1166
## as.factor(SIZE)6     0.5418     1.8456   0.20587    1.4260
## as.factor(SIZE)7     0.9220     1.0846   0.45944    1.8502
## as.factor(SIZE)8     5.8126     0.1720   1.29071   26.1766
## AVGSMK               1.0239     0.9767   1.01717    1.0306
## as.factor(SMSA)2     0.8481     1.1791   0.60018    1.1985
## as.factor(SMSA)4     1.3047     0.7665   0.69912    2.4349
## URBAN                0.1243     8.0463   0.02649    0.5830
## HTN_REP              0.8929     1.1200   0.70218    1.1354
## RBC                  0.9002     1.1109   0.74970    1.0809
## DRMI                 0.9511     1.0515   0.65098    1.3895
## DIAB                 0.9858     1.0144   0.61902    1.5699
##
## Concordance= 0.775  (se = 0.011 )
## Likelihood ratio test= 338.7  on 25 df,   p=<2e-16
## Wald test            = 269.3  on 25 df,   p=<2e-16
## Score (logrank) test = 301  on 25 df,   p=<2e-16
```

```
fit6_cont_f <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS  + as.factor(RACE) + GRADES + as.facto
                     HTN_REP + RBC + DRMI + DIAB, data = dta[dta$SEX == "Female",])
```

```
## Warning in coxph.fit(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 17,21 ; coefficient may be infinite.
```

```
fit6_cont_f %>% summary()
```

```
## Call:
## coxph(formula = Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS +
##     as.factor(RACE) + GRADES + as.factor(MARRY) + as.factor(SIZE) +
##     AVGSMK + as.factor(SMSA) + URBAN + HTN_REP + RBC + DRMI +
##     DIAB, data = dta[dta$SEX == "Female", ])
##
##   n= 4876, number of events= 223
##
##                          coef  exp(coef)   se(coef)      z Pr(>|z|)
## BOOZE                1.422e-02  1.014e+00  2.099e-02  0.677   0.4981
## AGEYRS               6.111e-02  1.063e+00  7.301e-03  8.370  < 2e-16 ***
## as.factor(RACE)2    -3.515e-01  7.037e-01  2.771e-01 -1.268   0.2047
## as.factor(RACE)3    -3.810e-02  9.626e-01  5.899e-01 -0.065   0.9485
## GRADES               3.648e-02  1.037e+00  2.195e-02  1.662   0.0965 .
## as.factor(MARRY)3    1.812e-01  1.199e+00  1.624e-01  1.116   0.2643
## as.factor(MARRY)4    2.890e-01  1.335e+00  2.519e-01  1.148   0.2511
## as.factor(MARRY)5    1.023e-01  1.108e+00  5.251e-01  0.195   0.8455
## as.factor(MARRY)6   -5.407e-01  5.824e-01  3.937e-01 -1.373   0.1697
## as.factor(MARRY)8    7.725e-02  1.080e+00  1.010e+00  0.077   0.9390
## as.factor(SIZE)2     5.776e-02  1.059e+00  2.565e-01  0.225   0.8218
## as.factor(SIZE)3     1.403e-01  1.151e+00  2.664e-01  0.527   0.5985
## as.factor(SIZE)4    -1.325e-02  9.868e-01  2.982e-01 -0.044   0.9646
## as.factor(SIZE)5    -3.917e-01  6.759e-01  5.549e-01 -0.706   0.4803
## as.factor(SIZE)6     5.865e-01  1.798e+00  4.052e-01  1.448   0.1477
## as.factor(SIZE)7    -1.499e-01  8.608e-01  3.995e-01 -0.375   0.7076
## as.factor(SIZE)8    -1.109e+01  1.534e-05  1.125e+03 -0.010   0.9921
## AVGSMK               2.910e-02  1.030e+00  4.956e-03  5.871 4.33e-09 ***
## as.factor(SMSA)2    -2.106e-01  8.101e-01  2.085e-01 -1.010   0.3125
## as.factor(SMSA)4    -4.605e-01  6.309e-01  3.278e-01 -1.405   0.1600
## URBAN                1.143e+01  9.180e+04  1.125e+03  0.010   0.9919
## HTN_REP              1.046e-01  1.110e+00  1.407e-01  0.744   0.4572
## RBC                  1.351e-01  1.145e+00  7.611e-02  1.776   0.0758 .
## DRMI                -3.575e-01  6.994e-01  3.875e-01 -0.922   0.3563
## DIAB                -2.536e-01  7.760e-01  3.007e-01 -0.843   0.3990
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## BOOZE                1.014e+00  9.859e-01    0.9734     1.057
## AGEYRS               1.063e+00  9.407e-01    1.0479     1.078
## as.factor(RACE)2     7.037e-01  1.421e+00    0.4088     1.211
## as.factor(RACE)3     9.626e-01  1.039e+00    0.3029     3.059
## GRADES               1.037e+00  9.642e-01    0.9935     1.083
## as.factor(MARRY)3    1.199e+00  8.342e-01    0.8720     1.648
## as.factor(MARRY)4    1.335e+00  7.490e-01    0.8150     2.187
## as.factor(MARRY)5    1.108e+00  9.027e-01    0.3958     3.100
```

```
## as.factor(MARRY)6 5.824e-01  1.717e+00    0.2692     1.260
## as.factor(MARRY)8 1.080e+00  9.257e-01    0.1493     7.819
## as.factor(SIZE)2  1.059e+00  9.439e-01    0.6409     1.751
## as.factor(SIZE)3  1.151e+00  8.691e-01    0.6826     1.939
## as.factor(SIZE)4  9.868e-01  1.013e+00    0.5501     1.770
## as.factor(SIZE)5  6.759e-01  1.480e+00    0.2278     2.006
## as.factor(SIZE)6  1.798e+00  5.563e-01    0.8125     3.977
## as.factor(SIZE)7  8.608e-01  1.162e+00    0.3934     1.884
## as.factor(SIZE)8  1.534e-05  6.520e+04    0.0000       Inf
## AVGSMK            1.030e+00  9.713e-01    1.0196     1.040
## as.factor(SMSA)2  8.101e-01  1.234e+00    0.5384     1.219
## as.factor(SMSA)4  6.309e-01  1.585e+00    0.3319     1.199
## URBAN             9.180e+04  1.089e-05    0.0000       Inf
## HTN_REP           1.110e+00  9.006e-01    0.8426     1.463
## RBC               1.145e+00  8.736e-01    0.9861     1.329
## DRMI              6.994e-01  1.430e+00    0.3272     1.495
## DIAB              7.760e-01  1.289e+00    0.4304     1.399
##
## Concordance= 0.716  (se = 0.016 )
## Likelihood ratio test= 138.9  on 25 df,   p=<2e-16
## Wald test            = 123.3  on 25 df,   p=6e-15
## Score (logrank) test = 131.1  on 25 df,   p=2e-16
```

```
z <- (coef(fit6_cont_f)[1] - coef(fit6_cont_m)[1])/sqrt(vcov(fit6_cont_f)[1,1] + vcov(fit6_cont_m)[1,1])
1 - pnorm(z)
```

```
##     BOOZE
## 0.7171777
```

Checking PH-assumption

```
# categorical version
fit3_q <- coxph(Surv(fu_time, cancer_death) ~ BOOZE_q + AGEYRS +
                    SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) + AVGSMK + as
                    HTN_REP + RBC + DRMI + DIAB, data = dta)
cox.zph(fit3_q)
```

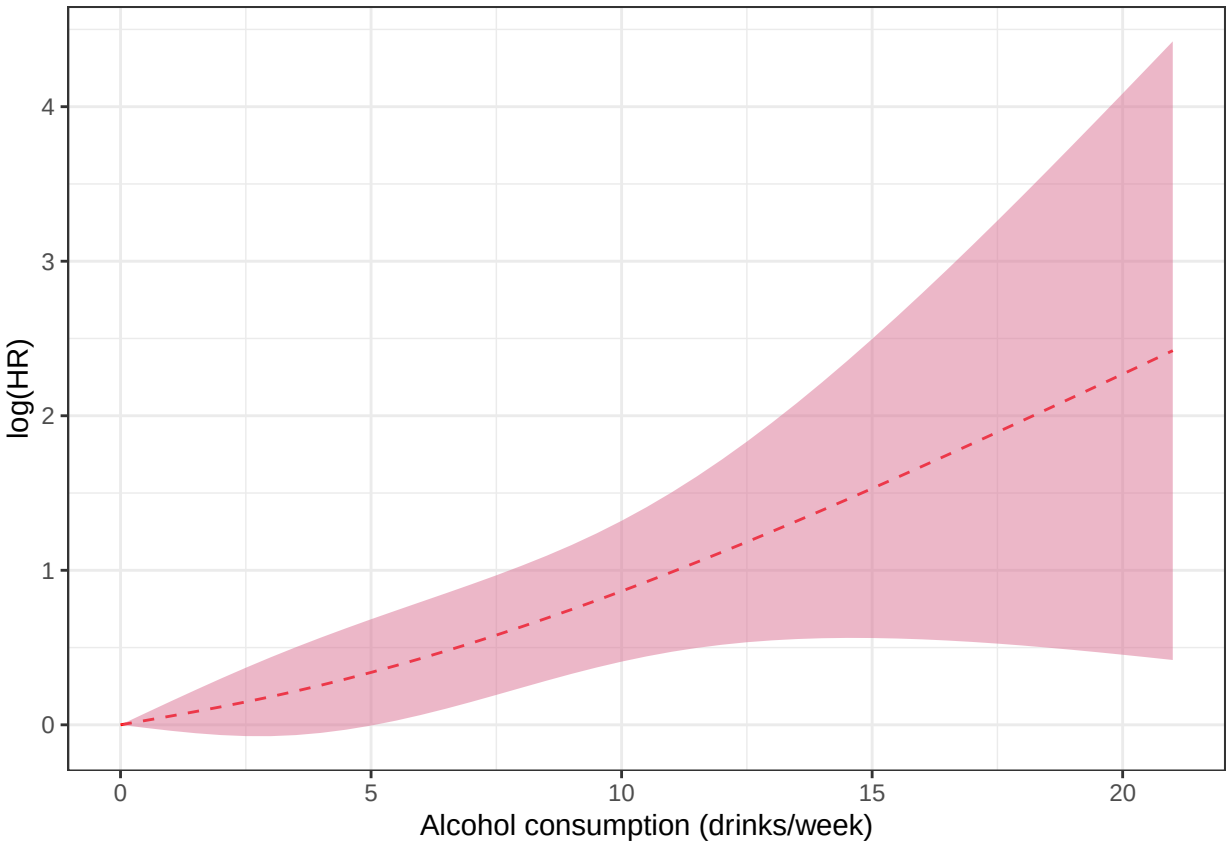
```
##                    chisq df     p
## BOOZE_q          2.92259  3 0.404
## AGEYRS           0.99688  1 0.318
## SEX              3.50873  1 0.061
## as.factor(RACE)  0.06268  2 0.969
## GRADES           2.06122  1 0.151
## as.factor(SIZE)  3.88985  7 0.792
## as.factor(MARRY) 5.51469  5 0.356
## AVGSMK           0.00156  1 0.968
## as.factor(SMSA)  0.40469  2 0.817
## URBAN            0.27417  1 0.601
## HTN_REP          0.25422  1 0.614
## RBC              0.40691  1 0.524
## DRMI             0.00317  1 0.955
## DIAB             3.01364  1 0.083
## GLOBAL          24.54144 28 0.653
```

```
# continuous version
fit3_con <- coxph(Surv(fu_time, cancer_death) ~ BOOZE + AGEYRS +
                      SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) + AVGSMK + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
cox.zph(fit3_con)
```

```
##                      chisq df     p
## BOOZE             5.34e-01  1 0.465
## AGEYRS            1.01e+00  1 0.315
## SEX               3.41e+00  1 0.065
## as.factor(RACE)   5.92e-02  2 0.971
## GRADES            2.05e+00  1 0.152
## as.factor(SIZE)   3.85e+00  7 0.797
## as.factor(MARRY)  5.46e+00  5 0.362
## AVGSMK            2.26e-05  1 0.996
## as.factor(SMSA)   4.17e-01  2 0.812
## URBAN             2.58e-01  1 0.611
## HTN_REP           2.58e-01  1 0.612
## RBC               3.77e-01  1 0.539
## DRMI              2.93e-03  1 0.957
## DIAB              3.05e+00  1 0.081
## GLOBAL            2.23e+01 26 0.672
```

Exploring Non-linearity

```
library(splines)
fit_nl <- coxph(Surv(fu_time, cancer_death) ~ ns(BOOZE, knots = 3) + AGEYRS +
                      SEX + as.factor(RACE) + GRADES + as.factor(SIZE) + as.factor(MARRY) + AVGSMK + as
                      HTN_REP + RBC + DRMI + DIAB, data = dta)
x <- seq(0,21, 0.5)
loghr <- ns(x, knots = 3) %*% coef(fit_nl)[c(1,2)]
vcov <- vcov(fit_nl)[c(1,2), c(1,2)]
var <- diag(ns(x, knots = 3) %*% vcov(fit_nl)[c(1,2), c(1,2)] %*% t(ns(x, knots = 3)))
se <- sqrt(var)
df <- tibble(x = x,
             y = loghr,
             ymin = loghr - 1.96*se,
             ymax = loghr + 1.96*se)
ggplot(df) +
    geom_line(aes(x = x, y = y), color = "red", linetype = "dashed")+
    geom_ribbon(aes(x = x, y = y, ymin = ymin, ymax = ymax),
                alpha = 0.5, fill = "#DB7093") +
    theme_bw() +
    labs(x = "Alcohol consumption (drinks/week)",
         y = "log(HR)")
```


log(HR)
0
1
2
3
4
0
5
10
15
20
Alcohol consumption (drinks/week)